Centrum voor Wiskunde en Informatica

REPORT*RAPPORT*

INS

Information Systems

INformation Systems

Towards making automated multimedia communicate more effectively: A multi-disciplinary approach

M.S. Cáceres

REPORT INS-E0410 NOVEMBER 2004

CWI is the National Research Institute for Mathematics and Computer Science. It is sponsored by the Netherlands Organization for Scientific Research (NWO).
CWI is a founding member of ERCIM, the European Research Consortium for Informatics and Mathematics.

CWI's research has a theme-oriented structure and is grouped into four clusters. Listed below are the names of the clusters and in parentheses their acronyms.

Probability, Networks and Algorithms (PNA)

Software Engineering (SEN)

Modelling, Analysis and Simulation (MAS)

Information Systems (INS)

P.O. Box 94079, 1090 GB Amsterdam (NL)
Kruislaan 413, 1098 SJ Amsterdam (NL)
Telephone +31 20 592 9333
Telefax +31 20 592 4199

ISSN 1386-3681

Towards Making Automated Multimedia Communicate More Effectively: A Multi-Disciplinary Approach

ABSTRACT
For over a decade, research has been conducted into artificial intelligence systems that have the ability to automatically generate multimedia presentations. Instead of drawing on discourses commonly involved in the creation of multimedia presentations, such as graphic design, the responsibility for insuring the communicative and aesthetic effectiveness of presentations generated by these systems has been restricted to a limited set of linguistic and print-centric discourses. As a result, users (and developers) are often disappointed that automatically-generated presentations lack much of the communicative coherence and multimodal aesthetic qualities of real-world multimedia. In this paper we show that defining a presentation's motivation for communication will determine the disciplines that should be involved in both its creation, and in the formulation of evaluation criteria for ensuring effective communicative and aesthetic outcomes. Such evaluation criteria, if implemented into one of these Intelligent Multimedia Presentation Systems has the potential to significantly improve the communicative efficiency of automatically-generated presentations.

1998 ACM Computing Classification System: H5.4 [Hypertext,Hypermedia]: Architectures, Navigation, User issues. I.7.2. [Document Preparation]: Hypertext/hypermedia, Multi/mixed media.
Keywords and Phrases: Multimedia semantics; intelligent multimedia presentation systems; automated presentation generation; semiotics; design; human factors
Note: Part of the research described here was funded by the Dutch national ToKeN2000/CHIME project.

Towards Making Automated Multimedia Communicate More Effectively: A Multi-Disciplinary Approach*

Marcos S. Cáceres
Department of Communication Design
Queensland University of Technology, Australia
marcos@datadriven.com.au

November 10, 2004

*This report is a revised version of the Honours Thesis with the same title, also by Marcos S. Cáceres, published by the Department of Communication Technology, Queensland University of Technology, Brisbane, Australia. This report was edited by Frank Nack (thesis supervisor) and Lynda Hardman.

Abstract

For over a decade, research has been conducted into artificial intelligence systems that have the ability to automatically generate multimedia presentations. Instead of drawing on discourses commonly involved in the creation of multimedia presentations, such as graphic design, the responsibility for insuring the communicative and aesthetic effectiveness of presentations generated by these systems has been restricted to a limited set of linguistic and print-centric discourses. As a result, users (and developers) are often disappointed that automatically-generated presentations lack much of the communicative coherence and multimodal aesthetic qualities of 'real-world' multimedia. In this paper we show that defining a presentation's motivation for communication will determine the disciplines that should be involved in both its creation, and in the formulation of evaluation criteria for ensuring effective communicative and aesthetic outcomes. Such evaluation criteria, if implemented into one of these *Intelligent Multimedia Presentation Systems* has the potential to significantly improve the communicative efficiency of automatically-generated presentations.

Contents

Chapter 1

Towards effective multimedia communication

The automatic generation of multimedia presentations has been a focus of multimedia research for over a decade. Research into the automated generation of multimedia presentations has resulted in a number of *Intelligent Multimedia Presentations Systems* (IMMPS), which are, essentially, knowledge-based systems that are able to generate multimedia presentations with minimal or no human intervention [6, 22, 73, 85, 97].

1.1 IMMPS research - motivation and problems

There are many commercially motivated reasons as to why researchers have been interested in developing IMMPS. André et al. [3, 6], Geurts et al. [45], and Roth and Hefley [85] (p15) argue that IMMPS have a vital role to play where an information domain is so large and dynamic that it would not be financially feasible or humanly possible to manually generate all possible multimedia presentations. Other researchers [86] argue that IMMPS could allow users with limited knowledge of the field of multimedia to easily produce multimedia presentations. More recently, van Ossenbruggen and Hardman [98] argue that, as we enter the era of mobile computing, there may be too many hardware/software combinations for systems that generate multimedia presentations by using templates to feasibly adjust to (see also [45]). Template-driven presentations are currently the most common means of generating multimedia on the World Wide Web [97]. Ideally, an IMMPS should adapt content by taking as input a user's profile, plus a device's profile, and adapting content for device independence and the user's accessibility needs – all this while retaining the presentation creator's intended message [3, 22, 99] (see also CC/PP [104] and WCAG 1.0 [102]). According to Kobsa et al. [58], there is a growing need for hypermedia systems that can handle such a degree of user/device adaptation and content tailoring. Kobsa et al. point to the economic benefits and consumer satisfaction that come from dynamic user modelling and content tailoring of web pages already seen on the World Wide Web.

Despite the various optimistic and forward-looking motivations for their development,

Figure 1.1: The sophistication of human designed interfaces, as in games, might influence the perception of a computer generated communication flow.

IMMPS have not proliferated into the commercial mainstream or widely onto the World Wide Web. Several reasons have been proposed as to why this limited success might be. For instance, van Ossenbruggen et al. [97] have argued that web-based adaptive multimedia requires different document and presentation abstractions from those currently used to publish documents on the World Wide Web, which are based on text-flow. Adaptive multimedia formatting models cannot be based on text flow and so the development of IMMP-like systems for the Web is proving increasingly difficult. As a result, Ossenbruggen and Hardman argue that new document and presentation abstractions may be needed to specifically support adaptive hypermedia [98].

In another instance, Nack [80] has criticised the multimedia research community for being out of touch with production and consumption of 'real world' multimedia, in particular the entertainment media of computer games. At first, it may seem unjustified to compare research level multimedia to today's multi-million dollar computer games productions. Nack, and also Schmitz [88], however, make the point that a user's exposure to computer-based entertainment media can shape their expectations for computer-based multimedia they experience. It also is apparent that the entertainment media has long ago departed from the passive 2D metaphor to an active 2D+ or 3D representation, while IMMPS generally have not[1].

To illustrate, the sophisticated multidimensional visual and auditory effects, as well as user interaction, experienced in computer games such as *Neverwinter Nights* [8], shown in Figure 1.1 on the left, can actively shape certain expectations users have about multimedia. IMMPS-generated presentations, such as the Cuypers-generated presentation on the right, lack the graphical sophistication through which computer games are able to coordinate multiple media and interaction into an active communication process. Most IMMPS only passively present information so lack the ability to engage users in an active communication process [5].

In Figure 1.2 we propose that, although entertainment and research industries employ

[1]There are, however, some notable exceptions. The PPP (Personalized Plan-based Presenter) system [3, 6], for instance, makes use of animated characters. PPP is discussed on in chapter 2.

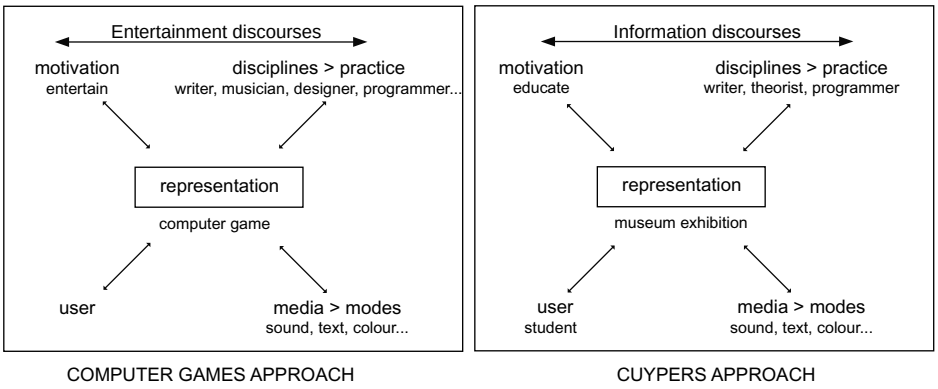

Figure 1.2: The motivations for communication dictates the disciplines involved in generating multimedia representations.

the same media and modes for expression, the *motivations* of the producers and consumers of a multimedia presentation differ. The motivation of the producers of multimedia, be it, for example, to educate or to entertain (or both simultaneously), will determine the various disciplines and discourses that designers will draw on when collaborating to generate a multimedia presentation. The entertainment industry, in developing games, has historically involved professionals from cognate artistic disciplines including creative writers, designers, musicians, and programmers. We argue, however, that the IMMPS research approach has generally been unrepresentative of the experts and knowledge from such diverse disciplines, whose practitioners commonly and effectively develop media for multimedia presentations. This under representation of potentially cognate disciplines limits the discourses surrounding IMMPS research and thus the potential to produce IMMPS that generate communicatively effective multimedia presentations. In addition, IMMPS research still struggles to capture the diverse motivations of users which could potentially utilize their presentations.

We therefore propose that, as a starting point, IMMPS need to identify and simulate the discourses and practices of disciplines that would normally be involved in producing media elements for multimedia presentations. Furthermore, given that IMMPS primarily output computer-based data *representations*, we argue that IMMPS may often need to model themselves around disciplines that explicitly deal with *representational theory*. That is, disciplines that study how and why representations, such as multimedia presentations, are produced and interpreted [28].

Attempts to integrate computer science with representational theory have previously been made by various researchers, including Laurel [64] and Nack and Hardman [81]. Laurel argues that dramatic theory and particular concepts from the theatrical domain can be applied to improve human-computer interface design and computer-based interactive methods (p1-33). Laurel also notes that it has been common practice amongst the computer science field as a whole to overlook disciplines that deal with representation theory,

such as theatre and graphic design, and instead opt for disciplines that are perceived as being more scientifically "serious"(p22), such as usability or the field human-computer interaction (HCI).

Nack and Hardman [81], on the other hand, propose semiotics as a way of integrating representational theory with IMMPS. It would seem that semiotics is a logical choice for the task, since, as Chandler [29] states, "semiotics involves studying representations and the processes involved in representational practices, and to semioticians, 'reality' always involves representation." We argue, however, that relying solely on dramatic theory or semiotic theory may not be enough to bridge the gap between IMMPS research and the creative practices that produce the media elements, or *content*, for multimedia. Again, in our view, the disciplines to include in the creation of a multimedia presentation should be primarily dictated by the motivation for communicating with a user.

We argue that IMMPS researchers have not acknowledged in their research the disciplines that explicitly deal with process of representation[2]. Instead, over the past decade, many IMMPS researchers have relied on, and often extended, the document abstraction model. This is not astonishing as this model is still widely applicable for a great number of multimedia presentations [96](p76-8). For instance, As Berners-Lee explains [20](p45), the most widely used hypermedia publishing standard on the Internet, the HyperText Markup Language (HTML) [103], was inspired from the document abstractions used in the Standard Generalised Mark-up Language (SGML) [53] – initially a print media publishing standard [96](p29). In fact, HTML was primarily proposed by Berners-Lee as a system to facilitate online documentation and hence was built around textual (print) presentation and structural requirements [20](p22)[3].

Irrespective of the origins of HTML, what is now generally professionally published on the web no longer looks or functions like print media documents [96](p10). Electronic documents nowadays contain digital videos, dynamic menus, embedded Flash [66] and Shockwave [65] movies, executable source code, and other dynamic media elements that supersede print-media as an active communication medium in many ways[4] [96](p2). In addition, even through the short time the web has existed, graphic/web designers have

[2]Here are a few works that actually address the disciplines we generally miss [100, 24, 78, 79].

[3]See in particular the elements and attributes of the HTML 2.0 specification document (1995), available at http://www.w3.org/MarkUp/html-spec/

[4]This is not to say that using proprietary multimedia technologies, such as Flash and Shockwave, on web-documents does not come with its own set of problems. Advocates of the Semantic Web and usability experts argue that proprietary formats such as Flash and Shockwave pose considerable usability, internationalization, and information processing problems on the web [82], [20](p180-1). (These issues are currently being addressed by Macromedia, the creator of the Flash and Shockwave formats [82]). New multimedia recommendations, such as SMIL [106, 101], XHTML [107], and SVG [42], are set to overcome these limitations by using an open XML framework, which natively supports internationalization, as well as by their ability to support various XML-based metadata standards/recommendations [17, 2, 55] and technologies. Using metadata and related processing technologies will allow multimedia to become machine processable, which is a prerequisite for the Semantic Web [97]. The Cuypers system is already overcoming the issues relating to proprietary multimedia formats as it natively built upon XML technologies and recommendations. See also our discussion of the CUYPERS system in chapter 2.

devised web-specific design ideologies which build upon the solid foundations of various traditional (modernist) design ideologies (such as Bauhaus, De Stijl, Swiss, etc.). According to Engholm [41], to create these new design ideologies, designers have, for some time now, been extending their discourses with progress made in the study of human-computer interaction, usability [41], and more recently accessibility [102].

The inclusion of rich media elements and these new design ideologies allows web designers the ability to easily *break* the document metaphor by, amongst other things, integrating interaction, adaptation, and temporal structuring of content [86], as well as creating new ways of structuring arguments in hypertext which differ significantly to those of traditional media [25, 62](p123) – such as print-media.

1.2 IMMPS - an approach from a different point of view

In this report we argue that the domain and motivation for an act of multimedia communication determine the various disciplines that should be involved in the creation of an effective multimedia presentation. For instance, if the motivation is to educate a person about artefacts from a museum domain through various digital media, then, museum educators, writers, graphic designers, sound designers, and so on, should be involved in constructing a multimedia presentation. However, if the motivation and domain changes, then so will many of the disciplines involved in the construction of a multimedia presentation. Furthermore, we argue that no matter which disciplines are involved in the design of a multimedia presentation, they will all impact on four particular processes of multimedia communication. We refer to these four processes as *generation*, *evaluation*, *collaboration*, and *signification*.

Firstly, in the process of *generation*, designers draw on their knowledge from their respective disciplines to design media elements for a multimedia presentation. For instance, a graphic designer will use skills from the discipline of graphic design to produce a particular graphic for a presentation.

Secondly, in the process of *evaluation*, these designers apply critical knowledge from their respective disciplines to evaluate a media element in a multimedia presentation in terms of quality of media and effectiveness of the communication design. Again, a graphic designer will draw on their understanding of fundamental design elements and design principles to decide whether their generated graphic meets the communicative criteria of the presentation [63, 108].

Thirdly, in the process of *collaboration*, designers from various disciplines rely on shared discourses and descriptive framework to effectively work together and evaluate the overall communicative quality of a generated presentation. For instance, the graphic designer, the writer and an audio designer may get together and critically evaluate the integration of their respective media elements in the overall presentation. Once the designers have finished putting together a multimedia presentation the final process may occur.

Fourthly, in the process of *signification*, a user derives meaning from a multimedia presentation based on perception, various cultural factors, and, to a lesser extent, their understanding of the processes of generation, evaluation, and collaboration.

To show the value in understanding multimedia in terms of the four processes discussed above, this report will show the following:

1. Multimedia is not simply the spatial-temporal amalgamation of different media, but that each medium comes with its own complex combinations of established disciplines and discourses that impact upon the generation, evaluation, collaboration of those who produce multimedia presentations. These disciplines may also impact how a user derives meaning from a presentation through the process of signification [61](p8, 67-85).

2. Given that each medium comes with its own established disciplines and discourses, multimedia designers (human or IMMPS) require a shared discourse in order to generate, evaluate, and collaborate in designing multimedia presentations.

Through this understanding of how multimedia is developed and communicated, our central objective for this report is to improve the communication efficiency of presentations generated by the Cuypers system through a descriptive framework that incorporates a number of established theories and discourses described mainly in Chapter 3 and 4. Within our work we draw upon the discourses of museum education, exhibition design, graphic design, hypermedia theory, and semiotics.

By "descriptive framework" we mean a set of description schemata and related manipulation rules that facilitate "sufficient expressive power to capture relevant knowledge of the design world under consideration and to allow complete and precise formulation of the requirements for a solution of a design problem" [76](p83). For the purposes of this report, we will substitute the words "design world" with multimedia presentation. The basic elements for our framework will be prepared from the vantage point of IMMPS research [85, 22], with the intent of representing artworks and concepts found in the domain of musea for the fine arts. When applied in the context of IMMPS, the framework will be shown to be useful for recursively looping through the processes of generation and evaluation. These recursive loops may result in multimedia presentations that may communicate more effective than those generated by current linguistic-based [68, 73] and print-based [83, 86] approaches commonly used in IMMPS research today.

The potential for our framework to communicate more effectively than current approaches is realized through the inclusion of disciplines that have remained mostly underrepresented, for one reason or another, in IMMPS research. We argue that these underrepresented disciplines, such as graphic design and sound design, have traditionally been included when manually constructing multimedia presentations [9] and thus should also be part of the common discourse of IMMPS researchers.

