International Journal of Software Engineering and Knowledge Engineering, Vol 9, N0 6 (1999)


# ENGINEERING VIA DISCOURSE: CONTENT STRUCTURE AS AN ESSENTIAL COMPONENT FOR MULTIMEDIA DOCUMENTS

STEPHEN J. MORRIS
City University, Department of Computing,
Northampton Square, London EC1V 0HB, UK
email: sjm@soi.city.ac.uk

ANTHONY C. W. FINKELSTEIN
University College London, Department of Computer Science,
Gower Street, London WC1E 6BT, UK
email: acwf@cs.ucl.ac.uk

Practical problems of multimedia document production require software engineers to provide an effective framework for inter-professional work. The paper distinguishes between abstract and physical media and hence provides the basis for definitions of multiple media and multimedia and the context for reviewing content structures proven in other disciplines. Such structures can act as a guide for production and the notion of the navigable discourse structure provides the essential means for testing content design. Combining these structures in a discourse driven process model and production method facilitates both the design of content and the development of associated software in an ordered and integrated manner, thus avoiding the pitfalls of ad hoc approaches. Investigation and testing of these concepts was effected via two case studies involving the production of two multimedia demonstrations of software engineering tools.

## 1. Introduction

Software engineering has as a primary task the provision of a framework for system development that encompasses the activities and concerns of all potential participants. This paper reports on major aspects of a research project intended to expand the standard software engineering framework so that it will incorporate new aspects necessitated by multimedia document production. We have been concerned, at root, with the incorporation of content generation activities and with all this might entail. How often do we see poorly engineered multimedia documents ? One important cause is inadequate attention to content structure. This can result not only in an inadequate product, but also in an uninteresting range of products. The prevalence of encyclopaedia derived CD-ROMs and other types of digital reference documents indicates that the full range of possibilities remains to be exploited. However the issue of content structure has received little attention, particularly within software engineering. Fortunately it is possible to draw on formal linguistics and discourse studies to provide a basic theoretical foundation, which can then underpin useful practical research.

This content based approach has important practical implications when viewing the process of production, and raises many basic questions. How to define any desired structure ? How to integrate such a structure with the software that will be needed to provide access operations and interfaces ? How to guide such an integrative production process ? How to test the product in terms of this desired structure, either in its final form or at some earlier stage ? How to provide for an evolutionary, even opportunistic, approach akin to that followed in traditional content creation activities ? How to accommodate the working methods of teams drawn from disparate professions ?

We have tried to address such questions within a strong theoretical structure and with practical production in mind. Practical investigations and testing took the form of two case studies of production. The purpose of these experimental productions was to recreate ‘live’ demonstrations of two software engineering tools, the viewer (Nuseibeh & Finkelstein, 1992; Finkelstein et al., 1992) and the System Architect’s Assistant (Ng, 1992; Kramer et al., 1993), as freestanding multimedia documents, the viewer?emo and the SAA Demonstration (SAAD)

respectively. Production used the same multimedia development system Macromedia (formerly MacroMind) Director in two versions (MacroMind, 1991; Macromedia, 1994). The project whose results are reported here offers a detailed and tested method for integration based on a new theoretical framework, practical experience and work in related fields. An overview of the project (Morris & Finkelstein, 1996) and a description of one detailed new technique (Morris, 1998) appear elsewhere.

From the software engineering point of view a multimedia end product will be some form of document, in the most general sense something furnishing information or evidence upon any subject, and will be unusual in having a specific communicative purpose. From the point of view of those who have been responsible for the content of traditional communicative objects such as books or films, the product is supported by a structure of software quite unlike the physical structure of traditional artefacts. Multimedia documents are the product of two separate activities which focus on their design as objects with some communicative purpose and their development as systems functioning via computer software. In order to maintain this important distinction, the term 'design' is restricted to activities related to the creation of document content, and the terms 'develop' and 'development' refer only to the generation of software. Successful production will take place only if the design and development activities are properly integrated. This integration requires both unifying concepts and practical techniques.

Successful integrated production requires a comprehensive view of media types that is applicable to all disciplines involved, which is the subject of Section 2. We summarise alternative and complementary views of multimedia systems, contrasting the technical functioning of systems supporting the manipulation of many media representations, formats or data types, with the communicative functioning of multimedia documents that combine content elements and access operations as systems. Integration also necessitates a understanding of possible content structures associated with communicative objects of all kinds; Section 3 traces the development of ideas of content or discourse structure from their origins in formal linguistics through to very recent definitions of digital genres. In order to support and test communicative functions, we here introduce the novel concept of 'navigable discourse structure' (or NDS) as the means of embedding content structure within a multimedia document system. This concept of an NDS, the subject of Section 4, has enabled the definition of a discourse driven process model for multimedia document production and, on this foundation, a detailed method for production. Section 5 sets out the model and includes parts of the method, illustrated with examples from the two case studies carried out as an integral part of the research. The last two sections, 6 and 7, review related work and present conclusions.

## 2. The nature of media

When observing the practical world of production, it is easy to caricature the participants as belonging to two groups of professionals, those from the software industry understanding only a data processing approach and those from publishing who are married to traditional physical objects and structures of content. Although such a divide is narrowing, there remains an essential need for a set of terms and definitions that will bridge this gap and provide a comprehensive framework for communication.

The purpose of this section is to discuss the nature of the media that may incorporated into any multimedia artefact or system. This is necessary in order to make clear the there are two fundamentally different categories of entity involved. One essential purpose of multimedia software engineering is to integrate the use of these entities into a whole that serves the communicative purposes of its producers. The two categories are the abstract media or systems of signs prepared by the designers of multimedia content and the physical media or data formats manipulated directly by software developers. Software engineers indirectly intervene in both and determine the manner of their coexistence. On the basis of this abstract / physical distinction it is possible to distinguish clearly between multiple media and multimedia and to differentiate between types of multimedia systems and documents.

## 2.1. *Abstract and physical media*

In general use the term medium denotes any kind of intermediate agency, means or channel. In the context of multimedia the term 'medium' is applicable with two specific meanings, setting aside this common usage. Any subject matter being communicated has an associated medium which is its carrier, or vector, in the physical sense. This sense is applicable to air waves for speech, to printing ink on paper for text, to photographic film for moving images, or to the variety of physical substrates that may carry digitally coded signals. The variety and sophistication of traditional physical carrier media (for example the multivarious types which may be used in cataloguing the objects in an art museum), the characteristics of the objects associated with them (for example the many forms which the book has taken), and their social and cultural impact (for example that of photographic film as used in the cinema) form an important area of study in itself. In the second sense, media are abstract; they are agencies for the communication of subject matter. As such they are often viewed as separate sign systems, the most common being natural language text, graphics, still and moving images, natural language speech, other sounds, and other sign systems such as computer languages (Morris & Finkelstein, 1993). In the classic semiotic definition (Saussure, 1983) the sign is a unit of relation between a signifier and what it signifies or conveys, the significate or signified.

Computer based digital media blur, but do not invalidate the distinction between the physical and the abstract. Digital codes provide an alternative means of representing the signifiers in sign systems without superseding them. These codes are not tied to any specific physical medium, but are supported on a new 'virtual physical medium', as created by one of the many kinds of device now in use for digital storage. The potential flexibility of this new type of medium, in terms of the multiplicity of codes or data formats that it can support, makes multimedia possible. However the relationship may not be clear between a medium, understood as a format within the technical arena of software development, and the same medium understood in the discoursive context of design activities.

Abstract media comprise the sign systems whose representations are necessarily carried on some physical medium. The practical use of such abstract media normally occurs within some form of discoursive structure which provides a framework for communication via speech or writing. Physical media comprise the physical substrates which may support any type of mark or variation of physical property that may act as a signifier within a sign system. The controlled manipulation of these physical media has always been a technical matter, whatever their ultimate purpose or use. The distinction between abstract and physical media is important because multimedia document production essentially comprises the manipulation of abstract media elements via their representations in elements of the common virtual physical medium, for brevity now referred to as the 'digital medium'. In this digital medium different formats are often identified as different media. These formats are an inadequate basis for design because they fail to differentiate satisfactorily between abstract media. Many digital formats may represent abstract media with different names and characteristics.

In the arena of discourse, and particularly where linguistic analysis has been applied, conglomerations of heterogeneous codes, which are the *langages* of Sausurre (Sausurre, 1983), displace media. It becomes essential to distinguish between the separate systems of signs used for communication. From the practical point of view such discrimination should allow the document producer to avoid the use of more than one language at a time and protect the user from the resulting confusion. The categorisation of abstract media, now correctly termed, becomes one of :

- *text* - written 'natural' language in any form;
- *graphics* - a two-dimensional representation of information using signs from a monosemic system whose meanings are specified or are known beforehand, or from a notation in the strict sense (Goodman, 1976);

~ *image* - a visual representation whose interpretation involves polysemic elements whose meanings are deduced from the collection in which they appear (Bertin, 1983); alternatively one that indicates nothing other than itself as its own 'pseudo-presence' (Metz, 1974);

~ *video* or *moving pictures* - sequences of graphics or images in animated form as technically defined;

~ *speech* - spoken natural language;

~ *sound* - any other naturally occurring or artificially generated sound proper to its source; and

~ *other sign systems* - numbers and mathematical symbols, computer languages and database entities.