1.3 Report overview

Chapter 1 introduced the motivation of IMMPS research and outlined the direction our work takes to improve aspects of IMMPS research. We initially proposed that the motivation behind an act of multimedia communication should determine the disciplines which should collaborate to generate and evaluate any multimedia presentation. In addition, we argued that all multimedia presentations are a form of representation and that this has particular implications for how multimedia is generated, evaluated, collaborated, as well as having an impact on what the presentation might signify to a user. As a result we propose a "descriptive framework", namely a set of description schemata and related manipulation rules that facilitate the capture of relevant knowledge of the design world under consideration and to allow complete and precise formulation of the requirements for solution of a design problem.

Chapter 2 looks at a number of already established IMMPSs. The aim of this analysis is to determine the strengths and weaknessis of the various approaches. The results of the analysis help us to develop our descriptive framework.

Since our aim is to improve the communication efficiency of the Cuypers system, which operates with data from a museum domain, Chapter 3 examines how various disciplines collaborate within a museum to put together an exhibition, which we will regard as a physical multimedia presentation. Such an analysis of a museum exhibition will serve to show that the collaborative unity between disciplines is a key factor to achieving common communicative goals.

Chapter 4 applies the results of Chapter 3 with respect to Cuypers' ability to generate multimedia presentations. We will show that the processes of evaluation and (re)generation will yield a multimedia presentation that may communicate more effectively than presentations currently generated by Cuypers. We will also show that our descriptive framework can be used to enhance many parts of the Cuypers system architecture. We will not, however, be formalising or implementing the suggested improvements into the Cuypers architecture as part of this research. As we will show, further research beyond the scope of this report, will be needed to achieve full implementation of our approach. This means that some sort of implementation of our approach, backed up with user testing, would naturally be required to validate our hypothesis. We will argue, however, that future application of our approach to Cuypers will demonstrate the advantages that a dynamic multidisciplinary approach can bring to the domain of IMMPS research in general and to the Cuypers system in particular.

Lastly, Chapter 5 concludes the report and gives a summary of the main contributions and recommendations made by this research.

Chapter 2

IMMPS research - an overview

In the previous chapter we outlined general problems of IMMPS research. In this chapter we aim to explain in a bit more detail why common IMMPS research limited itself by applying the document abstraction model as the basis of its discourse mechanisms. Our argument builds on a few key examples of IMMPS systems, namely the TEXPLAN [73] system, the WIP and PPP systems [3], the Cuypers system [56] and the Terminal Time System [70, 71, 36].

2.1 TEXPLAN

TEXPLAN is a reasoning system for generating text-based natural language explanations in any given domain. Given an object-oriented domain model, such as a street map, TEXPLAN can identify an entity, compare different entities, and explain a process by coordinating text, and simple graphical shapes through a graphical representation [73](p61-62). For instance, a user may query TEXPLAN as to how to get to a particular location on a street map relative to the user's current location. TEXPLAN would respond to the user's query by synchronising text and simple graphics over time that explains to the user how to reach their desired destination.

For the TEXPLAN system, Maybury [73] generalises upon a long tradition of computational linguistics, which stem from Searle's formalisation of (J. L. Austin's) speech acts [89], to propose a resulting set of communicative acts. According to Eagleton [38](p102-3), Austin's speech act theory demonstrates that spoken language's function is not merely to describe reality, but instead it is always *perfomative*. That is, language is always used to get something done. For instance, one performs an 'illocutionary' act by affirming something *in* saying 'you are under arrest'. According to Maybury, communicative acts theory covers a broader range of communication than speech act theory, including "rhetorical, linguistic, and graphical acts as well as non-linguistic auditory acts (e.g. snap, ring) and physical acts (e.g. gestures)" [73]. Similar to speech act theory, communicative acts nominally classify the rhetorical functions of media elements used in a multimedia discourse in terms of speech acts. A blinking circle, for instance, used in a presentation to attract a

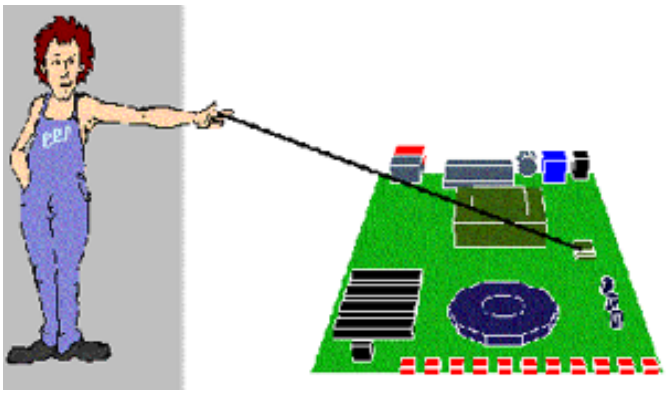

Figure 2.1: The PPP Persona explains to a user how to fix a modem.

users attention would be classified as a 'deictic', or indicating, act [73]. The following two systems also make use of communicative acts as a model for communication.

2.2 WIP and PPP

The WIP [4] and PPP (Personalized Plan-based Presenter) systems [3, 6] are two further examples of IMMPS that use language-centric theories. André et al. [6] state that WIP and PPP's approach to multimedia generation and evaluation is also through communicative acts, as well as through the extension of the linguistic theory known as Rhetorical Structure Theory [68] (RST). RST is a formal vocabulary used to describe a specific set of rhetorical relations that occur in (spoken or written) texts [68]. For instance, RST can be used to describe elaborations, motivations, evidence, concessions, and so on, in a written text. Written texts described in terms of RST result in a computable structure of nucleus and satellite relations that semantically represent the argumentation structure across linear text spans [68].

WIP is a fully automated IMMPS, meaning that it computationally generates images, animations, text and speech on the fly. WIP's motivation is to generate goal-driven presentations which can instruct users in maintaining and repairing technical devices such as a modem or a lawnmower [6]. André et al. note that "one limitation of the WIP system is that it merely generates the material to be presented such as text/picture combinations or animation sequences. It does not plan *when* and *how* to present this material to a particular user." This means that WIP is unable to utilise the temporal coordination of media and modes as a means of communication in a multimedia presentation. Another limitation of the WIP system is that it was not designed to deal with user interaction, thus, as André and Rist [5] point out, only the passive viewing of the generated presentations is supported.

To overcome the above two limitations, André et al. developed the PPP system. PPP is essentially an interactive hypermedia extension of the WIP system [5]. Similarly to WIP,

PPP generates almost all media elements on the fly. In PPP, however, this media is now temporally coordinated and explanations are partially delivered with the assistance of an animated character – the PPP Persona [7], shown in Figure 2.1. The Persona supports a wide range of pre-rendered gestures and is also able use synthesized speech as a mode of communication. PPP also supports hypermedia navigation structures through interactive elements and pop-up menus. A limitation of both WIP and PPP is their highly constrained domain of application: they are principally suited for generating presentations where the mode of communication is expository and where the domain model is closed.

2.3 Cuypers

The Cuypers [97] experimental system is one of the more recent IMMPS under development. Aside from being able to generate multimedia presentations with minimal human intervention, the Cuypers system is also designed as a tool to enable research into a range of issues that arise when multimedia documents are modelled and generated with XML and Semantic Web technologies (see [45, 69, 97, 98] for details). Cuypers is currently attempting to generate multimedia presentations that are able to *inform* a user about concepts and artefacts from the cultural heritage domain, namely the domain of musea for the fine arts. In order for Cuypers to achieve its information goal, it intelligently organises media elements from a multimedia database into a predefined rhetorical structure, and then present these elements to a user through the aid of certain communicative devices in the form of a multimedia presentation. This generation architecture, which we will discuss in greater detail later in this chapter, is shown in Figure 2.2.

The motivation for conducting research into Cuypers has been primarily driven by issues relating to modelling multimedia through XML. Given that Cuypers can generate XML-based presentations, research is now needed into how to generate presentations that communicate more effectively with a user. This is precisely the research problem this paper begins to address.

Cuypers utilises both RST and communicative acts [97] to generate presentations about artefacts from the fine arts. In Cuypers, the rhetorical structure of a multimedia presentation is described by its creators in terms of RST relations [45]. At runtime, Cuypers maps these RST relations to a number of distinct *communicative devices* [86], which Cuypers uses to spatially and temporally lay out a multimedia presentation. Rutledge et al. [86] define communicative devices as spatial temporal design patterns "of communicative acts that use or express a particular rhetoric". In Figure 2.3, for instance, Cuypers has used the *adjacency* communicative device to place the image next to the text [44](p11).

At the time of writing, Cuypers relies on a presentation structure which the creators of Cuypers have pre-specified as its discourse framework[1]. As a result, a Cuypers-generated

[1] Attempts are currently being made to rectify this problem by segmenting the Cuypers' architecture into specialised modules. It is envisioned that one such module could be used to specifically handle particular educational approaches.

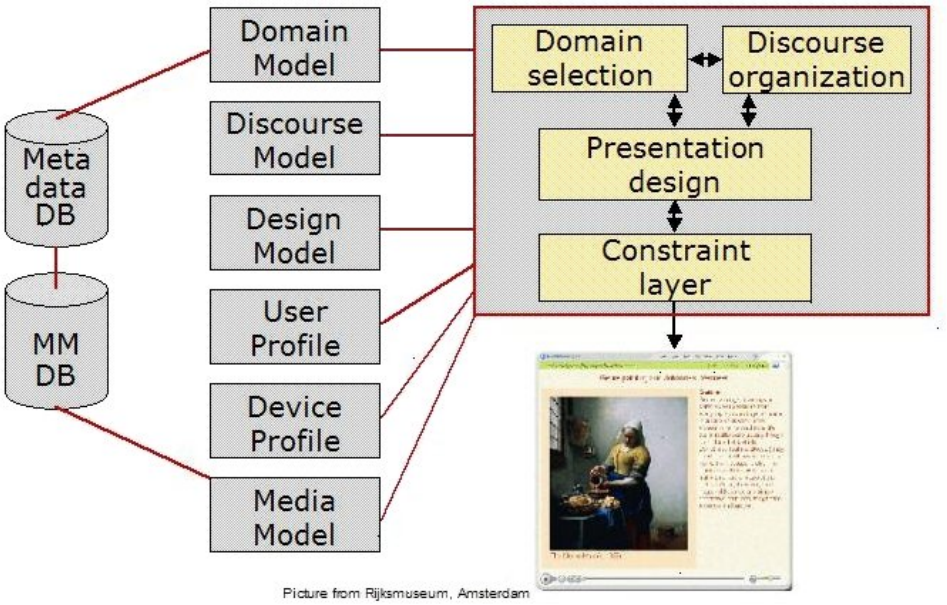

Figure 2.2: A view on the Cuypers architecture.

Figure 2.3: A Cuypers generated multimedia presentation, representing data from a museum domain.

presentations does weakly exhibit the principles of self-exploration, humour, or play – principals increasingly used in the museum domain [84](p20)[2].

We argue, therefore, that there are two problems with Cuypers. Firstly, the types of metaphorical representations result in multimedia presentations that always look and function like a printed document. Secondly, as a consequence of the first, these metaphorical representations limit how users interpret and interact with a multimedia presentation. In other words, metaphorical representations, such as a bookshelf, set the *discursive framework* (the world view) of the presentation to print-like documents. This would not be a problem if Cuypers were trying to simply represent static printed documents, nor would it be a problem if the content was suited to be shown as a print media representation. However, Cuypers tries to actively represent content from a museum domain by structuring media over time. Rutledge et al. note that, "hypermedia", of which multimedia is a special case that is temporally linear [77] [48](p2-3), "is distinguished from the other media", such as print media, "by... interaction, adaptability, temporal spatial structure"(p3). If this distinction between hypermedia and "other media" is valid, then, we will argue, *passive* print-based metaphors may not be ideal for representing certain discourses in a wide range of *active* multimedia representations. These kinds of limitations of metaphorical interface representations are well-known in the human-computer interaction community [64](p2-8).

2.4 TERMINAL TIME

Terminal Time [70, 71, 36] is an IMMPS performative artwork that generates documentary films based on audible audience responses. Terminal Time is a cinematic experience, generally exhibited to audiences within the context of a movie theatre. Three times throughout a thirty-minute screening, the generated documentary is stopped and the audience is asked to audibly respond to a set of multiple-choice questions with responses such as "A. life was better in the time of my grandparents" or "B. life is good and keeps getting better every day" [71] (see also Figure 2.4). The audience usually responds to these questions by clapping and shouting. A computer with a directional microphone measures the loudness of the audience response. Terminal Time uses the loudest response to influence the ideological theme (gender, race, technology, class, religion, etc.), and narrative arc (the ideological progress or decline the narrative will take)to generate the documentary.

From all the IMMPS discussed so far, Terminal Time is perhaps the system where "communicative goals" are used closest to the meaning as it is used in IMMPS research domain: "the purpose or intentions for communicating information or the task to be performed by the user of the information" [85](p15), as it coordinates various *means of representation* to achieve signification.

As described by Domike [36], Terminal Time's primary communicative goal is to challenge the authoritarian myth and objectivity generally attributed to the documentary

[2]See, for instance, the *constructivist* approach [50] –"how the learner constructs meaning out of experience"– which is commonly used in museum education. See also Dufresne-Tasse [37] in relation to andragogy in museums.

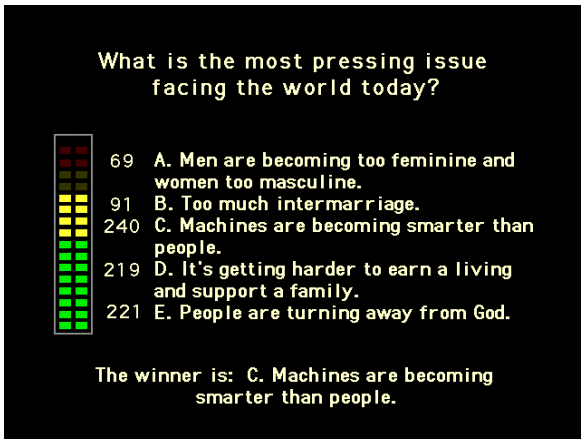

Figure 2.4: A questionnaire screen generated by Terminal Time.

genre [43](p165-72). Terminal Time challenges this mythology by ideologically selecting historical events, reinterpreting them through an exaggerated ideological perspective, and re-presenting these historical events through the stylistic conventions of the documentary genre. Its secondary communicative goal is to show how human-computer interaction can be value-laden as opposed to value-free. That is, this second goal is also realized through Terminal Time's ideologically-exaggerated representation of a finite set of historical events, which, overall, its creators hope will evoke an uncomfortable and reflective emotional response in members of the audience [36].

Terminal Time is composed of a number of distinct system components: a knowledge-base of historical events, a multimedia database that contains edited video sequences about the historic events and music sequences that can establish particular emotional states, ideological-goal trees, rhetorical devices, and a rule-based natural language generator to facilitate the communication with the audience.

The knowledge-base contains an ontological classification of 134 historical events that have occurred over the last one-thousand years. The system uses these historical events as the means of abstractly representing the narrative that best fits the audience's selected ideological themes. Exactly which historical events are selected from the knowledge-base depends on the ideological-goal trees. These trees act as deterministic narrative fragments. The trees are said to be deterministic because the creators have pre-specified the ideological attributes that a historical event must exhibit to be included as support for an ideological goal. For instance, if the audience's responses reflect a "hardcore anti-religion rationalist" ideological theme, then an ideological-goal tree will make the system select a historical event that meets an ideological precondition that "*religion leads to evil*". The First Crusades would be such an event [36]. The biggest achievement of Terminal Time is that it can establish a variety of story lines on the fly.

Once the system has established the story line of the presentation, it selects the relevant video clips and music sequences from the multimedia database. The video clips, which the

creators have pre-edited into thirty-second modules, include complete sequences of content that are designed in a way that they can be easily concatenated by use of filmic transitions such as cuts, fades and dissolves [36]. During video production the creators made sure that the style of each video clip follows the stylistic conventions of the documentary genre. Thus, at this stage all that Terminal Time does is stringing together manually edited video clips in the sequential order provided by the story engine. This means, however, that Terminal Time explicitly relies on the way that the system creators have edited the video clips to be able to present the story.

Additionally, the system selects music[3] from the multimedia database to reflect the rhetoric function of each sequence. The way Terminal Time achieves that is to relate parts of the ideological-goal tree via a "mood attribute" with music sequences that provide the same attribute within their content description. According to Domike [36], music in the multimedia database is annotated as 'happy', 'sad', etc. and actively used by the ideological-goal trees "as a mood elevator or depressor, depending on the arc of the narrative." In short, the system implicitly selects the music as a persuasive device with a clear rhetorical function in the narrative [36]. The music's rhetorical function is said to be implicit to the system because its inclusion depends on the ideological-goal tree explicitly stating what 'mood' a piece of music should have.

Finally, the rule-based natural language generator pieces together textual equivalents of the ideological-goal tree. The texts shown to the audience at the beginning of a sequence as well as at the end (here the alternative path directions are included) are composed of a mixture of text fragments that the creators refer to as rhetorical devices[4]. An example of a rhetorical device is: "but progress doesn't always yield success." Aside from their rhetorical function, these rhetorical devices, of which the system contains 281 [70], is to verbally express the narrative arc of the ideological-goal trees in the form of natural language. The system synthesises the rhetorical devices to speech at runtime, which are then used to narrate the documentary on the expression plane.