In the technical arena the so-called media, in fact information formats, normally consist of:

~ *text* - all forms of written language;

~ *graphics* - vector based or bilevel representation of an image of a type otherwise drawn with a pen or a similar instrument;

~ *image* - any multilevel or bitmap based image;

~ *sound* - all audio signals (sampled and quantised if digital);

~ *video* - linear sequences of graphics or images that provide the illusion of motion when displayed in rapid succession (usually in excess of twenty five images per second);

~ *numeric data* - all forms including mathematical symbols;

~ *database entities* and *attributes* (in general sense); and

~ *hypertext* and *hypermedia* - a problematic media type consisting of linked elements of other types.

This categorisation is the minimum necessary for discussion of multimedia, but is by no means comprehensive for all systems. Many systems may require further individual definition because of the particular types of format or computational objects used, for example, in the Medusa system discussed below. The discoursive division of abstract media is more appropriate in this study because of the overall design objective.

The relationship between the categories in each field may be one-to-one or one-to-many. For example, text in the discoursive field has an exclusive one-to-one relationship with text in the technical field, while text in the technical field may either be a representation of text in the discoursive field or may be related to some other sign system, for example a programming language. The formats available in the Medusa multimedia architecture (Wray et al., 1993), shown in Figure 1, provide an example. The abstract media shown on the left have a variety of possible relationships, both one-to-one and one-to-many, with the segments and messages that may be carried as part of the bit stream supported within the Medusa architecture, the virtual physical medium shown on the right. MIME (Borenstein & Freed, 1993) provides another example within the technical arena. The Multipurpose Internet Mail Extensions define multimedia additions to the original framework for electronic mail (Crocker, 1982). As in the case of Medusa, this mixture of seven simple and complex media precludes clear and unambiguous relationships with single abstract types.

The distinction between the discoursive and technical arenas highlights a further important difference between design and development, in addition to the contrasting approaches to production resulting from professional and cultural factors. The principal concern of design is the manipulation of abstract media with the view to the generation of content with some external communicative purpose. In development however the concern is with the manipulation of the digital medium with a view to the generation of software which will direct, or control in some way, other elements within the same digital medium.