The advantage of Terminal Time is that it applies a rhetoric canon to generate a biased cinematic experience in a particular style (documentary) for mass audiences. Important is that the system may select the same historical event, and hence the same video clip, to either visually support or contradict an ideological theme. From a semiotic perspective, this means that the imagery of a video clip by itself, that is, without the accompanying music and speech, would be quite open for interpretation by the audience. It is only through the composition of the music and synthesized voice with the video clip that a video clip's meaning is understood by the audience [14, 15, 43]. In other words, the way that media and modes are co-ordinated has an effect on the overall semantics of any given composition, which validates the semiotic standpoint that every composition needs careful consideration if a particular communicative goal is to be met [14, 39, 43, 60, 61].

A problem with Terminal Time is that the rhetoric rules for generating the argument are embedded in the material and the rules for its organisation. This makes it impossible

[3]Film makers always use music to serve a rhetorical function [40](p157-216), [49](p38), [26]

[4]These texts are edited text pieces provided by the creators.

to use either the material or the rhetoric engine for another application. Thus, Terminal Time is, as other knowledge intensive but closed AI applications from the nineties, content driven.

The Cuypers system, on the other hand, is structure driven, meaning that the particular rhetoric rules as well as stylistic presentation strategies are made explicit. In this way the approach taken in the Cuypers system potentially facilitates the generation of complex discourses on-the-fly. The risk embedded in the Cuypers is, however, that the system cannot rely on "perfect material", as provided in Terminal Time, which might result in a presentation that triggers undesirable associations in the user and thus creates misdirecting meanings.

2.5 Summary

In this chapter we discussed the strengths and weaknesses of implemented IMMPS. As we conclude this overview on IMMPS research, it is important to note that of the four IMMPSs discussed, only PPP and Terminal Time are able to effectively handle human-computer interaction and hypermedia structures within a generated presentation[5]. It needs to be mentioned, though, that the Cuypers system follows a different approach to the other three systems described, as it does not rely on presentational material predefined for the particular application but provides explicit mechanisms to handle large knowledge spaces, such as the web. In that respect it is difficult to effectively compare the different approaches.

As we stated in chapter 1, the primary objective of this thesis is to improve the communication efficiency of the Cuypers IMMPS. In chapter 3, we will step aside from the IMMPS domain and investigate a physical museum. Doing so will allow us to examine how multiple disciplines collaborate together to create a museum exhibition, which we will regard as similar to generating a multimedia presentation. From investigating the museum domain, we will be able to derive the basic elements for our descriptive framework. We will then, in chapter 4, show how these basic elements can be used to enhance the communicative efficiency of a multimedia presentation.

[5]Martinez [69](p45-7) implemented some linear hyperlinking components into Cuypers. However, further research is still needed to make Cuypers' presentations truly interactive.

Chapter 3

A museum exhibition as multimodal communication - an analysis

The assembly and presentation of a museum exhibition is without doubt a multidisciplinary activity [19](p33). A whole range of creative professionals, including museum directors, curators, exhibition designers, graphic designers, conservators, security officers, education officers, editors (writers), production staff, maintenance staff and external consultants, plan methodically how visitors experience an exhibition [19](p78-9), [57](p38), [46, 93]. The process of communication through exhibiting may include, amongst other things, physically re-arranging the layout of rooms, strategically placing artefacts at particular distances from the visitor, planning at what pace visitors traverse through a particular space, as well as designing the descriptive labels, audio, textures, and even smells, that an audience will experience throughout an exhibition space [19]. In doing so, Trant [93] notes, collaborating professionals are "challenging us with new ideas about [artefacts'] meaning and makeup and offering new interpretations and juxtapositions."

The motivation behind exhibitions, however, is not simply to assert facts about artefacts [84]. In most cases, there is clear motivation to educate and entertain visitors through the incorporation and coordination of various media and modes that engage the visitor through their sensory *modalities*[1] [84](p15-45), [19](p64). That is, the sensory channels through which a person perceives the world.

In this chapter, we attempt to understand how professionals apply knowledge from curatorship, exhibition design, graphic design and narrative theory to coordinate various media and their modes to reach a clear communicative goal – the presentation and communication of an exhibition. According to Kräutler [59](p59), semiotic theory has influenced museums to move beyond passive mechanical transmission of information to the public. In addition, he explains that "research regarding museums as socio-semiotic phenomena can

[1]According to Mitchell [76](p2), Aristotle identified five sensory modalities: visual, auditory, tactile, olfactory, and gustatory. Mitchell notes that, today, kinaesthetic experiences are commonly appended to this list. We are aware that authors in IMMPS research use the term modalities differently; what IMMPS researchers [22, 72, 4] call modalities, is called *modes* in this paper but also by other researchers [19, 29, 60, 61].

yield valuable tools to improve and enhance the pertinence and meaning of their public-oriented activities." So, we will also rely on semiotic theory to examine the process of representation as it occurs in the museum domain.

We divide our discussion in two parts. First, we investigate environmental sign systems, such as maps and posters, that are commonly found throughout museum exhibitions but do not contribute to an exhibition's narrative. These types of artefacts provide the means to discuss the basic elements required for the process by which meaning is established (semiosis). The artefact we build our discussion on is a poster. The discussion of the first part provides a detailed description of elements and related processes applied to presentation forms as being produced by the Cuypers system.

The second part of our investigation extends the findings of the first part by analysing what is required to extend the single event of an environmental sign system to a sequence of events *re*-presented from a particular point of view through media and modes over time. This part of the discussion helps us to establish a framework for parts of the Cuypers system that are either under development or still in a design phase, such as the discourse model. As the case for this part of the discussion, we use the *Van Gogh Gauguin* exhibition, which was held at the Van Gogh Museum in Amsterdam between 9th of February and 2nd of June, 2002.

3.1 Environmental signs - communication through composition

Environmental signs, such as maps and posters, are commonly found throughout museum exhibitions. Environmental signs are multifunctional; rather than directly contributing to an exhibition's narrative, these signs help to geographically orient visitors by suggesting the direction in which to traverse through an exhibition [19](p99). In other words, this type of sign system represents a static composition with a limited and mostly determined meaning.

In fact, environmental signs represent a semantic system, in which a *sign*[2] is understood as any element perceived implicitly or explicitly to have a communicative intent; that is, anything that an entity perceives to *stand for* something other than itself [39](p48). A semantic system we understand as a particular modality-based structuring of content that an audience understands [39](p48-150), [43](p64-84). Further on we use the word *code* instead of semantic system.

In this section we examine a poster from the *Van Gogh Gauguin* exhibition to discuss how collaborating professionals make use of codes and signs as devices to reach a number of intertwined communicative goals. At the end of the analysis we aim to have established a first set of elements and mechanisms for our descriptive framework.

[2]From a Saussurian semiotic standpoint, for something to be a sign, it must be composed of both a signifier and a signified [92](p27-31). The signifier is what one perceives as a sign, and the signified is conceptually what the sign means.

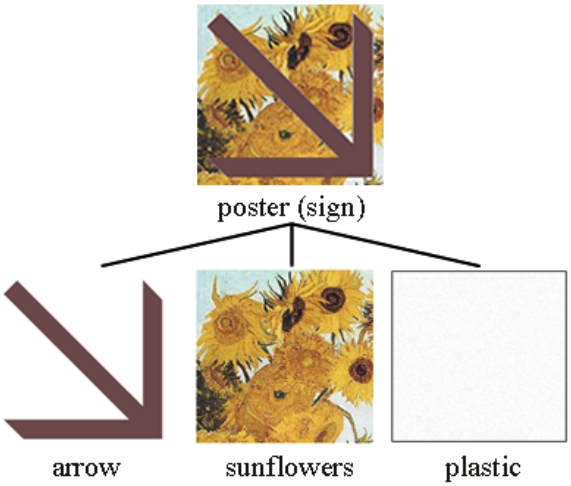

Figure 3.1: The poster (a sign) is composed of three signs: the arrow, the sunflowers, and the plastic.

3.1.1 The poster

Upon entering the *Van Gogh Gauguin* exhibition, a visitor is presented with a large (approximately 1.5m in diameter) colourful plastic poster resembling the one shown in Figure 3.1. Printed in the foreground of this poster is a large brown arrow, which is both centred and slightly smaller in size than the poster. Printed in the background of the poster is a zoomed-in image of colourful sunflowers. This poster is composed of three signs:

1. The brown arrow centred in the foreground, which indicates a direction to the visitor.
2. The colourful sunflowers in the background, which relates semantically to the *Van Gogh Gauguin* exhibition.
3. And, the plastic material on which the other two signs are printed, which indicates to the visitor the utilitarian function the sign.

A visual decomposition of the poster into signs is shown in the lower part of Figure 3.1. In the following sections we discuss the individual function of each of these signs, in the order arrow, sunflowers and plastic material. We conclude with a section that analyses the interplay between the three signs.

3.1.2 The arrow sign

Of the three signs that compose the poster, the arrow is as Kress and Van Leeuwen [60](p212) would describe, the most *salient* sign.

Salience "The degree to which an element draws attention to itself, due to its size, its place in the foreground or its overlapping of other elements, its colour, its tonal values, its sharpness or definition, and other features" [60](p225). In addition, Chandler [27] notes that salience may also depend on genre. He argues that a person may draw from their previous knowledge of texts within the same genre "to sort salient from non-salient narrative information in an individual text."

In the poster the arrow is made salient through the contrast of the dark-brown colour of the arrow against the lighter yellows and blues of the sunflowers (boldness of the arrow), the contrast of the textural plainness of the arrow against the complex curving lines and texture of the sunflower (consistent appearance), its size and its position on the foreground plane of the poster (what is known in graphic design as a *figure-ground* relationship).

Figure-ground describes the relationship between forms in a composition of 'figures' perceived as occupying space on a particular 'ground' [108](p47).

Thus, forms which are perceived as occupying space are referred to as *figure*, the space in which they occupy is referred to as the *ground*. Any number of forms can be figures on a particular ground. However, the degree to which a figure stands out may be discussed in terms of salience.

The salience of the arrow is a controllable compositional system dependent on *formatting objects* and *information value*, as described by Kress and Van Leeuwen [60](p203-212).

Formatting objects The actual nominal modes[3] involved in the *expression of content* [39](p48). This includes, for instance, nominal modes such as colour, boldness, padding, font-family and line width on the medium of paper. Formatting objects are not limited to visual modes; they are also identifiable across all media and modes.

If we consider the arrow to be a simple graphical element, then we can understand its modes of boldness, colouring, and relative size as being affected by *formatting objects* [96](p24). These formatting objects can be understood as being similar to those applied to formatting of text or other media elements (akin to CSS [23], XML/XSL [105], DSSSL [54], and Hypermedia Formatting Objects [69](p22-29)) (see also Ossenbruggen [96])).

Information value is the emphasis created by the placement and ordering of the spatial or temporal elements within a composition [60](p203-12).

The information value, or the semantic relevance as it is termed by other semioticians [33], varies depending on the placement of the sign in the spatial context.

To summarise, the arrow has the highest information value and is the most salient sign of the composition. The arrow's salience is achieved through its graphic design features, which include boldness, colouring, and size, which create a strong contrast against the

[3]By *mode* we mean a particular form in which content is expressed – such as 'the mode of writing'.

sunflowers in the background. The arrow's information value is achieved by its dominant size in relation to other signs in the composition, as well as its central position in the poster.

How the visitor knows that the poster relates to the exhibition is predominantly, although not solely, the function of the sunflowers.

3.1.3 The sunflowers sign

The sunflowers depicted in the poster are a familiar reprint of one of Van Gogh's paintings. Although only part of the painting is shown, it is obviously recognisable as one of Van Gogh's artworks: the energetic brushstroke style and use of colour, which are culturally synonymous with Vincent Van Gogh [60](p236-7), feature significantly in this sign. This is of significance: to the visitor who knows the details of the relationship between the painters Vincent Van Gogh and Paul Gauguin, the sunflowers symbolise the relationship between the two painters – sunflowers are a common subject matter of both artists. But even to those visitors who would not make the latter connection between the flowers and the painters, the stylistic conventions used in the poster are associable with the rest of the promotional signage of the exhibition. (We will return to look at these stylistic associations later in the discussion). So we can say that the sunflowers are standing for more than just Van Gogh's real *sunflowers* painting – they are standing for the whole *Van Gogh Gauguin* exhibition. In other words, the graphic designer who designed the poster has explicitly used the sunflowers to act as a *rhetorical trope* [31], [39](p280-3), [87](p181).

Rhetorical tropes allow signs to become part of larger systems of meaning through a semantic association [31]. Rhetorical tropes are a fundamental part of any sign system. The most common rhetorical tropes include metaphor, metonym and synecdoche [39, 92].

A metaphor functions by transposing domains to create an analogy or a simile (i.e. 'he's sluggish' or 'my love is like a rose') [31], [39](p279-86).

A metonym conjures up a whole concept by close indexical association (i.e. 'all hands on deck', where 'hands' may stand for people) [31], [92](47-48), [39](p279-86).

A synecdoche is a special case of metonymy involving the substitution of part for whole,genus for species or vice versa [31].

The importance of rhetorical tropes is that they are mechanisms to generate 'imagery' with connotations over and above any literal meaning. Thus, rhetorical tropes can, once being described as transformation rules, be used to improve the collection of signs as well as their presentation. The particular kind of rhetorical trope that the sunflowers create a *synecdoche*, where the part of the sunflowers painting is standing for the whole exhibition.

The final element left to examine in the poster is the plastic, which is the sign on which the other two elements have been composed.

3.1.4 The plastic sign

The plastic is the medium on which the graphic designer composes the arrow and the sunflowers. Here, the plastic can be understood as acting as a *composition* – that is, a medium on which this group of signs can be composed with communicative intent.

Composition is formed by a sign or a group of signs composed into a unified signifying structure, within a particular frame [108](p345), across one or many media. In the words of Kress and Van Leeuwen [60](p181), compositions are "the way in which the representational and interactive elements are made to relate to each other, the way they are integrated into a meaningful whole."

The plastic also forms a frame for the composition, giving it a context in which to signify. So the function of the plastic material is triple: it is (1) a sign, (2) it is a compositional element, and (3) it provides a *frame of reference* [108](p347).

Frame of reference The natural boundary imposed on the composition by the termination of a medium. It "marks the outer limits of design and defines an area within which the created elements and left-over blank space, if any, all work together" [108](p44).

The boundary of the plastic material, for instance, forms the frame of reference in the case of the poster.

A sign or a set of signs may also create an implicit frame within a composition independent of the frame of reference. This type of framing we will call *internal framing*. A composition may contain one or many internal frames, it all depends on how signs are grouped and distributed in terms of information value within the composition. Frames connect signs together into a signifying relationship, and disconnect those signs from other elements in a composition [60](p183).

Internal framing is created from perceivable frame lines, either explicit or implicit, that perceptually and semantically connect a set of related signs within a composition. Internal framing may also disconnect a set of signs from other elements in a composition [60](p183). Internal framing is interrelated with the salience and the information value of a sign [60](p183). Both types of framing (internal and reference) always occur as part of a composition.

As a sign, the plastic material also carries a considerable amount of information [12, 61](p79-81). Kress and Van Leeuwen [61](p79-81) [60](p232) and others[4] point out that

[4]In Kress and Van Leeuwen [61](p232) and also in Chandler [30]. The most prominent advocate of medium as a sign is Marshall McLuhan [74](p7), [75] who in 1964 famously proclaimed "The medium is the message." McLuhan's proclamation has had wide ranging implications for how media is critically studied and constructed, including in the museum domain where museums are now understood as media [19](p37-43), [18] (museums act as a medium between visitors and artefacts; as a medium the museum construct and control the message). See in particular *museum, media, message* [51] and [19](p37-43).

any material – *the medium* – on which a sign is realised carries a varying signifying value. Plastic is a widely available and commonly discarded material, and so has low cultural value [12](p98). As a medium, the plastic signifies to the visitor to decode the poster as a simple sign, and not as something with high cultural value and worth contemplating over (as the visitor might if the poster had been inscribed on a canvas with oils). In short, a medium functions not only as a sign in its own right but also acts as a compositional device and a framing device.

So far we have seen how each sign in the poster functions independently, now we discuss how the poster functions as a whole and as a part of the museum.

3.1.5 The poster - revised

The point of the poster is that it indicates to the visitor which directions to chose to find the ideal path through the exhibition space. For instance, within the context of the Van Gogh museum, the poster was located near a set of escalators. Here the arrow clearly indicates to the visitor that he or she should go down to continue through the rest of the exhibition. Thus, the designer's deliberate repetition of form and styling achieves that every time a visitor encounters the poster, its function is the same. In other words, the designer achieves cohesion.

Cohesion A coherent repetition of semantically (or functionally) related signs or modes which create coherence inside and across a composition.

If we take the exhibition as the overall composition then it becomes important that functional elements, such as the poster, stay coherent over time. In the case of the poster that can be achieved through the constant use of attributes of formatting objects, such as colour, font family, etc. Take colour-based cohesion as an example. First of all its function can be purely pragmatic in a composition [61](p57-9), [94](p81). It is important to note that colour is applied here not as a sign but as a label, as colour by itself does not have any initial meaning, or referent. In order for it to function as a label it must first be bound to a proper sign such as an icon, symbol, or index. In our example the sign is the arrow. Thus, the process of cohesion is actually related to the labels (attributes) of a sign and hence cohesion functions as part of the visual rhetoric[5] which contributes directly to the communicative process.

For the poster to communicate effectively the person visiting the museum needs, therefore, to identify him or herself as being addressed by the poster. The visitor has what in semiotics would be understood as a *discursive position*, which depends on the surrounding domain and *discourse*.

[5]Colour may also exhibit what Tufte [94](p81) refers to as a "*decorative*" or "*enlivening*" function. That is, "colour as beauty". Colour as beauty is rhetorical as it always functions to reinforce a particular ideological position within any discourse – this we will call an *aesthetic function*, which forms part of dynamic aesthetic codes that are particular to social class and culture [38](p18-9), [43].

Discourse is "a system of representation consisting of a set of representational codes (including a distinctive interpretative repertoire of concepts, tropes and myths) for constructing and maintaining particular forms of reality within the ontological domain (or topic) defined as relevant to its concerns." [28]. Discourse, goes beyond simplistic verbal assertions [52][6], encompassing *everything* that is used in a representation/communication process. This includes all media and their modes, as well all signs experienced in any form and in any modality, be it spatial, temporal, haptic, optical, gesture, acoustic, olfactory, gustatory, kinaesthetic, or a mix of them. Thus, discourse encompasses the context, those who produce, and those who participate in it.

Understanding how discourse works and what it means is of particular importance when discussing museum, where, as Roberts [84](p2) points out, "language about facts and certainties has been replaced by language about context, meaning, and discourse."

Given the large number of visitors attending the exhibition, collaborating professionals must compositionally structure the poster so to communicate with the greatest number of people possible. To achieve this communicative goal, collaborating professionals must use a set of *codes* that are simple enough to be understood – or *decoded* – by a mass audience. In the poster, such codes include the arrow as a directional guide and the flowers as standing for the exhibition. These kinds of codes are known as *broadcast codes*[7] [43](p73-6).

Broadcast code is a code whose meaning is homogeneous amongst a heterogenous audience [28].

The use of imagery from Van Gogh and Gauguin's artworks as a broadcast code is not just reserved for the posters. Much of the promotional signage for the *Van Gogh Gauguin* exhibition uses imagery taken from the artworks of the two painters. This consistent repetition of familiar elements functions to *integrate* texts relating to the exhibition, forming a particular discursive *identity*. "This enables the public to identify quickly museum material and, through frequent exposure, maintain the identity of the museum as an active organization in the public's mind" [19](p25).

Integration uses codes "that serve to produce text, to place the meaningful elements into a whole, and to provide a coherent ordering amongst them."[60](p212). Integration devices, which act more or less as rhetorical tropes, are the deliberate use (and repetition) of modes that function to integrate signs into a related set through semantic association to a wider structure of meaning.

[6]More complex argumentative mechanisms are described in rhetoric, which has begun to move away from formalisations of what constitutes a 'good' argumentation structure and more towards understanding the informal dynamics of argumentation construction in everyday discourses. A detailed description is out of the scope of this paper. The interested reader is referred to [39]. See also our comments on tropes in Section 3.1.3 and for computational issues see [68].

[7]This is, as opposed to the sign containing *narrowcast codes*. Narrowcast codes are domain specific and usually require a reader to be trained in order to decode their significance [43](p73-6).

Integration differs from cohesion. On the one hand, cohesion is tightly bound to the function of a sign within a discourse. For example, the poster is cohesively used several times throughout the exhibition to indicate to the visitor in which direction the exhibition continues. On the other hand, integration is concerned with making texts identifiable as part of a larger structure of meaning. For example, the visitor decodes the poster as being part of the *Van Gogh Gauguin* exhibition, and not as being part of the set of signs showing where to exit the museum in case of an emergency. Emergency exit signs are socially ubiquitous and standardised broadcast codes which contain their own aesthetic function, discourse and discursive positions, and because of this, they are widely understood by visitors.

3.1.6 Summary

Once a curator chooses to repeat the poster through the exhibition it will act as a cohesive device. Cohesion also occurs internally within the compositional space of the poster, where the colours of the composition of the arrow and the sunflowers distinctively set each element apart, creating a degree of coherence. The poster, along with other signs within the museum context, evokes and reinforces the position of 'the visitor' while at the same time establishing the discursive position of the museum. And finally, the sunflowers and styling of the arrow act as integration devices, making the poster identifiable to the visitor as part of the *Van Gogh Gauguin* exhibition.

The poster is in a way the level of multimedia presentation the Cuypers system can generate automatically at the time of writing this paper. In Chapter 4 we will show how the established elements of our descriptive framework can improve Cuypers' communication efficiency.

The intention of the following section is to investigate larger presentation structures, such as an exhibition, as those complex structures that Cuypers, at some stage, should be able to generate.

3.2 The exhibition as a narrative

Roberts [84](p134) argues that a narrative process within a museum exhibition "aims to establish not truth but meaning; explanation is achieved not through argument and analysis but through metaphor and connection." The *Van Gogh Gauguin* exhibition's narrative is mostly concerned with the emotional and artistic relationship between the artists Vincent Van Gogh and Paul Gauguin, as well as the relationships the two painters had with other people during a nine-week period in which the two painters collaborated.

Narrative Based on [34](p236), [60](p43-78), [16](p109-14), [28], [92](p118), [38](p92), [35](p28-45), we propose a narrative to be a sequence of events *re*-presented by a narrator to a narratee from a particular point of view through media and modes over time.

We can qualify this definition by asserting that narratives are fundamentally comprised of a *plot*, *story events*, and a *sequence* [34][8].

Story The events, which take place in space and time, that make up the narrative.

To illustrate, some of the story events in the exhibition include the time Van Gogh spent in Paris from 1886, the time which Van Gogh and Gauguin spent painting together in Arles in 1888, as well as the time that Gauguin spent in the tropics of Tahiti post Van Gogh's death in 1891, until his own death in 1903. For a narrative to be coherent, this sequence of events must be semantically and causally linked together by an underlying plot [34](p4-7).

Plot The causal relationship between characters and the story, as wells as between one story event and the other events in a narrative [34](p4-7).

For instance, in the case of the exhibition, part of the plot is formed from the emotionally-charged events surrounding Van Gogh and Gauguin relationship, which, the narrator of the exhibition argues, eventually lead Van Gogh to brutally sever his ear. The plot motivates the sequence in which the story events are put together within the narrative [34](p5).

Sequence How the events in the story are spatially ordered in relation to the overall narrative[9] [34](p7-16).

In the case of the exhibition, the narrative sequencing of story events are generally conveyed in a chronological order. However, there are a number of occasions when the narrative jumps out of order in terms of time and space. For instance, the first story event the narrator describes is how the two artists met, as well as giving a brief overview of their artistic approaches; this event is followed by a reversal in time and space to the city of Paris, where the narrator tells the visitor of the different effects that living in Paris had on the two artists. The story events, plots, and narrative sequence are shown and told to the visitor through the process of narration.

Narration How the story events and plots are represented through signs with respect to the media and modes used to narrate the narrative.

[8]Cobley breaks down narrative structures even further by asserting implied positions for the reader, which he calls the *implied reader*, as well as an implied position for a producer of a narrative, which he calls the *implied author*. He further expands this by also distinguishing between the implied reader and the real reader, just as we did in this paper in terms of the visitor (as a discursive position) and the real visitor. We limit this discussion of narratives to just narrator and narratee positions.

[9]A sequence of events can be arranged in any number of ways, as long as there is a plot to hold the causal relations between events. Quentin Tarantino's film *Pulp Fiction* (1994) [90], for instance, has become a more or less quintessential example of a narrative in which the story events are presented in a non-linear fashion yet works well as a narrative.

Figure 3.2: The Booklet, the digital transceiver, and wall-labels are some of the elements through which the exhibition narrative is conveyed.

There may be any number of narrators in a narrative, all of which speak from their own particular point of view. An *omniscient* narrator is one who is not directly involved in the action, but retells stories as if they were a god-like being or a fly on the wall [34](p100-4). The omniscient narrator is all knowing, able to jump to any place at any time – including into the minds of the protagonists and other characters. As we will show, this is the method of narration used in the *Van Gogh Gauguin* exhibition. A narrator who directly speaks to the visitor is known as an *intrusive* narrator [34](p108)[10]. For instance, a narrator speaking directly to an audience in a second-person voice would be considered intrusive.

A narrator always narrates to a *narratee*.

Narratee The ideal entity to which a narrative is told [34](p236).

The narratee is not necessarily the visitor, but another discursive position that is established between the constructed narrator and the visitor.

The narrative process of the *Van Gogh Gauguin* exhibition takes place through various signs, codes, and architectural forms that collaborating museums professionals have composed to communicate in unison, some of which are shown in Figure 3.2. Prior to entering the two rooms in which the exhibition is held, the staff of the museum provide the visitor with a booklet and a digital transceiver to guide them through the exhibition. As the narrative is *told* to the visitor through the transceiver and the booklet, the artefacts become the means through which the narrative is *shown*. In other words, text in the booklet, audio

[10]Intrusive narration is generally seen in novels; however, less so in films and on television, where this kind of narrative voice – when a character speaks directly into the camera – is mostly used for comical effect [34](p190-9).

emitted by the digital transceiver and the artefacts act as the *narration devices* within the exhibition.

Narration devices Signs specifically used to narrate a narrative.

As the visitor enters the room in which the exhibition is held, the first thing he or she encounters is two paintings mounted on a wall. As the visitor continues walking through the exhibition, the walls periodically change colour signalling a change in the time and space of the narrative. In other words, the function of rooms and walls is to act as *narrative containers*.

Narrative containers The elements that spatially support the narrative, such as rooms and walls.

As can be seen in Figure 3.2, a large wall label identifies the subject matter of the two paintings. Inscribed on top of each painting is also a plaque – a group of labels that assert a number of claims about a particular painting. The wall labels and the plaques make up part of the exhibition's *labelling system* [19](p167).

Labelling system The function of each label in a labelling system is to impart information in a concise and comprehensible way [1](p130). The labelling system, however, functions as a self-contained communication system and also in conjunction with the narration devices and the architectural forms of an exhibition [19](p167-8)[11] Belcher[19](p130) also notes that "the content of a label is an educational matter, but its visual presentation is an organic part of the whole exhibition it" and so, "it should conform in color, scale and location with the overall scheme". In other words, the labelling system must make use of integration devices.

To gain a further understanding of how exhibition designers convey the narrative to the visitors, we will now look at the narration devices, the narrative containers, and the labelling system in greater detail.

3.2.1 Narration devices

In the case of the exhibition, many different narrators and narratees are established through the multilingual support provided by the narration devices; both the digital transceiver and the booklet were available in a number of European languages. Although the visitor may consult them together, the digital transceiver and the booklet independently convey different narratives.

To most effectively communicate the narrative, each narration device makes use of the various modes available to its medium. The digital transceiver, for instance, incorporates various auditory signs, which represent a particular space and time. The booklet, on the other hand, makes use of images and text to show and tell the narrative.

[11] See in particular Roberts [84](p47-79) for an in-depth discussion on the historical evolution of labelling systems and interpretive practices. See also Belcher [19](p147-68) in relation to how multiple disciplines work together in constructing a labelling system for a museum exhibition.

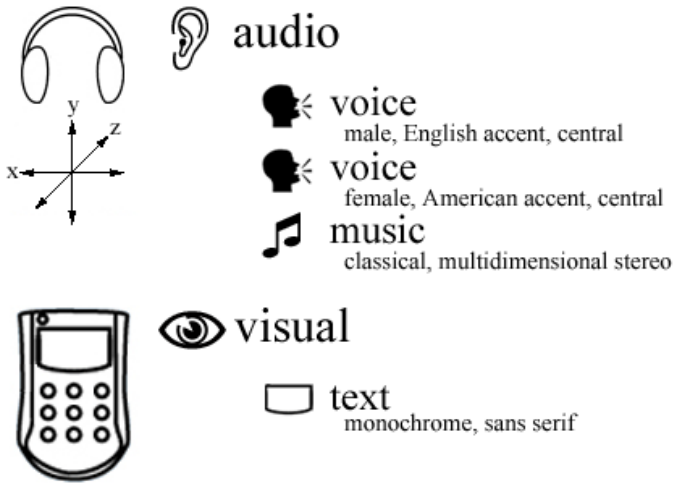

Figure 3.3: The media, modes, and modalities of the digital transceiver.

Booklet

The narrator constructed in the booklet narrates unobtrusively from an omniscient third-person position. Consider, for instance, the narrators omnipresence in the following passage from the booklet (italicised for emphasis):

> "Once in Arles, *Van Gogh became entranced* by the plan of founding an artists' colony intended to give French art a new stimulus, the 'Studio of the South'. *Misunderstood* artists such as himself, Gauguin, and Bernard [a friend of the two artists] were to join forces and found a society of kindred spirits. As an art dealer, Theo [Van Gogh's brother] would promote the interests of this brotherhood in Paris."

The narrator's omnipresence is demonstrated in the clause "Van Gogh became entranced". Here, the narrator extracts a subjective state of being, "entranced", and proposes it as the way things were psychologically for Van Gogh at a particular instance in time. The narrator also makes a value judgement by stating that the artists were "misunderstood", presumedly, by the society of the time. As far as the narrator is concerned, the validity of all assertions are to be taken by the visitor as true [34](p236-7).

The digital transceiver

The digital transceiver, illustrated in Figure 3.3, plays back a composition of aural signs that have been structured over time to convey parts of the narrative to the visitor.

The headphones of the transceiver transmit the voice of two people. One is a male voice, the other is female. The function of the female voice is to instruct the visitor, yet directly serves no function in narrating the narrative. The function of the male voice is

that of narrator. The male voice is clearly that of an older male, spoken in a bourgeois English accent – educated, refined, and sophisticated with strong historical and academic connotations. Kress and Van Leeuwen [61](p82) state that, "voice can be made to mean on the basis of *provenance* and on the basis of experiential meaning potential [ie. broadcast codes]".

Provenance Signs being imported by designers "from other contexts (another era, social group, culture) into the context in which we are now making a new sign, in order to signify ideas and values which are associated with that other context by those who import the sign" [61](p10) (see also Barthes in relation to *myth* [10]). In a similar way to rhetorical tropes, the use of provenance allows signs to represent ideological and semantic systems through a visitor's culturally-constructed (*mythic*) understanding of how the world works.

Broadcast codes can be easily created by a designer's use of signs that conjure up provenance. In the exhibition context, the provenance of the male voice provides visitors with the means to mythically evaluate the authority of the information they are receiving.

Artefacts

As we previously stated, one of the roles of the artefact within the narrative is to be the means through which the story events are shown. Artefacts being used to progress a narrative in the manner as being done in the *Van Gogh Gauguin* exhibition, which is of the type *thematic* exhibition [19](p66). According to Belcher, "thematic exhibitions start with a story line and draw upon objects to illustrate the theme.. . . The theme is the linking element between objects which are strung along a liner sequence, like beads on a necklace."

Thematic exhibitions are but one type of exhibiting method available to museum curators and designers. See Belcher [19](58-66) for a detailed description of other exhibition types and how these types are often hybridized.

3.2.2 Narrative containers

As stated previously, the function of narrative containers is to spatially support the exhibition narrative. This means that rooms, for example, provide the means of clustering information and thus work as a structural devices on a macro level of the narrative (a room is a chapter). On the other hand, narrative containers can also serve on a narrative micro level, such as walls can be interpreted as sections.

Rooms

A room is an architectural form consisting of walls and at least one doorway. We can initially view the space in which an exhibition takes place as a single room. In the case of a temporary exhibition, such as the one under discussion, the exhibition designer may segment the space by erecting walls. This temporary measure effectively creates smaller

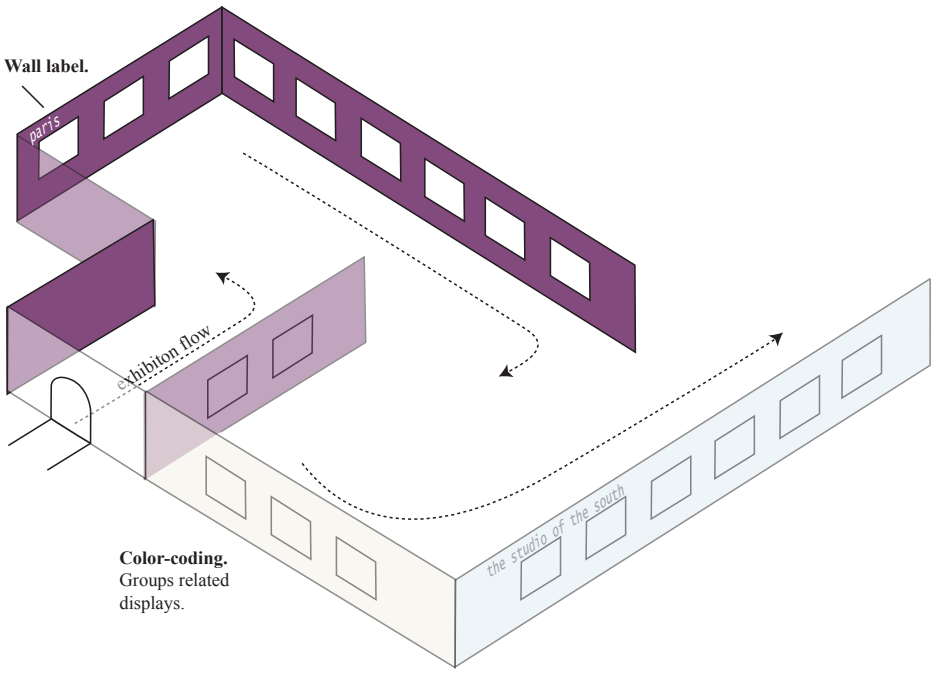

Figure 3.4: A chain exhibition. In this instance, colour is used as a means of grouping related artefacts.

rooms in which the curator may exhibit semantically related artefacts and employ other signs and architectural forms that will optimally facilitate the showing and telling of the exhibition narrative [19](p47-51), [84](p131-52).