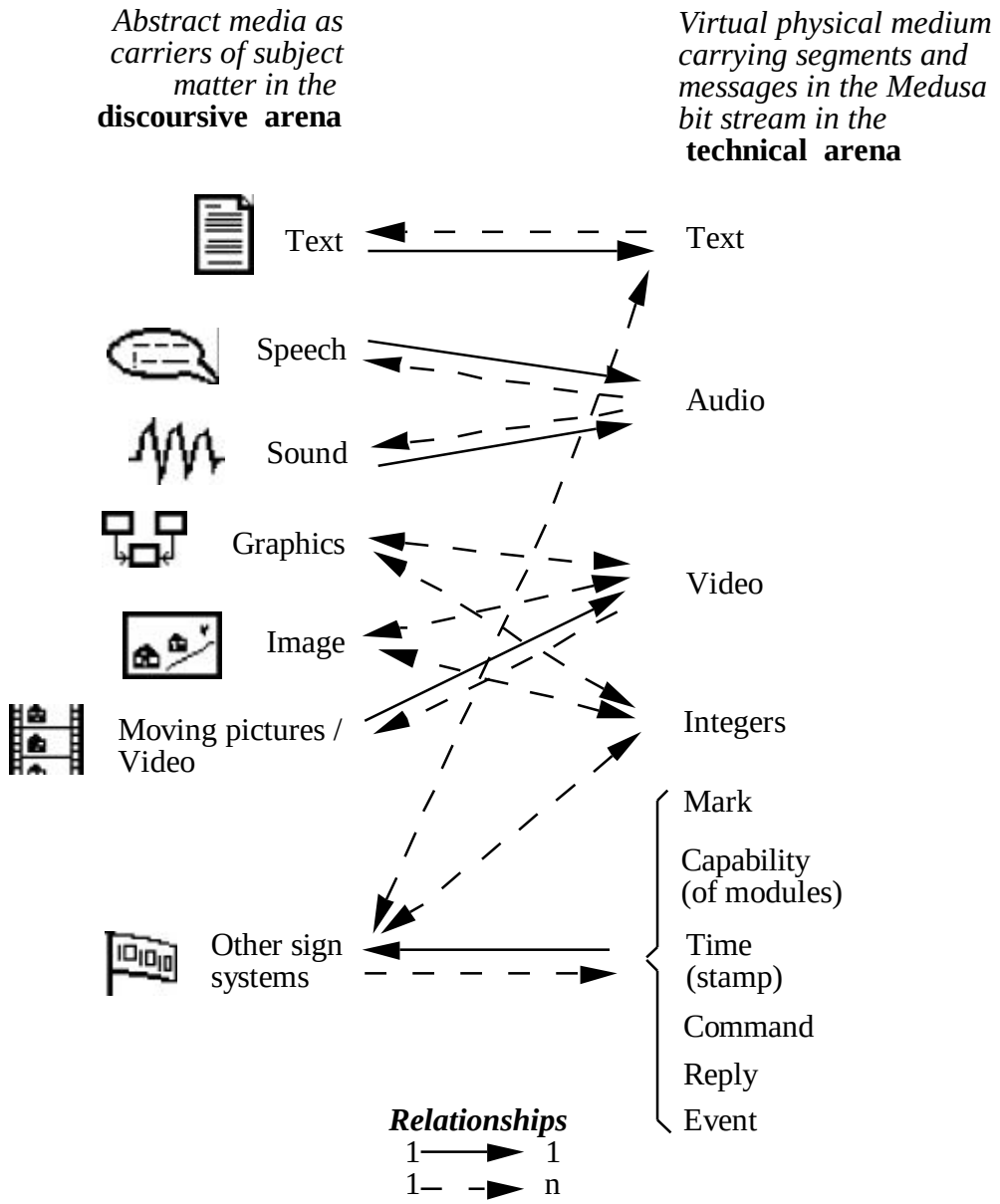



**Figure 1 :** Relationships between abstract and physical media in the Medusa multimedia architecture

Some notion of the distinction between the physical and the abstract is present in other general definitions of media. Blattner and Dannenberg see media in general terms as carriers of information which may or may not have a specific physical form. They consider them to be best understood as multiple layers of encoding, in this context signifying every type of representation from the structure of a novel to the format of a particular typeface. The subject of the representation or coding is the information content. They contrasted medium with mode, whose general concerns are the styles or states of interactivity, the dynamics of interaction (Blattner & Dannenberg, 1992).

Although the single virtual physical medium is a useful concept in making the distinction between content carrying abstract media or sign systems, physical media, quite literally physical carriers or means, remain an important factor for engineers building multimedia systems. This is particularly true in the increasingly important field of distributed systems for the processing of multimedia data. In this field storage, transmission, interchange and presentation media provide the physical means of storing, transmitting, interchanging, and reproducing or acquiring data respectively. The term ‘representation media’ refers to the particular codes used, for example ASCII characters, and the term ‘perception media’ is used to describe the nature of the medium

as perceived by the user (Blair & Stefani, 1997). This last definition comes closest to the concept of the abstract medium or sign system. The primary purpose of multimedia systems in this context is the handling of a variety of representation media, i.e. separate data or information formats, in an integrated manner. Although the means of achieving this integration and its quality is a major concern outside the scope of this paper, it is important to note that the type of system required has very different objectives and characteristics from those of the discoursive system discussed above.

The case of hypertext and hypermedia as a complex combination of data formats is an exception at present because there is no clear discoursive equivalent. This omission is partly the result of the absence of a substantial corpus of such artefacts for analysis. There is also the open question of the essential nature of hypertext and hypermedia as part of a virtual physical medium, or as modes of discourse or styles of rhetoric. A paradox lies in the fact that hypermedial interactivity increases digressive flexibility at the same time as prejudicing the homogeneity and integrity of any traditional discourse structure. But comprehensible discourse structure, however created or assimilated, is a prerequisite for a communicative object (or system) to perform its fundamental communicative function. For those seeking a 'reader-centred' rather than a 'writer-centred' approach to literature study (Landow, 1992), this paradox presents a fruitful opportunity. For those hoping to exploit new technology to enrich conventional modes it introduces a new uncertainty that must be addressed as part of the production process.

## 2.2. *Multiple media and multimedia*

Multimedia, loosely defined as physical objects, or events that combine a number of abstract and physical media, has a long history. By the late nineteenth century Wagnerian opera had already realised the concept of the *gesamtkunstwerk*, the synthesis of all the arts in a single work; contemporary documentaries regularly integrate as many as four distinct abstract media; combinations of both physical and abstract media, for example in collages, have been common. However, without the digital medium, integration can only take place over a limited range of either physical or abstract media. The digital media makes possible any combination of abstract sign systems that can be represented in digital form. In recognition of this important qualitative change, the term 'multimedia' is applied here only to combinations of abstract media within a computer environment carried in the single digital medium. The term 'multiple media' then applies outside the computer environment where different abstract media may each require at least one physical medium as a carrier.

A live demonstration of a software engineering tool, such as the viewer (Nuseibeh & Finkelstein, 1992), and a computer based recreation of such a demonstration, for instance the viewer?emo, illustrate the differences between multimedia and multiple media. Both are composite media objects, with different structures but the same communicative purpose, which is to show what the viewer does and why it is important.

The live demonstration is an example of multiple media. The demonstration incorporates several abstract media including still and moving images, graphics, natural language speech, computer languages, and text. The carriers for these abstract media include printing on paper, sound waves and the virtual medium provided by the computer. The combination of all these physical and abstract media takes place at the time of the demonstration and at the discretion of the presenter. The multiple media are held together by the actions and utterances of the presenter, who provides co-ordination by the selection and timing of these elements.