For the *Van Gogh Gauguin* exhibition, the exhibition designer structured the walls using a method known as *chain* [19](p112). As can be seen in Figure 3.4, a chain exhibition involves erecting walls in such as way as to enclose a finite number of artefacts within a space. These spaces not only limit what visitors can see at any one point, but also suggest a logical sequential flow for the exhibition [19](p99). As the visitor traverses the exhibition, the size and repetition of these spaces creates a sense of *rhythm*.

Rhythm A pattern-based repetition of elements over space and/or time. Rhythm can be roughly classified into either *alternating* or *progressive* rhythm [63](p104-7) (see also Thiel [91](p221)). *Alternating* rhythm is formed by a binary repletion of elements, such as the alternating columns of a classic Greek architecture [76](p27-30). *Progressive* rhythm, on the other hand, changes at regular intervals forming sequential patterns, such as a fade-in/fade-out transition in a movie. An extensive vocabulary exists in the discipline of music to describe various subclasses of both kinds of rhythm [91](p224) [63](p102). However, we can only acknowledge that these terms exist and consider their classification beyond the scope of this work. When rhythm is consistent, a design is said to be *harmonious*.

According to Belcher [19](p112-5), breaking up an exhibition into rhythmic modules works so that the visitor does not become bored and agitated with the exhibition. It gives the delivery of the narrative a sense of motivated pace.

In the case of the *Van Gogh Gauguin* exhibition, roof, floor, and lighting designs of rooms also have an important communicative role within the exhibition. The neutral colour and texture of the floor, and the neutral colour of the lighting, function to not detract attention from the artefacts on display, the narration devices, and the labelling system [19](p125). In another instance, however, these design elements may be used to directly manipulate meaning of the exhibition narrative by changing the way meaning is constructed by the visitor [19](p123-46).

Walls

A wall is an architectural form on which collaborating professionals compose artefacts, their corresponding plaques, colours, and wall labels into a meaningful composition. Together, these elements are used in conjunction with the narration devices to facilitate showing and telling of the exhibition narrative.

A wall creates a natural frame of reference. By using a colour, collaborating professionals are able to extend this frame of reference by bridging together multiple walls into a single continuous text. This use of colour can be seen in Figure 3.4, where five adjoining walls are painted purple to group together the paintings that correspond to a particular set of events in the narrative. The wall label and the narration devices suggest which events are being depicted across a particular number of colour-grouped walls.

As depicted in Figure 3.4, the wall label that is attributed to the five purple walls is the proper noun 'Paris'. Within the context of the exhibition, 'Paris' acts as a metonymic rhetorical device. Paris is a metonym as it stands for Van Gogh's and Gauguin's experiences in Paris at a particular period in time, and only indirectly stands for the French city. The metonymic function, and hence the intended semantics of the noun Paris in this instance, will further become clear to the visitor once he or she refers to the booklet and/or listens to the audio tour delivered on the digital transceiver [14].

In the case of the purple wall, colour is functioning as a label: it carries the semantics of the metonym Paris. The colour terminating or meeting another colour forms a *transition*.

Transition A transition marks a shift in time and or space within a narrative [49](p135), [40](p187).

To illustrate how colour is used a transition, after his time in Paris, Van Gogh moved to Arles in the south of France to establish a painting colony. This transition in time and space is represented in the exhibition narrative by changing the wall colour from purple to blue and including the wall label 'The studio of the south'. The combination of these elements also forms a code, which the visitor understands as a transition to a different set of story events within the narrative.

The colours that the exhibition designers' selected include purple, yellow, and blue – colours that will form cohesion with surrounding paintings as they are well-known to have been used by Van Gogh. From our observation, there appears to be no relationship between the colours of walls and particular events and/or causal relationships in the narrative. The accompanying wall labels and narration devices denote what a particular wall colour signifies at any particular point in time and space — this is done never by the colour alone. In this sense, the walls are redundantly acting as containers for a particular number of story events within the overall narrative.

3.2.3 Labelling system

Labels at the exhibition come in two forms, as *direct labels* and *indirect labels*.

Indirect label An indirect label is a sign that functions as a code and links to key where further information may be obtained [95](p98-101).

By looking at Figure 3.5, it can be seen that the number 2 in the plaque is a code which corresponds to a key in the booklet and, if the headphone icon is present, to a key combination on the digital transceiver.

Direct Label A direct label is a noun-string or sign description of an element or a group of elements.

For instance, in Figure 3.5, the label '*Self portrait at the easel*' directly labels the artefact.

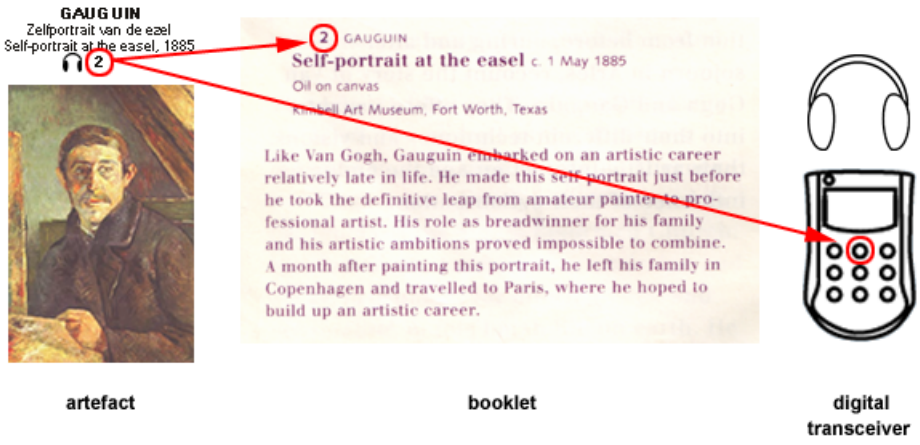

Figure 3.5: The number 2 in the plaque is an indirect label corresponding to the booklet, and the digital transceiver.

Indirect labels can also function as direct labels, and vice versa. For instance, as previously discussed, one function of the number 2 is to denote the reference number of the artefact in the exhibition. Another function of the number 2 is to act as a deictic (indicating) device that implies the linearity of the exhibition layout and flow.

Plaques

A plaque is a sign on which collaborating professionals visually compose a specific set of metadata [47] about an artefact into a code. As shown in Figure 3.6, plaques are not structured using linguistic conjunctives, such as 'or' or 'and'. Neither are they structured by means of linguistic qualifiers, such as 'creator:' or 'title:'. Rather, the plaque is structured by means of a broadcast code. How the visitor can be sure that he or she is deciphering this information correctly depends on how often they have come across such a data structure before within their cultural setting [43](73-74). As a broadcast code, most visitors should understand the word GAUGUIN as implying that Paul Gauguin painted the artefact, and not that GAUGUIN is the title of the painting.

According to Belcher [19](p149-50), the information that goes on a plaque needs to be balanced between what is needed to convey a narrative, and information that the visitor might want to know about an artefact. The only information available on the plaques at the exhibition includes who created the painting, the title of the artefact, the year in which it was created, and the artefact's reference number. Together with the wall label, this information is sufficient for the visitor to access the booklet and the digital transceiver for supplementary information about an artefact.

A graphic designer's use of salience and information value plays an important role in how the visitor decodes the plaque [19](p160-1). By looking at Figure 3.6, it can be seen

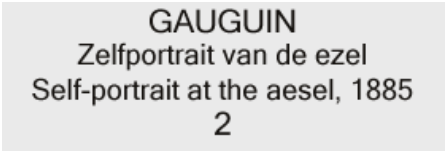

Figure 3.6: An example illustration of a plaque from the *Van Gogh Gauguin* Exhibition.

that the last name of the artist and the number of the artefact are more salient than the artefact's titles. These elements have a higher information value as they have been made salient in relation to other elements within the composition.

Given that we now have a definitive understanding of the elements that make up an evolving multimedia presentation (the reader should be reminded that this is how we see the exhibition in the context of this paper), we can propose the following definition for the *narrative flow of a multimedia presentation*.

Narrative flow of a multimedia presentation A motivated spatiotemporal discourse with a communicative function that is sustained through architectural forms, codes, and signs that surround a group of strategically positioned artefacts [84], [19](p37-57), [60](p243-62), [93]. It is a medium, which always serves to construct meaning in between the artefacts and the visitor [19](p37-57), [18](p403), (see also [51] for an extensive discussion). As a medium, it comes with its own particular modes of articulation, ideological positions, and practices through which various professionals construct meaning [19](p37-57), [51]. The narrative mode is one of the modes that can be used within a multimedia presentation to form semantic relations between artefacts through the processes of showing and telling. The actual process of narration takes place through various signs acting as narration devices.

3.3 Summary

In this chapter, we analysed the *Van Gogh Gauguin* exhibition, in order to establish the elements for our descriptive framework, which should help to improve Cuypers' ability to generate multimedia presentations. We used semiotic theory as the basis for our investigation because semiotic theory has influenced museums to move beyond passive mechanical transmission of information to the public.

We divide our discussion in two parts. First, we investigate environmental sign systems that are commonly found throughout museum exhibitions but do not contribute to an exhibition's narrative. We used the example of a poster giving directions to discuss the basic elements required for the process by which meaning is established (semiosis). The form of a poster was chosen because it is the level of multimedia presentation the Cuypers system can generate automatically at the time of writing this paper. As a result we established a set of relevant communication elements and their attributes, as well as processes (e.g. rhetoric tropes) that can manipulate these elements.

The second part of our investigation extended the established findings by those elements and related processes that are required to extend the single event of an environmental sign system to a sequence of events *re*-presented from a particular point of view through media and modes over time. The results of that investigation should help to shape the ongoing research in which the Cuypers system is embedded, such as the discourse model.

In the following chapter, we will apply the findings of this chapter to regenerate a multimedia presentation generated by the Cuypers system. The aim is to show that the elements of our descriptive framework can be used to make the communication of multimedia presentation more effective.

Chapter 4

A Cuypers presentation evaluated and redesigned in terms of the description framework

Having acquired the basic elements of our description framework through the investigation of discourses in the museum domain (see Chapter 3), we may now return to the digital domain to evaluate and redesign a multimedia presentation generated by the Cuypers system.

The objectives of this chapter are as follows:

1. Identify all the signifying elements, rhetorical relations, and motivations behind a presentation currently generated by the Cuypers system.
2. Evaluate the communicative effectiveness a multimedia presentation generated by the Cuypers system.
3. Put forward a new presentation that builds upon the criticisms made in the evaluation process.

In undertaking these three objectives, we aim to demonstrate that future implementations of the description framework directly into the Cuypers architecture could potentially yield presentations that may communicate more effectively than those currently being generated.

4.1 The Presentation – a structural analysis

The key to performing an effective evaluation will be to break down the structure of a Cuypers presentation into the smallest signifying units, or *sememes* [39](p70), to reveal all components of communication. The static visual elements of the presentation, as they are visually distributed on the screen, are shown in Figure 4.1. This 'screen dump' represents

Figure 4.1: A screen dump from a presentation generated by the Cuypers system. Note: the presentation also contains a spoken commentary (same as the text) and a number of other images showing examples of the clair-obscur technique. These images follow the one shown in this figure.

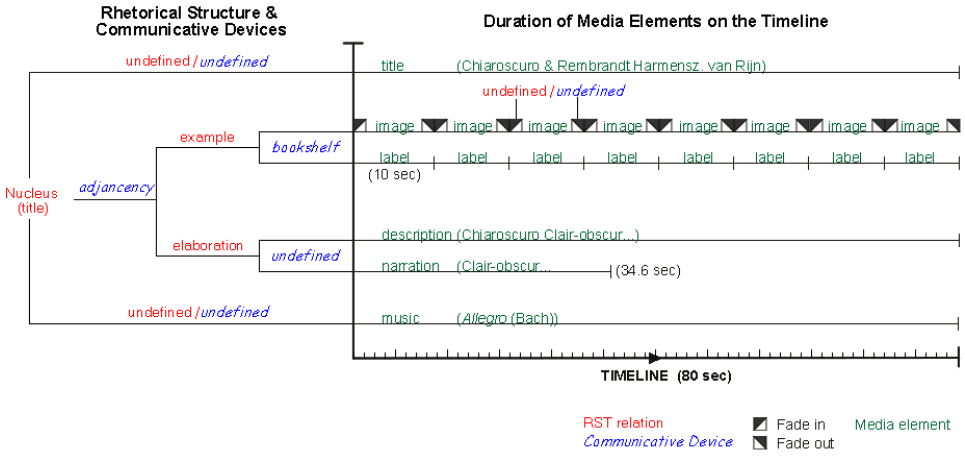

Figure 4.2: The temporal structure of the Cuypers generated multimedia presentation, partly adapted from [97].

the presentation at some random moment in time, using some user-defined generation parameters (such as screen size and particular skill level). Since this layout does not change drastically throughout the presentation it adequately serves as the basis on which to evaluate the general visual layout of elements for the whole presentation (and every other presentation which Cuypers generates using these parameters, since the system currently uses the same rhetorical template for every presentation it generates) [44](p10). Temporally, this presentation occurs across an 80 second time-line along which the following multimodal elements are organised:

- Static:
  - Visual elements: 8 images (presented sequentially over time)
  - Textual elements: title, heading, description, 8 labels
- Temporal:
  - Auditory elements: music, narration
  - Visual transitions: fade in, fade out

The temporal structure of the presentation is represented in Figure 4.2. This figure also depicts the *rhetorical structure* of the presentation in terms of *Rhetorical Structure Theory* [68] and *communicative devices* [86] used by Cuypers to perform the spatiotemporal layout of media elements in the presentation[1].

[1]At the time of writing, ongoing research and discussions on issues relating to communication and a formatting model for hypermedia have resulted in the communicative devices model being superseded by the *Hypermedia Formatting Objects* (HFOs) model, which is described in detail in [69](p22-39).

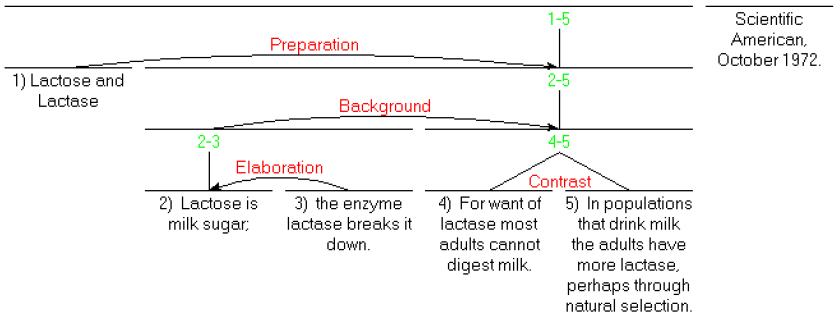

Figure 4.3: An example of an simple RST analysis showing nuclear and satellite rhetorical relations, adopted from [67].

As was briefly discussed in Chapter 1, Rhetorical Structure Theory, or RST, is a formal vocabulary used to describe a specific set of rhetorical relations that occur in written texts [68]. "Because RST generally provides an analysis for any *coherent carefully written text*, and because such an analysis provides a motivated account of why each element of the text has been included by the author, it gives an account of textual coherence that is independent of the lexical and grammatical form of the text" [67] (italics included for emphasis). Written texts that are described in terms of RST result in a computable structure of nucleus and satellite relations that represent the argumentation structure of *text spans* [68]. An example RST analysis is shown in Figure 4.3, which can be read from left to right by following the numbers 1 through to 5. Note that text spans are in a *linear temporal sequence* and *mono*-modal. This is somewhat in conflict with multimedia, where, although also linear, multiple elements may communicate concurrently through multiple modalities.

Geurts [44](p10) states that, in order to generate a presentation, Cuypers takes as input an RST structure encoded as XML. A representation of the RST structure that Cuypers uses is shown in Figure 4.4. This RST structure contains no styling information and so styling relies on a combination of *communicative devices* and decisions made by the Cuypers system. An implication of such a process is that anything not described as part of the RST structure must therefore be hard-coded into the output format of the final presentation. As shown in Figure 4.4, the music's rhetorical function is not related to other elements in the presentation, since it is not considered by the creators of Cuypers to have a formal rhetorical function in the presentation.