The viewer?emo is an example of multimedia. It uses the same abstract media as its live predecessor, but they are carried together in a physical medium holding digital signals. This single physical medium is created, stored, manipulated and displayed exclusively in the computer. Integration of the abstract media takes place at the discretion of the developer and the resulting structure determines the manner of presentation to the audience.

Any single abstract medium, or combination of such media, is feasible in multimedia. Any abstract medium can find a representation in multimedia if there is a digital format suitable

for the corresponding sign system. Any combination of abstract media elements is also theoretically feasible because each element is represented by data blocks that can be linked by pointers. The exploitation, as a means of presenting data, of purposefully created links explicitly represented to users gives rise to the concepts of hypertext and hypermedia.

### 2.3. *Multimedia documents as systems*

Multimedia technology now supports both networked and broadcast communication, thus facilitating the general communicative activities discussed in the previous section. The technology also includes multimedia applications that allow the creation of individual multimedia objects. The potential beneficiaries from this second branch of multimedia technology, the subject areas or fields previously constrained by analogue technology, will be the primary users for multimedia objects. The term ‘document’ will be used for these objects because they will have specific communicative purposes and because common usage of that term already embraces a range of functions and a variety of media. Like traditional documents their multimedia successors will be able to function in a variety of ways: showing, pointing out or proving something, furnishing information, evidence or proof. In the most general sense they will function as abstract communicative objects carried in some physical form.

When considering production of multimedia, those responsible naturally turn to existing documents. In any general statement about developing such documents, for example ‘We are going to produce an X about Y’, X refers to some type of communicative object and Y is its general subject matter, usually not a specific title. Conventional terms used for X carry with them assumptions about three factors: the physical carrier medium being used, a single dominant abstract medium (or a limited range of them), and the physical mechanism providing access to the physical medium and underlying support for a relevant mode of discourse.

In traditional artefacts the physical medium will comprise standard media elements. The standard elements of the book have two characteristics, printing on paper and sheets of common size. Printing on both sides of each sheet creates two pages, *recto* and *verso*. Size, weight and reading distance are among the factors which limit within broad margins the number of sheets in any particular object. In all cases display size is fixed at the time of production. No separate access mechanism is required. The simple operation of moving a single sheet, 'turning the page', allows the reader to progress forward or backward, to the next or previous pair of pages in the fixed sequence. Binding sheets in a block with an easily sensed edge allows movement to the beginning or end of the book, or to any arbitrary approximate position. It should be noted, however, that some printed content, usually text, is required to signal the correct orientation of the binding, most often vertical, and that the particular natural language represented in the text will determine whether left or right pages are turned during reading.

Cinematic moving pictures employ photographic film, each frame having a uniform size. The basic element comprises a sequence of frames, sufficient in number to give the necessary illusion of movement when displayed successively at a fast enough rate, normally twenty-four per second. This basic element is sometimes called a 'clip'. The mechanism for display is an essential and separate prerequisite and there is a single mode of operation for viewing. The proportions of the displayed image are fixed, but its exact size may vary according to distance between source and display surface. The access mechanism provides a single operation, inherited from the basic element of the medium, which is to display each frame in their fixed sequence until the mechanism is stopped. The reverse is only permitted with moving images recorded on analogue or digital video tape. Only in the digital case is it possible to implement an operation allowing movement to any defined arbitrary point. No moving image mechanism allows movement to an approximate position. In the case of the book there is no need for a separate display operation.

If, however, the document X is a multimedia document, then this will necessarily imply that the physical medium is exclusively digital, possibly even in preliminary stages of development, secondly that the abstract media may be multiple both in their form and in their

relationships, and finally that no assumptions can be made about the access mechanism. As a consequence there is no standard mechanism to support any particular mode of discourse in multimedia objects. This is the case both during development and on completion. Replication of any conventional discourse structure will require the construction of such a mechanism. Unlike any conventional artefact a multimedia document may accommodate any single abstract sign system or combination. It will lack any standard access operations, and it will be disassociated from any particular mode of discourse or culturally defined structure.

The presence of the two major components, media and software, as an organised group means that the document will meet the most general criterion for designation as a system. It will also meet more specific requirements for: a clear boundary, an operating environment, a homogeneous character and some emergent property unique to the system as a whole rather than to any of its individual components (Checkland, 1981; Checkland & Scholes, 1990). In addition to diverse potential media, multimedia offers a potentially infinite variety of software operations, and hence a variety of external structures. This external structure defines the user interface and provides the clear boundary required of a system. The users are an audience and, in communicative terms, the environment within which a system must operate. The purpose of the document can only be achieved by its operation as a single entity, homogeneously, within this environment. The document will also have communicative properties 'emerging' from the available combinations of media elements and software operations. The necessity for such emergent properties, in this case being properties of the multimedia document as fully realised, now plays a major part in systems thinking (Kronlöf et al., 1993).

As finally completed the multimedia document will need to be effective both as an abstract communicative object and as a system combining media elements and software operations. This double potential and joint requirement are fundamental determinants in the production process. A separation between physical and abstract categories has a number of advantages. It opens the way to the consideration of the purely abstract discoursive structures which underlie most communicative objects. It also facilitates a clear demarcation of the group of issues that relate only to physical carrier characteristics, for example digital storage and retrieval. In combination the two categories serve the very useful practical function of facilitating the definition of complex structures employing particular physical media tied to specific sign systems or semiotic languages. Such languages may themselves be associated with individual modes of discourse and with culturally defined macrostructures for such modes.

Having established the relationship between abstract and digital media, it is possible to consider the manner in which such media might be deployed to provide a coherent communicative whole. The study of rules and structures with this purpose has a long and complex history, stretching from Aristotle to contemporary researchers of genre and register. The next section reviews developments in linguistics and many related fields, initially in order to clarify the plethora of similar terms currently in use.

## 3. Content Structure

Faced with the practical problem of specifying and then designing any multimedia artefact, the software engineer needs to be able to define a suitable framework. If the artefact is to serve some communicative or discoursive function, this framework must reflect some form of content structure. Development of such structures is not a commonplace of software engineering at present. It has, however, been the subject of much study in the past. The results of these studies offer practical support to software engineers in the form of content frameworks of various kinds.

The purpose of this section is to trace the development of ideas related to content structure, that is how abstract media elements may be organised to some communicative purpose. It is necessary to begin with formal linguistics and traditional views of discourse and then to follow the impact of semiotics in shifting the emphasis to social and cultural context. Starting with general pragmatics, part of the triadic division of formal linguistics including the syntax and

semantics of natural language, we conclude with a diversity of potential genres of multimedia documents, each pragmatically defined within a community of use or interpretation.

### 3.1. *Formal linguistics*

A traditional division of the field of linguistics (Morris, 1938 ) takes *syntactics* to involve the purely formal study of the relations of signs to one another, *semantics* to involve the relation of signs to the objects they denote (or may denote) and *pragmatics* to be the science of the relation of signs to their interpreters, dealing with the psychological, biological, and sociological phenomena which occur in the functioning of signs. Recently pragmatics has been described more succinctly as the study of linguistic acts and the contexts in which they are performed, one of its central goals being to characterise the features of the speech context which help determine which proposition is expressed by a given sentence (Stalnacker, 1972).

Beyond helping to determine the assignment of lexical and sentential meanings, the context provides a specification of the universe of discourse. The wider field of discourse, that is the field encompassing any construct larger than a single sentence, is important to linguists if only because it begs one basic question, whether the essential communicative function of dealing with such larger units, discourse competence, is part of linguistic competence or something else. The analysis of discourse has generally fallen under pragmatic competence, given the general view of language that divides language into (truth conditional) *semantics* and (non truth conditional) pragmatics (Prince, 1988). However the enormous diversity of communicative purpose that any stretch of language may serve has meant that discourse analysis has been a subject of interest to many other disciplines including ethnomethodology, computational linguistics and literary theory, to name only three examples. One central contention of this paper is that it is also important to software engineers.

From the purely linguistic point of view the study of discourse is concerned not just with the properties of linguistic representation but also with the non-linguistic factors that determine what message is conveyed by the linguistic form and whether it counts as an acceptable contribution to the communicative enterprise (Blakemore, 1988). As a consequence the notion of grammar needs to be extended so that it can account for the well-formedness of discourse, so that it can predict whether a given utterance is an acceptable contribution to a text. The idea that elements of a well-formed discourse are bound together by principles of connectivity or textual unity is fundamental to most approaches. Such principles are potentially significant to software engineers because they might provide the framework for understanding (and producing) the structure of any form of text or discoursive document. An multimedia artefact with communicative intent would form such a document.

### 3.2. *Traditional coherence of content*

Some authors assume that coherence of content can be explained in terms of coherence relations between propositions. One view of this type (Johnson-Laird, 1983) argues that there are two levels of representation for discourse : a superficial propositional format close to linguistic form, and a mental model that is close to the structure of the events or states of affairs that are described in the discourse. Coherence relations, according to this argument, hold between the latter rather than the former. This view has found favour in the field of human computer interaction because of the direct analogy between conversational discourse and dialogues between user and computer concerning system and application states.