As can be seen in Figure 4.4, the rhetorical relations established by the creators of Cuypers between elements in the presentation include *example* and *elaboration*. If these relationships are taken at face value – that is, without scrutinising the lexical, semantic, or spatial attributes of the elements – their relationship can be understood as: 'the images are an *example* of the title and the description and the narration are an *elaboration* of the title'. In other words, just in terms of RST, the creators of the Cuypers consider the title the most important part of the argument being presented; all other elements in the

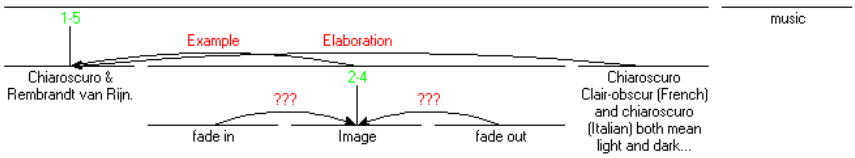

Figure 4.4: A rhetorical structure that Cuypers takes as input to generate all presentations.

presentation satellite the title.

As we previously discussed in Chapter 1, the communicative devices used in the presentation generated by Cuypers include *adjacency* and *bookshelf* [44](p11). The adjacency device basically tells Cuypers that elements should be spatially (visually) "as close to another as possible" [44](p11). The device, however, does not define exact distances or orientation for the element being displayed — decisions that are left up to Cuypers to solve by means of *quantitative* and *qualitative* constraints [45]. On the other hand, the *bookshelf order* device, states that elements should be placed from left to right and top to bottom in some particular predefined order relevant to the domain or context (ordered by date, for instance).

In the case where a bookshelf device cannot fit elements spatially, as is the case in the current presentation, elements must be distributed temporally in a sequence, one after the other [44](p11), [86]. This is shown in the distribution of images, text labels, and transitions in Figure 4.2. How long elements should remain on the screen before being replaced by others remains undefined by the bookshelf device. So, Cuypers decides the appropriate screen time of each element. In the case of the current presentation, this occurs in 10 second intervals in which an image, a label, and transitions between the images are coordinated into an temporal semantic unit: an image fades-in at the same moment when a label is shown, after 8 seconds the image fades out and the label disappears to be replaced by a new label and the next image fades in. The bookshelf device gives no specification as to how far on-top or next-to each element should be to one another. Again, such decisions are left for Cuypers to sort out through its quantitative and qualitative constraint solver [44](p38).

As we stated in chapter 1, *Cuypers does not have the means to self-evaluate the quality of a generated presentation in terms of some established design criteria.* The quality of the presentations relies solely on the rhetorical structure and communicative devices, which may not result in an optimum design solution.

To summarise, we have identified all the elements that make up the presentation both on the spatial and temporal levels (Figure 4.2). In addition, we have identified all the rhetorical relations described in terms of RST between elements as defined by the creators of the Cuypers system (Figure 4.4). By examining the definition of particular communicative devices used in the presentation, we identified which elements Cuypers positions in space and time. We also pointed out that Cuypers relies on communicative devices and a rhetorical structure which only function to spatiotemporally *present* information, but not necessarily to *communicate* it.

Finally, in all the previous chapters we argued that producers of multimedia try to achieve various communicational goals through the end product. This means necessarily that various discourses will be involved not only during the development of the application but also during the application of the tool by the end user. The problem with the current Cuypers system is that the underlying assumed communicative goals are not made explicit, a fact that is reflected in the weak user and discourse models. The aim of the ongoing discussion in this chapter is to provide the means to overcome such problems.

Having analysed the macro structure of a Cuypers presentation we now take a more detailed look at the micro structure of a Cuypers generated presentation screen, as presented in Figure 4.1. This analysis forms the basis for our suggestions for a redesign of Cuypers presentations applying the description framework established in Chapter 3. Thus, we actually apply the facts and rules of our description framework not only for the purpose of a critical analysis but also for the generation process.

4.2 Evaluation

We evaluate a Cuypers presentation using a three-phased approach. In the first phase, the presentation will be deconstructed into basic visual elements and examined as a static visual composition. The evaluation will involve, for instance, examining figure-ground relationships and the information value of static compositional elements. In effect, the initial phase of the evaluation will provide an analytical discussion of the communication processes as it takes place purely through the visual layout.

The second phase of the evaluation involves looking at the presentation from a temporal point of view. That is, how does communication unfold over time, and what function, if any, does time play in meeting the presentation's communicative goal? This stage of the analysis focuses on elements such as the use of music and narrative elements in the presentation, revealing the effectiveness elements that dynamically signify over time.

The third phase of the evaluation focuses on discipline-specific issues in the presentation. For instance, does the presentation work well from a museum education standpoint? From each of the three phases, we will derive a number of criticisms, which we will address in the redesign process outlined in the last section of this chapter.

4.2.1 Evaluating the presentation as a visual composition

The first phase of the evaluation focuses exclusively on the layout of elements as shown in Figure 4.1. For reference and clarity, the elements of the composition have been labelled and framed in Figure 4.5.

Frame of reference and figure-ground relationships

As a visual composition, the presentation is composed of various visual elements forming five figure-ground relationships within one frame of reference. In this instance, the flesh-coloured background denotes the overall frame of reference for the composition; it delimits

Figure 4.5: Elements of the visual composition which are being evaluated.

the user's attention to the limits of the composition and implies that everything contained inside the frame is somehow related and should be considered a unified semantic unit.

Looking at the composition in terms of figure-ground relationships, we can classify the five figures on the flesh-coloured ground:

- The tile: "Chiaroscuro & Rembrandt Harmensz van Rijn",
- the image,
- the label: "Self Portrait (1661)",
- the heading: "Chiaroscuro",
- and the description of chiaroscuro.

The underlying motivation of the presentation is to suggest an order in which the user perceives the elements in the composition. As can be seen from the RST structure of the presentation (Figure 4.4), the title is the nucleus of the argument being presented and so it should be the first element that the user perceives. Less important is the textual description, which functions to rhetorically elaborate on the title and even less important are the images, which rhetorically function as examples of the title. In other words, the RST structure facilitates the relation of compositional elements on the basis of framing and information value (combining both results in the salience of an information unit) to ultimately achieve communication [60](p183).

The RST intended salience of title, description and explanation does not, however, stand out in Figure 4.1, which in fact gives a probable reading order of explanation, description and title. There are many reasons, for example that

- there is only one font used throughout the whole presentation: Arial,
- there is little variance in the font size of textual elements, which ranges only between 14, 16 and 20 pixel points,
- the image area consumes an essential part of the available space and is prominently positioned (the line of reading in western societies goes from top-left corner to bottom-right corner),
- the image area is more contrasting and more colourful.

Salience is also relevant to the order in which a user will perceive elements within internal frames. Internal frames are formed when elements are either connected or disconnected from other elements in a composition by proximity and alignment, "signifying that they belong or do not belong together in some sense" [60](p183). There are five internal frames formed within the composition, and all imply some sort of semantic relationship (these frames are shown in detail in Figure 4.5):

1. Title: the title forms its own frame, as it is neither aligned or in proximity to other elements in the composition.
2. Image + label: the proximity and shared left alignment of the image and the label form a single internal frame.
3. Heading + description: the heading and the description are also in close proximity and share left-alignment, forming another internal frame.
4. (Image + label) + (heading + description): these internal frames are linked together through close proximity and top alignment, forming a further internal frame.
5. Title + ((image + label) + (heading + description)): the title, being centre-aligned across the other internal frames, forms a taxonomical relationship with the other internal frames of the composition.

It can be argued that, although the elements in point 5 are presented as taxonomically related, the internal frame formed around the five elements has little *information value* in relation to the communicative goal. That is, these elements have little semantic relationship to each other and serve a questionable function in meeting the communicative goal. The lack of information value becomes apparent once the presentation is evaluated at a lexical and visual-layout level.

On a lexical level, the title, "Chiaroscuro and Rembrandt Harmensz. Van Rijn", asserts that the composition is about two entities: a concept (chiaroscuro) AND a person (Rembrandt). However, the textual description does not support the assertion made by the title because it never makes reference to the person 'Rembrandt'. The description reads as follows:

> "Clair-obscur (French) and chiaroscuro (Italian) both mean 'light-dark'. Both terms are used to describe strong contrast of light and dark shading in paintings, drawings and prints. Although the effect had already been used for many years, the term only came into fashion in the late sixteenth century. The Originally, the word came from Italy. The painter Caravaggio (1573-1610) made chiaroscuro his trademark. He was a master of at painting illuminated scenes in dark settings."

Although the description text clearly defines the chiaroscuro technique, it actually refers to another painter, the Italian painter Michelangelo Merisi da Caravaggio; an artist of another time (1573-1610) and place (Caravaggio, Italy). In other words, the description does not directly support the communicative goal.

The same lack of support for the communicative goal occurs on the visual-layout level. The image may validate that Rembrandt did in fact use the chiaroscuro technique in his work, but does not educate the user as to *how* Rembrandt used the chiaroscuro technique in the image shown, or in any other of his works.

So, even though the image and the description are visually composed as if they are semantically related, the description never makes reference to the image. In short, the contents and visual layout of the description text and the image only weakly support the communicative goal of the presentation.

Summary of criticisms

From evaluating the presentation as a visual composition, we derived the following criticisms:

1. Incorrect use of style can decrementally influence elements in the presentation leading to an incorrectly perceived order in relation to the rhetorical structure.
2. Incorrect use of internal framing can form unrelated signifying relationships.
3. Content does not directly meet or support the communicative goal[2].

4.2.2 Evaluating the presentation as a temporal composition

The second phase of the evaluation focuses exclusively on the temporal layout of elements as shown in Figure 4.6. The dynamic temporal elements which we evaluate include:

- The images and labels presented in a temporal sequence with the use of transitions.
- The narration (the description narrated word-for-word by a female voice).

[2]It needs to be mentioned here that the Cuypers system does not have control over the content on a deep-semantic level. The reader needs to be reminded that missing reference points are a general problem in the automatic generation of multimedia presentations using "random" material. There is and will be the risk that material is combined that is not fully complementary, resulting in misperception by the reader.

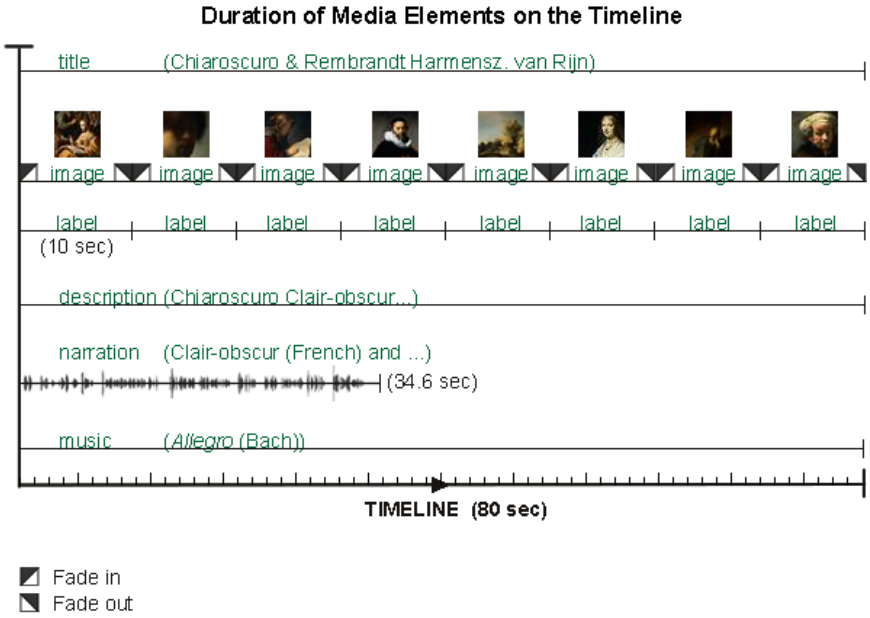

Figure 4.6: The temporal layout of the presentation.

- The music (a work composed by the classical music composer Johann Sebastian Bach (1685-1750)).

Elements presented in a parallel temporal sequence will always be perceived by a user as being semantically connected. Because the images and the narration played back in parallel, when the user watches the presentation they will most likely perceive these two elements as being semantically connected, when in fact, they are not. For instance, from 23.30sec to 29.20sec, the narrator says "the painter Caravaggio, 1573 to 1610, made chiaroscuro his trademark. He was a master of painting illuminated scenes in dark settings." At the same time as these two sentences are being narrated, the third and fourth images in the image sequence are displayed (see Figure 4.6 above). There is no reason for a user not to think that image number four (portrait of Johannes Wtenbogaert) is an example of Caravaggio's work, when in fact the work was created by Rembrandt. Such a phenomenon may occur because of the overlap of rhythms between the narration and the sequence of image.

The same kind of ambiguous synchronisation relationship occurs with the music being played in parallel to the narration and the sequence of images (as also shown in Figure 4.6). The user may assume that there is some utilitarian or communicative relationship between the music and the other elements in the presentation (e.g. the music represents the music of the time), when in fact, there is not. Bach was born in 1685, 14 years after Rembrandt's death, so it was impossible for Rembrandt to have ever listened to a Bach piece. Given that there is no clear functional reason for the music to be part of the presentation, we can only conclude that the music plays an aesthetic function that acts to reinforce an archaic ideological position relating to classical music and the 'fine' arts. Such re-contextualisations

are considered by museum professionals as misleading users [84].

Issues relating to rhythmic harmony become apparent once one evaluates the rhythm of temporal elements in the presentation. For instance, the fade in/fade out transitions used on the images forms a progressive rhythm, while the replacing of the labels one after another forms alternating rhythm. Because these elements are internally framed by proximity, the mix of alternating and progressive rhythms breaks the harmony of the design. Further harmony (and cohesion) is broken by textual labels constantly visually changing x and y axis positions as they are replaced throughout the temporal sequence.

Summary of criticisms

From evaluating the presentation as a temporal composition, we derive the following criticisms:

1. Temporal synchronisation of semantically unrelated elements form ambiguous signifying relationship and information values.
2. Music's aesthetic function reinforces archaic ideological position.
3. The mix alternating and progressive rhythm breaks the harmony of the temporal design.
4. Position changes of labels breaks cohesion and harmony of the design.

4.2.3 Evaluating discipline-specific concerns

In this final phase of the evaluation process, we discuss some concerns that different disciplines, such as multimedia and hypermedia researcher or educationalists, may have towards the current presentation.

Multimedia and hypermedia development

The way that arguments are constructed in hypertext/hypermedia differ significantly from the way they are structured in other media [25, 62], including multimedia. Johnson and Blair (in [25](p86)) describe hypertextual argumentation as "digressive, rhetorical, repetitive, ill-organised, incomplete, and multi-functional". Again, this is in sharp contrast to arguments structured in the presentation using RST, which, according to Mann and Thompson [67], are supposed to be "coherent carefully written text[s]".

All content for the presentation originally comes from the Rijksmuseum website, a hypertextual context. Because the content has been taken from a hypertextual context, it inherently uses hypertextual argumentation structure. In the presentation, the ill-organised nature of this hypertextual content is evident when one examines the last two sentences of the description: "The painter Caravaggio (1573-1610) made chiaroscuro his trademark. He was a master at painting illuminated scenes in dark settings." Within the context of the Rijksmuseum website, including the two sentences about Caravaggio makes sense

because the word "Caravaggio" is highlighted as a hyperlink. In the context of the Cuypers presentation, however, these two sentences have no semantic relationship to either the title, the sequence of images, or to the communicative goal[3].

Education in a museum environment

Though Cuypers as a system is merely understood by the creators as an "information provider" it seems that the system designers also have other communication goals in mind. Take the following quote as an example:

> "The user (studying art history) just asked the system to explain the use of the *chiaroscuro* technique (strong context of light and dark shading) in the paintings of Rembrandt van Rijn" Geurts [44](p9).

From a museum education standpoint, in order for a presentation to be used as an effective pedagogical tool it should support either self-exploration, humour, or play [84](p20). The current presentation does not demonstrate any of these qualities.

Part of the reason why the presentation does not support self-exploration or play is because of its multimedia structure. Hardman [48](p2) states that, multimedia is a "collection of multiple units of information that are constrained by temporal synchronization relationships" and hence multimedia is constrained in terms of interaction to starting, stoping and pausing. The presentation needs to be designed differently if it is to support at least self-exploration.

Summary of criticisms

From evaluating discipline-specific concerns, we derived the following criticisms.

1. Hypertextual rhetorical structure used in a multimedia context.
2. No support for varying discourse principles depending on the communication goal.

Having a better understanding of the specific problems of Cypers presentations we attempt now to redesign the same presentation space according to the description framework developed in Chapter 3.

4.3 Redesign of a Cuypers presentation

Figure 4.7 demonstrates a redesign of the presentation to address the criticisms of the presentation as a visual composition. Ideally, Cuypers would generate something similar to this screen as a response to the user's initial query about how Rembrandt used the

[3]Again, the Cuypers system does not have control over the content on a deep-semantic level, which makes the reuse of material in a different context difficult.