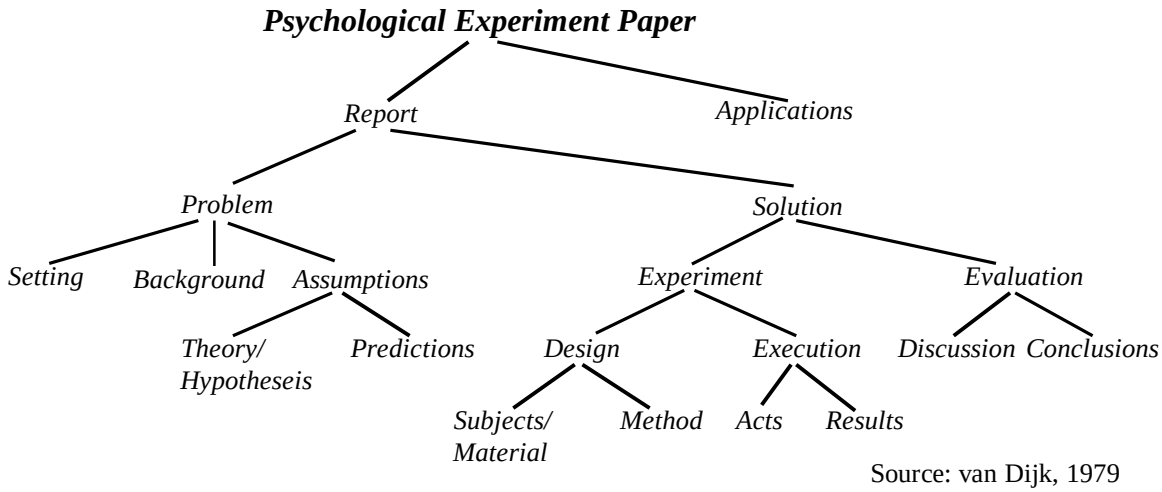



**Figure 2 :** 'Superstructure' of a psychological experiment paper (van Dijk, 1979)

An alternative view (Van Dijk, 1977; 1997a) suggests that in order to account for the well-formedness of discourse it is necessary to include in grammar a pragmatic component with rules relating sentence-context pairs to interpretations at a global level of semantic description or 'macro-structure' of propositions. He appeals to the notion of a frame, in the sense of a representation of the knowledge people have of stereotypical events and structures, and proposes the notion of discourse 'superstructures' (van Dijk and Kintsch, 1983). Figure 2 shows an example. Such structures offer a powerful composition tool that has yet to be fully exploited in multimedia development. Our work provides a software engineering framework within which they may be exploited profitably.

Approaching the problem from a third angle, the analytical method based on rhetorical structure theory (RST) offers a fine grained approach to the examination of semantic relationships within text (Hovy & McCoy, 1989). Production of the viewer?emo showed its applicability to the analysis of scripts for multimedia productions. The basic premise tested by such work was that it is typical, although not universal, for discoursive text to be hierarchically and functionally organised. It underpins the principle of "constructivity" (van Dijk, 1997a). Discourses are constructive in the sense that their constructive units may be functionally used, understood or analysed as elements of larger ones, thus also creating hierarchical structures. This applies to meaning and interaction as well as to forms.

The principle of "constructivity" accompanies that of "sequentiality", the principle that the accomplishment of discourse is largely linear and sequential in the production and understanding of talk and text. This principle is essential to the maintenance of continuity, particularly that of context (Blakemore, 1988). Within planned discourse there is continuity of context. In such 'planned' discourse the aim of the producer is to maximise relevance over the discourse as a whole, whereas in 'unplanned' discourse the goal is simply to make each utterance relevant. The idea of maximising relevance of coherence as a whole has not been a major concern of software engineers, in the sense that a user is usually offered by a system a range of functions which may be exploited in arbitrary order, or not at all.

### 3.3. *Discourse types*

The issue of coherence or relevance of a software system output has, in the discoursive sense, been relatively simple to address until the advent of multimedia documents. This has been the consequence of the relatively simple nature of such output from a discourse structure point of view. Much is simply data that is descriptive of some object or entity within the system domain. Explicit description is the principal communicative function performed by the large class of computer applications designed to store and retrieve information in databases. Description

renders explicitly the essential properties of an object for their own sake and for no other purpose.

Description is one of the major types of discourse in a traditional classification that segregates discourse, by convention and long practice, into a number of modes: narrative, description, argumentation, exposition and other much less used forms such as lyric poetry (Chatman, 1990). This categorisation in effect provides a general division according to communicative purpose. Production of a discourse in the modes of argument or exposition will involve the use of language, whether in verbal or other forms, so as to persuade or influence others. To treat a computer based document as a communicative object also presupposes that the computer itself should be treated not as an independent entity, but as an intermediary, itself a sophisticated medium of communication in the most general sense (Winograd & Flores, 1986; Andersen, 1990).

Argument, for the purpose of discourse, is defined as any process of reasoning intended to establish or subvert a position. Although this would encompass types of argument based on formal logic, the emphasis here is on the informal types commonly used in general intercourse (Perelman & Olbrechts-Tyteca, 1968). Argument and exposition may overlap because the latter may subsume the former. Argument presupposes a difference of opinion, while exposition merely requires an absence or confusion of opinion, whose clarification implies an argument in favour of its own preferability (Chatman, 1990).