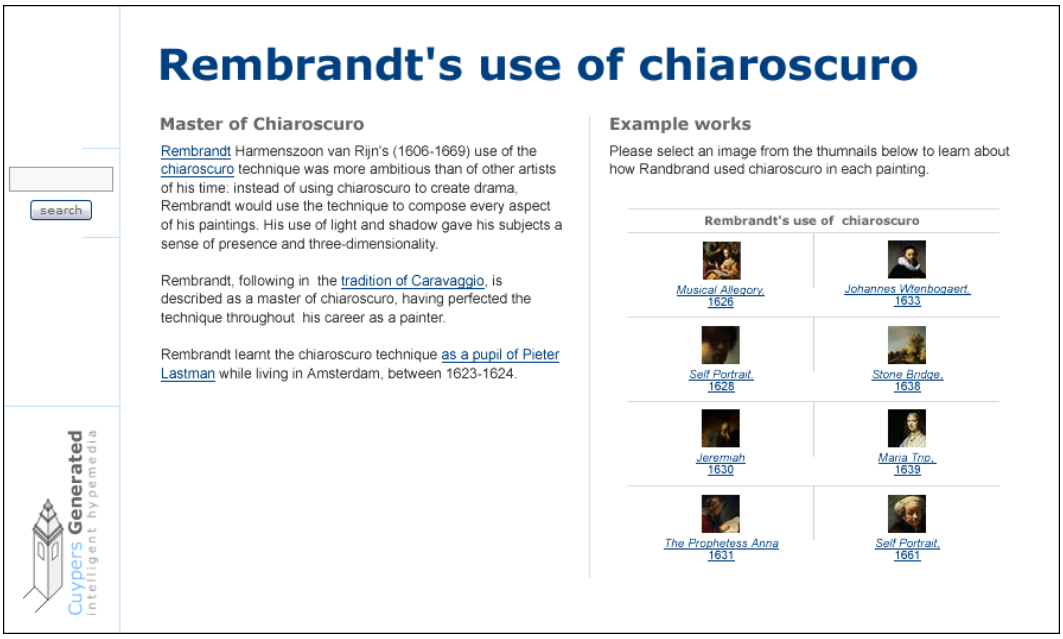

Figure 4.7: A possible design solution for dealing with compositional criticisms of the presentation.

Figure 4.8: A possible design solution for dealing with both spatial and temporal criticisms of the presentation.

chiaroscuro technique in his works. Figure 4.8, on the other hand, demonstrates a possible solution for addressing the criticisms of the presentation as both a temporal and spatial composition. Again, ideally, Cuypers would generate something similar to this second screen as the result of the user selecting the '*Musical Allegory*, 1626' image or hyperlink from the table seen in Figure 4.7.

By systematically stepping through the criticisms outlined in sections 4.2.1, 4.2.2 and 4.2.3, we will explain how the elements of our description framework were used to change the contents and the spatiotemporal properties of elements in the redesigned presentation and hence significantly enhance the communicative efficiency of presentations generated by the Cuypers system.

4.3.1 Addressing the criticisms of the presentation as a visual composition

In the evaluation of the presentation as a visual composition, we made the following criticisms about the original presentation.

Criticism 1 - Unstructured use of salience may cause the user to perceive elements in the presentation in the incorrect order in relation to the rhetorical structure

In the redesigned presentation, in order for elements to be perceived in the desired order, it was necessary to manipulate the formatting objects and internal framing of the original presentation. The order of perception we sought in Figure 4.7 was:

1. Title
2. Headings
3. Description
4. Table of images

For the title to achieve a significant degree of saliency over other elements in the composition, we selected a near-black colour that would contrast strongly against the new white background. In addition, the title is isolated from all other elements and framed with 32 pixels margin. This isolation and contrast allows the title to confidently compete for perceptual dominance against other very salient elements, such as images. To make the title even more salient, we used the Verdana (bold) font at 32 points, instead of the Arial (bold) font used in the original presentation. Verdana is a much bolder font than Arial at the same font-size. So, when Verdana is used in combination with a colour that will contrast against a ground, it can make textual elements relatively more salient than would be possible with Arial.

To make the headings less salient than the title, we made their font-size around 40

To make description texts less salient than the headings, we used the Arial font at 12 point. We used the same grey font colour for the headings and the description text to form a visually cohesive tie between them.

In order to reduce the salience of images, we made them significantly smaller than in the original presentation. In the original presentation, only one relatively large image was displayed on the screen. In the regeneration, all the images are grouped into a table structure and presented as a taxonomical subset of the heading "Example works". In order to make the image table a taxonomical subset of this heading, the table of images was centre-aligned and made 35 pixels smaller in width than the width of the descriptive text above it. Making the images a subset of a heading lowers their overall salience in the composition.

The way salience has been used in the redesigned presentation addresses specifically the order in which we want the user to perceive elements of the composition: firstly, the title; secondly, the headings; thirdly, the description and fourthly, the example works table. We achieved this aim by making elements bigger or smaller, lighter or darker, or by using of internal framing to connect and disconnect elements from each other. By effectively creating an order of perception, we have adorned each element in the composition with its own relative information value, which in turn increased the overall information value (and communicative efficiency) of the redesigned presentation.

Criticism 2 - Incorrect use of internal framing

In the original presentation, internal frames that were not semantically connected were not sufficiently unconnected to be visually perceived as separate (as was the case with the description and the example image - see Figure 4.1). In the redesigned presentation, disconnection between internal frames is made explicit through the use of visual lines. For instance, in Figure 4.7, we place a light-blue line exactly in between the two content frames to explicitly disconnect the description on the left hand of the screen from the description and table of images on the right hand side of the screen. The blue line, however, does not split the framed relationship between the title and the two content frames. In addition, we top-aligned the blue line with the two frames, and made the blue line taller than the height of either frame, so not to imply that the blue line is connected to either frame.

Criticism 3 - Content in the presentation does not support the communicative goal

In order for the textual content of the presentation to support and meet the communicative goal, it was necessary to create a new description and rewrite the title[4]. For the title to support the communicative goal, it was necessary change it from "Chiaroscuro and Rembrandt Harmens. van Rijn" to "How Rembrandt used chiaroscuro". Through flipping

[4]The approach taken here stretches the current Cuypers approach quite drastically, as Cuypers takes information as it is retrieved from the database. Here we just suggest what can additionally performed with respect to automatic text adaptation. Further research has to show if such an approach is feasible.

the order of the words, and through the substitution of the original "and" for "how" and "used", the title now gives a more semantic overview about what the user will find on the screen (Figure 4.7). The words "Harmens. van Rijn" have also been omitted from the title as, contextually, the metonym "Rembrandt" is sufficient to stand for the artist's full name.

The new description principally states how Rembrandt generally used the chiaroscuro technique in all his paintings:

> "Rembrandt Harmenszoon van Rijn's (1606-1669) use of the chiaroscuro technique was more ambitious than of other artists of his time: instead of using chiaroscuro to create drama, Rembrandt would use the technique to compose every aspect of his paintings. His use of light and shadow gave his subjects a sense of presence and three-dimensionality."

Having explained how Rembrandt generally used the chiaroscuro technique in his paintings (which directly meets the communicative goal), this paragraph is followed by complimentary information about Rembrandt's mastership of the chiaroscuro technique:

> "Rembrandt, following in the tradition of Caravaggio, is described as a master of chiaroscuro, having perfected the technique throughout his career as a painter.
>
> Rembrandt learned the chiaroscuro technique as a pupil of Pieter Lastman while living in Amsterdam, between 1623-1624."

Providing complementary information provides the user a wider discourse in relation to their original query. If the user would like to know specifically about the chiaroscuro technique, they have the option of selecting a hyperlink to its definition. If the user would like to know specifically how Rembrandt used the chiaroscuro technique in one of his paintings, then they may select and image or hyperlink from the table of images on the right hand side of the screen. As stated previously, Figure 4.8 is the result of the user selecting either the image or the hyperlink labelled "Musical Allegory, 1626".

Figure 4.8 meets the new communicative goals of 'how did Rembrandt use chiaroscuro in the painting *Musical Allegory*' through the use of both spatial and the temporal coordination of elements[5].

4.3.2 Addressing the criticisms of the presentation as a temporal composition

In the evaluation of the presentation as a temporal composition, we made the following criticisms about the original presentation.

Criticism 1 - Temporal synchronisation of semantically unrelated elements form ambiguous signifying relationship and information values

In the redesigned presentation, we addressed the issue synchronising temporal elements in such a way that elements now semantically complement each other. In order to do

[5]Most of the content for Figure 4.8 was adapted from the Rijksmuseum website.

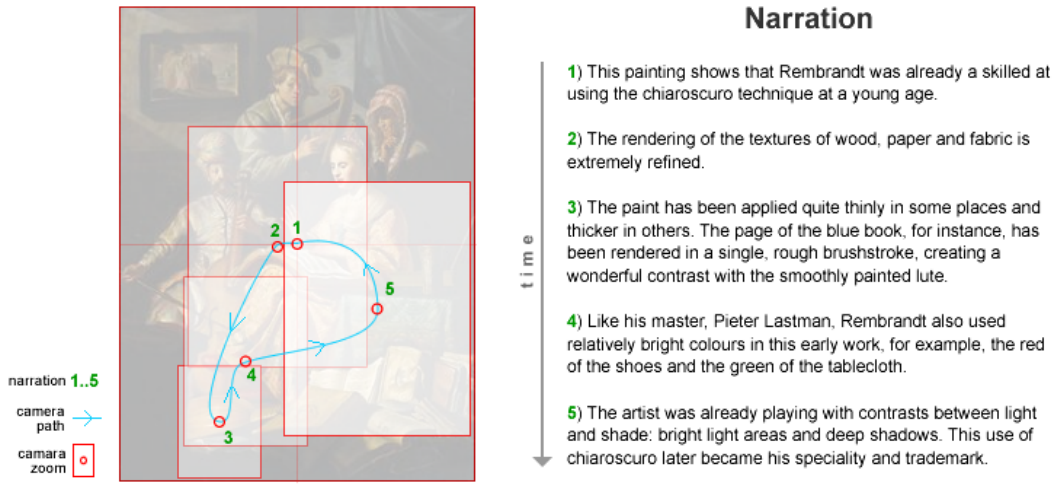

Figure 4.9: Example of possible solution to synchronisation of temporal elements, based on semantic relationships that are formed through temporal synchronisations of modalities.

this, we embedded a mini multimedia presentation inside the redesigned presentation. The mini presentation is displayed in Figure 4.8, framed on the left hand side. In this mini multimedia presentation, we have synchronised over time the movement and zoom of a simulated camera with some narration: as show in detail in Figure 4.9, as the narrator speaks, the camera would simultaneously move and zoom into particular areas of the picture which are directly related to what is being spoken by the narrator. Note, generating the multimedia presentation automatically requires better ways of annotating the original material, so that the presentation engine not only knows which material can be used for such a presentation style but also how to manipulate the material. This requires more research.

Criticism 2 – Music's aesthetic function reinforces archaic ideological position

We included appropriate music into the presentation.

Criticism 3 - The mix alternating and progressive rhythm breaks the harmony of the temporal design

We removed those transitions that broke the harmony of the original presentation.

Criticism 4 - Position changes of labels breaks cohesion and harmony of the design

In the redesigned presentation, elements, such as labels, which moved around in the original presentation are now presented at the same position.

4.3.3 Addressing discipline-specific criticisms

In the evaluation of discipline-specific criticisms, we made the following criticisms about the original presentation.

Criticism 1 – Hypertextual rhetorical structure used in a multimedia context.

To overcome the criticisms of hypertextual structures being used in a multimedia presentation, we designed the whole presentation as a hypermedia presentation. The rhetorical features of the unstructured nature and redundancies of hypertext are evident throughout both Figures 4.7 and 4.8. For instance, Rembrandt's affiliation with Pieter Lastman is mentioned on both screens. In addition, both screens provide hyperlinks to the terms 'Rembrandt' and 'chiaroscuro'.

Criticism 2 – Missing interaction

In the redesigned presentation, we focused on implementing self-exploration. We interpret self-exploration to be a user's ability to explore the contents of a presentation without having a strict linear navigation structure imposed upon them by the creators of the presentation: in effect the user is free to create their own interpretations/meanings from the text through their own exploration. Self-exploration is made possible through the inclusion of visual hyperlinks and various interactive tools. For instance, on both screens, many visual hyperlinks are included throughout textual descriptions as a means for the user to explore the contents related to the communicative goal. To support the user in freely reaching a communicative goal through their own active self-exploration, unrestricted paths to exploring the contents have been included. For instance, the user may explore in any order how Rembrandt used the chiaroscuro technique in each of his painting. Alternatively, the user may choose not to view any of the paintings at all, and may simply explore other areas of interest. This ability to explore the content freely is achieved through hyperlinks such as 'Rembrandt', 'Rembrandt as a pupil of Pieter Lastman', or 'tradition of Caravaggio', or even through the search tool provided on the left.

To further support self-exploration, in Figure 4.8 the redesigned presentation includes four additional interactive tools:

- A set of presentation control buttons and a slider,
- a hand tool,
- a zoom tool,
- and a button to explore the image 'full-screen'.

The user may use the presentation control buttons and slider for controlling the playback of the mini multimedia presentation show in Figure 4.8. The function of the buttons

includes play/pause and stop. The function of the slider is to freely jog through the mini-presentation. The function of the hand tool is to freely explore the image in the mini presentation by dragging the camera along the x and y axis. The function of the zoom tool is to explore the image shown in the mini-presentation in greater/lesser detail. The function of the 'full-screen' button allows the user to explore the image with a significantly larger viewing area without the interference from other multimodal elements.

Possibly an unlimited number of other tools and features, such as forward and back buttons, advance search methods, could be designed into the presentation to aid in self-exploration (and usability) and as a means of better meeting the communicative goal of the redesigned presentation. We have, however, constrained the redesign to address the basic criticisms outlined in this chapter within a limited set of multidisciplinary discourses.

In other words, we are here confronted with those problems described in chapter 3 in the context of the exhibition as a narrative. The problem with this type of communication aids is that they are very difficult to be achieved automatically, as they require an understanding of the multimedia material on a deep-semantic level. We are aware that a great deal of research is required with respect to making content accessible as well as manipulating it.

4.4 Conclusion

In this chapter we decomposed a typical Cuypers presentation into its basic signifying elements and evaluated as well as redesigned it based on our description framework we developed in Chapter 3. Through the evaluation process, we were able to critique the spatial and temporal quality of the presentation. In our redesign we proposed a new set of designs that addressed all criticisms raised in the evaluation process. As a result, the new hypermedia presentation should theoretically communicate more effectively than the original multimedia presentation generated by Cuypers.

Chapter 5

Conclusion

The aim of this paper has been to demonstrate how the use of theory in graphic design and semiotics can help to improve the communicative efficiency of multimedia presentations generated by Intelligent Multimedia Presentation Systems (IMMPS). The system that has been the focus of this paper is Cuypers, an IMMPS currently developed at the Multimedia and Human Computer Interaction group at CWI in Amsterdam.

Initially, it was shown that the current approaches used by IMMPS researchers were limited in their communicative efficiency. In particular, we demonstrated that existing IMMPS lack much of the communicative sophistication currently seen in commercial-grade multimedia and hypermedia products. In order to overcome these limitations, we argued that IMMPS researchers need to better consider the motivations for constructing a multimedia presentation. Identifying the motivation for constructing a presentation may determine the disciplines that should be involved in the processes of generation, evaluation, and collaboration and hence result in more effective communicative outcomes. In addition, we argued that all multimedia presentations are a form of representation and that authors of multimedia may commonly need to draw on theories of representations to generate and evaluate presentations. Since multimedia authors draw upon theories of representation, we deduced that IMMPS will more than likely have to base their generation and evaluation processes on theories of representation to meet their motivated communicative goals.

As Cuypers generates multimedia presentations using data from the domain of musea for the fine arts, we studied how effective multimedia communication is constructed within a museum exhibition. For the discussion, we deconstructed a static visual composition of an exhibition as well as the overall design of the related exhibition. It was shown that every element in a multimedia presentation, and all those disciplines involved in putting one together, function in unison to reach motivated communicative goals. In doing so, we demonstrated that multimedia is not simply the spatio-temporal amalgamation of different media, but that each medium comes with it own complex combination of established disciplines and discourses that impact upon the generation, evaluation, signification, and collaboration of all media elements in a multimedia presentation.

As an outcome of the study, a description framework was formulated by drawing on the discourses of various key disciplines involved in building a museum exhibition. We

did this to show the validity of our initial assertion that effective multimedia generation is always a multidisciplinary activity and that designers require a shared discourse and a multi-disciplinary description framework in order to generate, evaluate, and collaborate in designing multimedia presentations.

We demonstrated the validity of the description framework as a tool for evaluating and generating multimedia presentations by applying it to the analysis and the redesign of a Cuypers generated presentation. The evaluation covered three levels: spatial, temporal, and discipline-specific aspects of the presentation. We then redesigned the multimedia presentation, solving and improving upon all criticism identified in the evaluation process. The outcome of the redesign was a presentation that could potentially communicate more effectively than presentations currently generated by the Cuypers system.

The next step forward for Cuypers would be for its creators to attempt to implement into Cuypers' architecture all functional aspects proposed in this paper – this, however, remains a subject for further research and beyond the scope of this work.

We have, nevertheless, shown that there is great potential to vastly improve the communicative efficiency of IMMPS such as Cuypers through bringing together of multidisciplinary discourses. What is innovative in our approach is our attempt to extend IMMPS discourses beyond traditional artificial intelligence, linguistic, and print discourses, and to consider the experience and knowledge of disciplines that have been, for a very long time, effectively doing in multimedia development what IMMPS are now moving towards.

Given the academic, educational, and commercial demands for highly adaptive and interactive hypermedia experiences, it is surprising that IMMPS-like systems are not yet greatly in demand. This lack of demand may reverse soon, given the majestic adaptability, flexibility, and tailorability promised by the arrival of the Semantic Web and its related technologies [21]. If IMMPS do come into fruition, in our belief, they will have fundamental impact on the way designers think about constructing and delivering hypermedia.

Bibliography

[1] Adams, P. R. The organisation of museums: practical advice. In *The Organisation of Museums: practical advice*, 4th ed. UNESCO, France, 1960.

[2] Adobe. *XMP - The Extensible Metadata Platform*. Adobe Systems Incorporated, September 2001. Available at http://www.adobe.com/products/xmp/.

[3] André, E. WIP and PPP: A Comparison of two Multimedia Presentation Systems in Terms of the Standard Reference Model. *Computer Standards & Interfaces 18*, 6-7 (December 1997), 555–564.

[4] André, E., Finkler, W., Graf, W., Rist, T., Schauder, A., and W.Wahlster. WIP: The Automatic Synthesis of Multimodal Presentations. In *Intelligent Multimedia Interfaces* [72], pp. 75–93.

[5] André, E., and Rist, T. Towards a new generation of hypermedia systems: Extending automated presentation design for hypermedia, 1995.

[6] André, E., Rist, T., and Muller, J. Wip/ppp: Automated generation of personalized multimedia. In *Multimedia 96, 4th ACM international Conference* (1996), pp. 407–408.

[7] André, E., uller, J., and Rist, T. The ppp persona: A multipurpose animated presentation agent, 1996.

[8] Atari. Neverwinter nights, 2002.

[9] Bailey, B. P., Konstan, J. A., and Carlis, J. V. Supporting multimedia designers: Towards more effective design tools. proceedings of multimedia modeling. In *Proceedings of the International Conference on Multimedia Modeling 2000 (MMM00)* (Nagano, Japan, November 13-15 2000), pp. 267–286.

[10] Barthes, R. Myth today. In *Mythologies* [11], pp. 109–159.

[11] Barthes, R. *Mythologies*. Vintage, Paris, [1957] 2000.

[12] Barthes, R. Plastic. In *Mythologies* [11], pp. 97–99.

[13] Barthes, R. *Image-Music-Text*. Fontana, London, 1977.

[14] Barthes, R. The photographic message. In *Image-Music-Text* [13], pp. 15–31.

[15] Barthes, R. Rhetoric of the image. In *Image-Music-Text* [13], pp. 32–51.

[16] Barthes, R. Structural analysis of narratives. In *Image-Music-Text* [13], pp. 79–124.

[17] Beckett, D., Miller, E., and Brickley, D. Expressing simple dublin core in rdf/xml (dcmi recommendation), 2002. http://dublincore.org/documents/dcmes-xml/.

[18] Belcher, M. Communication through museum exhibitions. In *Manual of Curatorship: a guide to museum practice*, J. M. A. Thompson, Ed. Butterworths, London, 1984, pp. 403–411.

[19] Belcher, M. *Exhibitions in Museums*. Smithsonian, Washington, D.C., 1991.

[20] Berners-Lee, T. *Weaving the Web*. Orion Business, 1999.

[21] Berners-Lee, T., Hendler, J., and Lassila, O. The Semantic Web. *Scientific American* (May 2001).

[22] Bordegoni, M., Faconti, G., Maybury, M., Rist, T., Ruggieri, S., Trahanias, P., and Wilson, M. A Standard Reference Model for Intelligent Multimedia Presentation Systems. *Computer Standards & Interfaces 18*, 6-7 (December 1997), 477–496.

[23] Bos, B., Lie, H. W., Lilley, C., and Jacobs, I. Cascading Style Sheets, level 2 CSS2 Specification. W3C Recommendations are available at http://www.w3.org/TR, May 12, 1998.

[24] Brooks, K. M. *Metalinear Cinematic Narrative: Theory, Process, and Tool*. PhD thesis, MIT, April 30th 1999.

[25] Carter, L. M. Arguments in hypertext: a rhetorical approach. In *Proceedings of the eleventh ACM on Hypertext and hypermedia* (San Antonio, Texas, USA, May 30 – June 3, 2000), ACM, ACM Press, pp. 85–91.

[26] Chandler, D. The 'grammar' of television and film, 1994. Available at http://www.aber.ac.uk/media/Documents/short/gramtv.html.

[27] Chandler, D. Working within genres. In *An Introduction to Genre Theory*. 2000. Available at http://www.aber.ac.uk/media/Documents/intgenre/intgenre2.html.

[28] Chandler, D. Glossary of key terms. In *Semiotics for Beginners* [32]. Available at http://www.aber.ac.uk/media/Documents/S4B/sem-gloss.html.

[29] Chandler, D. Modality and representation. In *Semiotics for Beginners* [32]. Available at http://www.aber.ac.uk/media/Documents/S4B/sem02a.html.

[30] Chandler, D. Paradigms and syntagms. In *Semiotics for Beginners* [32]. Available at http://www.aber.ac.uk/media/Documents/S4B/sem03.html.

[31] Chandler, D. Rhetorical tropes. In *Semiotics for Beginners* [32]. Available at http://www.aber.ac.uk/media/Documents/S4B/sem07.html.

[32] Chandler, D. Semiotics for beginners, 2001. Available at http://www.aber.ac.uk/media/Documents/S4B/semiotic.html.

[33] Chandler, D. Syntagmatic analysis. In *Semiotics for Beginners* [32]. Available at http://www.aber.ac.uk/media/Documents/S4B/sem04.html.

[34] Cobley, P. *Narrative: the new critical idiom*. Routledge, London, 2001.

[35] Davis, L., and McKay, S. *Structures and Strategies: And introduction to academic writting*. Macmillan, Melbourne, 1996.

[36] Domike, S., Mateas, M., and Vanouse, P. The recombinant history apparatus presents: Terminal time. In *Narrative Intelligence* (2002), M. Mateas and P. Sengers, Eds., John Benjamins, Amsterdam.

[37] Dufresne-Tassé, C. Andragogy (adult eduction) in the museum: a critical analysis and new formulations. In Hooper-Greenhill [51], pp. 245–275.

[38] Eagleton, T. *Literary Theory*, 2nd ed. Blackwell, Oxford, 1996.

[39] Eco, U. *A theory of semiotics*. Indiana University Press, 1976.

[40] Eisenstein, S. M. *The Film Sense*. Harcourt, New York, 1975 [1942].

[41] Engholm, I. Digital style history: the development of graphic design on the internet. *Digital Creativity 13*, 4 (2002), 193–211.

[42] Ferraiolo, J. Scalable Vector Graphics (SVG) 1.0 Specification. W3C Recommendations are available at http://www.w3.org/TR/, 4 September 2001.

[43] Fisk, J. *Introduction to Communication Studies*, 2nd ed. Routledge, London, 1990.

[44] Geurts, J. Constraints for Multimedia Presentation Generation. Master's thesis, University of Amsterdam, 2002.

[45] Geurts, J., van Ossenbruggen, J., and Hardman, L. Application-Specific Constraints for Multimedia Presentation Generation. In *Proceedings of the International Conference on Multimedia Modeling 2001 (MMM01)* (CWI, Amsterdam, The Netherlands, November 5-7, 2001), pp. 247–266.

[46] Gilliland-Swetland, A. J. *Enduring Paradigm, New Opportunities: The Value of the Archival Perspective in the Digital Environment.* Council on Library and Information Resources, February 2000. Available at http://www.clir.org/pubs/reports/pub89/contents.html.

[47] Gilliland-Swetland, A. J. Introduction to metadata: setting the stage, 2000. Available at http://www.getty.edu/research/institute/standards/intrometadata/pdf/swetland.pdf.

[48] Hardman, L. *Modelling and Authoring Hypermedia Documents.* PhD thesis, University of Amsterdam, 1998. ISBN: 90-74795-93-5, also available at http://www.cwi.nl/∼lynda/thesis/.

[49] Harrington, J. *The Rhetoric of Film.* Holt, New York, 1973.

[50] Hein, G. E. Evaluating teaching and learning in museums. In Hooper-Greenhill [51], pp. 189–203.

[51] Hooper-Greenhill, E., Ed. *Museum, media, message.* Rutledge, London, 1995.

[52] Howarth, D. *Discourse.* Open University, Buckingham, 2000.

[53] International Organization for Standardization. Information Processing — Text and Office Information Systems — Standard Generalized Markup Language (SGML), 1986. International Standard ISO 8879:1986.

[54] International Organization for Standardization/International Electrotechnical Commission. Information technology — Processing languages — Document Style Semantics and Specification Language (DSSSL), 1996. International Standard ISO/IEC 10179:1996.

[55] International Organization for Standardization/International Electrotechnical Commission. MPEG-7: Context and Objectives, 1998. Work in progress.

[56] IW3C2. *The Tenth International World Wide Web Conference* (Hong Kong, May 1-5, 2001).

[57] Kaplan, F. E. S. Exhibitions as communication media. In Hooper-Greenhill [51], pp. 37–58.

[58] Kobsa, A., Koenemann, J., and Pohl, W. Personalized Hypermedia Presentation Techniques for Improving Online Customer Relationships. In *The Knowledge Engineering Review* (2001), vol. 16, pp. 111–155.

[59] Kräutler, H. Observations on semiotic aspects in the museum work of otto neurath: reflections on the 'bildpädagogische schriften' (writing on visual eduction). In Hooper-Greenhill [51], pp. 59–71.

[60] Kress, G., and van Leeuwen, T. *Reading Images: the grammar of visual design.* Routledge, London, 1996.

[61] Kress, G., and van Leeuwen, T. *Multimodal Discourse: the modes and media of contemporary communication.* Routledge, London, 2001.

[62] Landow, G. P. *Hypertext 2.0: the Convergence of Critical Theory and Technology.* John Boltimore, 1997.

[63] Lauer, D., and Pentak, S. *Design Basics*, 5th ed. Harcourt, 2000.

[64] Laurel, B. *Computers as Theatre.* Addison-Wesley, Massachusetts, 1993.

[65] Macromedia. Director 8.5. http://www.macromedia.com/software/director/.

[66] Macromedia. Flash MX. http://www.macromedia.com/software/flash/.

[67] Mann, W. C. An introduction to rhetorical structure theory (rst), August 1997. available at: http://www.sil.org/ mannb/rst/rintro99.htm.

[68] Mann, W. C., Matthiesen, C. M. I. M., and Thompson, S. A. Rhetorical Structure Theory and Text Analysis. Tech. Rep. ISI/RR-89-242, Information Sciences Institute, University of Southern California, November 1989.

[69] Martinez, O. R. Design dependencies within the automatic generation of hypermedia presentations. Master's thesis, Technical University of Catalonia, June 30, 2002. Published as CWI technical report INS-R0205.

[70] Mateas, M., Vanouse, P., and Domike, S. Generation of ideologically-biased historical documentaries. In *AAAI 2000* (2000), pp. 36–42.

[71] Mateas, M., Vanouse, P., and Domike, S. Terminal time: An ideologically-biased history machine. In *AISB Quarterly: Special issue on Creativity in the Arts and Sciences* (Summer/Autumn 1999), pp. 36–43.

[72] Maybury, M. T. *Intelligent Multimedia Interfaces.* AAAI Press, 1993.

[73] Maybury, M. T. Planning multimedia explanations using communicative acts. In *Intelligent Multimedia Interfaces* [72], pp. 59–74.

[74] McLuhan, M. *Understanding Media: the extensions of man.* Routledge, London, 1964.

[75] Mcluhan, M. The relation between environment and anti-environment. In *Media Research: technology, art, communicaiton*, M. a. Moos, Ed. OPA, Amsterdam, 1967, pp. 110–120.

[76] Mitchell, W. J. *The logic of architecture.* MIT Press, Cambridge, 1990.

[77] Moore, D. M., Burton, J. K., and Myers, R. J. Multiple-Channel Communications: The Theoretical and Research Foundations of Multimedia. In *Handbook of Research for Educational Communications and Technology.* The Macmillan Press, 1996, pp. 851–875.

[78] Nack, F. *AUTEUR: The Application of Video Semantics and Theme Representation in Automated Video Editing.* PhD thesis, Lancaster University, 1996.

[79] Nack, F. Catch me, said the meaning, and turned out to be a hydra. In *Workshop on Narrative and Interactive Learning Environments* (Edinburgh, Scotland, August 30 - September 1, 2000).

[80] Nack, F. Play the game. *IEEE MultiMedia 8*, 1 (2001), 8–10.

[81] Nack, F., and Hardman, L. Denotative and Connotative Semantics in Hypermedia: Proposal for a Semiotic-Aware Architecture. Tech. Rep. INS-R0202, CWI, March 2002.

[82] Nielsen, J. Flash: 99% bad, 2000. http://www.useit.com/alertbox/20001029.html.

[83] Reichenberger, K., Rondhuis, K. J., Kleinz, J., and Bateman, J. Effective presentation of information through page layout: a linguistically-based approach. In *ACM Workshop on Effective Abstractions in Multimedia* (November 4, 1995), ACM.

[84] Roberts, L. C. *From Knowledge to Narrative: educators and the changing museum.* Smithsonian, London, 1997.

[85] Roth, S. F., and Hefley, W. E. Intelligent multimedia presentation systems. In Maybury [72], pp. 14–58.

[86] Rutledge, L., Davis, J., van Ossenbruggen, J., and Hardman, L. Interdimensional Hypermedia Communicative Devices for Rhetorical Structure. In *Proceedings of the International Conference on Multimedia Modeling 2000 (MMM00)* (Nagano, Japan, November 13-15, 2000), pp. 89–105.

[87] Saeed, J. I. *Semantics.* Blackwell, Oxford, 1997.

[88] Schmitz, P. Multimedia goes corporate. *IEEE Multimedia* (July - September 2002), 2–5. edited by Frank Nack.

[89] Searle, J. R. *Speech Acts: an essay in the philosophy of language.* Press Syndicate, Cambridge, 1969.

[90] Tarantino, Q. Pulp fiction, 1994.

[91] Thiel, P. *Visual Awareness and Design: an introductory program in conceptual awareness, perceptual sensitivity, and basic design skills.* Washington University Press, 1981.

[92] Thwaites, A., Davis, L., and Mules, W. *Tools for Cultural Studies.* Macmillan, South Yarra, 1994.

[93] Trant, J. The web of collaboration: new technologies and new opportunities, 22 September 2001. Available at http://www.amico.org/docs/papers/2001/.

[94] Tufte, E. R. *Envisioning information.* Graphic Press, Connecticut, 1990.

[95] Tufte, E. R. *Visual Explanations: images and quantities, evidence and narrative.* Graphic Press, Connecticut, 1997.

[96] van Ossenbruggen, J. *Processing Structured Hypermedia — A Matter of Style.* PhD thesis, Vrije Universiteit, Amsterdam, The Netherlands, April 10, 2001. Also available on http://www.cwi.nl/∼jrvosse/thesis/.

[97] van Ossenbruggen, J., Geurts, J., Cornelissen, F., Rutledge, L., and Hardman, L. Towards Second and Third Generation Web-Based Multimedia. In *The Tenth International World Wide Web Conference* [56], pp. 479–488.

[98] van Ossenbruggen, J., and Hardman, L. Smart Style on the Semantic Web. In *Semantic Web Workshop, WWW2002* (May 2002).

[99] van Ossenbruggen, J., Hardman, L., and Rutledge, L. Hypermedia and the Semantic web: A research agenda. Tech. Rep. INS-R0105, CWI, May 31, 2001.

[100] Venkatesh, S., and Dorai, C. Bridging the semantic gap in content management systems: Computational media aestetics. In *COSIGN 2001* (September 2001), pp. 94–99.

[101] W3C. Synchronized Multimedia Integration Language (SMIL) 1.0 Specification. W3C Recommendations are available at http://www.w3.org/TR/, June 15, 1998. Edited by Philipp Hoschka.

[102] W3C. Web content accessibility guidelines 1.0. W3C Recommendations are available at http://www.w3.org/TR, May, 5, 1999. Edited by Wendy Chisholm, Gregg Vanderheiden, and Ian Jacobs.

[103] W3C. XHTML 1.0: The Extensible HyperText Markup Language: A Reformulation of HTML 4.0 in XML 1.0. W3C Recommendations are available at http://www.w3.org/TR/, January 26, 2000.

[104] W3C. Composite Capability/Preference Profiles (CC/PP): Structure and Vocabularies. Work in progress. W3C Working Drafts are available at http://www.w3.org/TR, 15 March 2001. Edited by Graham Klyne, Franklin Reynolds, Chris Woodrow and Hidetaka Ohto.

[105] W3C. Extensible Stylesheet Language (XSL) Version 1.0. W3C Recommendations are available at http://www.w3.org/TR/, 15 October 2001, 2001.

[106] W3C. Synchronized Multimedia Integration Language (SMIL 2.0) Specification. W3C Recommendations are available at http://www.w3.org/TR/, August 7, 2001. Edited by Aaron Cohen.

[107] W3C. XHTML 1.1 - Module-based XHTML. W3C Recommendations are available at http://www.w3.org/TR/, May 31, 2001. Edited by Murray Altheim and Shane McCarron.

[108] Wong, W. *Principles of form and design.* Van Nostrad Reinhold, New York, 1993.