Narrative is a sequence of events characterised by plot, character and setting. It is unique among discourse types in having a 'doubly temporal logic' entailing both an 'external' movement through time, the duration of presentation, and also an 'internal' progression, the duration of the sequence of events that constitutes the plot. Chatman also emphasises the importance of double functions not only of media but also of discourse types, which may form an underlying structure or an overriding surface representation (Chatman, 1990), for example the surface narrative form of fables or of television commercials which present arguments in clear stages. Hierarchical structures for both narrative and argument have been generally agreed, but with formal variations (van Dijk, 1979).

### 3.4. *Extensions beyond text*

One significant recent shift has been the incorporation of additional sign systems into the field of these discussions, the most prominent examples being the use of "multi-modal" texts incorporating visual images and written language (Kress et al., 1997). In many fields of commercial and academic interest the use of particular media is much more than a coincidence of general activity; there is a close link between subject matter and particular media. In the base case primary source material exists in a variety of media. Ethnographers and art historians must examine contemporary and historic objects created in the widest variety of media. They must also employ different media as essential records of both objects and events. The value of different media remains in the subsequent stages of their work. At the same time traditional technologies narrow significantly the range of media available for the communication of results, possibly until there is only printed text. All fields employing social science research techniques may suffer similar impediments, including studies of human computer interaction and software requirements elicitation. In such cases multimedia offers the possibility of new kinds of communicative object that do not suffer from the physical constraints imposed by older media. In any field where a variety of media may better represent any part of the subject matter, multimedia offers this important potential benefit. In general any communicative activity which demands the use of more than one medium may benefit from multimedia (Morris & Finkelstein, 1996). This trend requires a significant shift of emphasis away from that placed on text and speech.

### 3.5. *Broadening of 'context'*

The traditional classification of discourse types discussed above, whilst retaining a restriction to the domain of text and speech, also inherited from linguistics the notion of context incorporated within the definition of pragmatics. A second recent shift has involved a significant widening of the scope and definition of context. To some authors the study of discourse should now deal with all the characteristics of the social situation or of the communicative event that may systematically influence text or talk (van Dijk, 1997a). The expansion of the field in this direction implies that linguistic pragmatics in their original sense could become a sub-discipline of discourse studies.

As a consequence of this expansion and the resulting incorporation of many types of communicative activity, for example conversations, advertisements, poems and news reports, it has been necessary to attempt to redefine discourse types in a more fine grained way, commonly using the term "genres". The combination of this expansion with the increasing use of the digital medium has now created interest in the study of genres of digital documents. A recent definition considers a digital document to be any socially and contextually complete semantic unit of communication - including text, video, audio, hypermedia, multimedia, and computer-mediated communication - which is created, stored, and transmitted via digital media (Yates and Sumner, 1997). Although this definition seem to miss the essential issue of multimedia, it highlights the breadth of colonisation. The term 'register' introduces a further refinement of the idea of context (Eggins & Martin, 1997), based on the assumption that texts are not neutral encodings of content but semiotic constructions of socially constructed meaning. The two main dimensions of variation between texts viewed in this way are 'register' or the context of situation, and 'genre' or the context of culture.

In this newly expanded territory genres become heterogeneous categories of such artefacts, which can be characterised by formal properties of format, language or presentation, by schemes of content organisation, by social function and by the communities who use and interpret them (Nunberg, 1997). Examples now become as diverse as a special tailor-made report generated by a database system, a project meeting, a social science research paper or even a table of contents or index on a WWW page (Crawston and Williams, 1997).

From the point of view of the work reported here, schemes of content organisation, essentially hierarchically and functionally organised, will be an essential component in any principled approach to the production of multimedia documents. Whilst accepting the interest of this recent lower level diversification, it is important to note the higher level definitions of basic discourse types and structures provided by Chatman and van Dijk, all of which will be implicit in some way in any specific lower level genre. Further structures still have to be examined in depth pragmatically in order to provide a wide range of heuristics. In the interim we advocate the specific inclusion of discourse structures within multimedia production and the use of navigable discourse structure, the subject of the next section, as an essential component of any complete software engineering approach.

## 4. Navigable discourse structure

This section introduces the idea of a navigable discourse structure (NDS), the actual discoursive structure generated in, and by, a completed multimedia document. Satisfactory production requires that the participants should be able to distinguish between any abstract discourse structure which might provide a basis for a design of the object and the discoursive structure actually realised within the computer environment. Introduction of navigable discourse structure provides the necessary bridge between communicative object and software system and between intended and actual structure, opening the way for a dynamic approach to document design and development.

The navigable discourse structure of a document can be treated in two ways, as a property of the document that emerges consequent upon its other characteristics and without the direct

intervention of its producers, or as a property that is directly reliant upon the existence of particular elements and upon specific interventions by the producers. The view taken here is that the nature of the multimedia artefact, without standard access operations and disassociated from any particular mode of discourse or culturally defined structure, requires the latter approach. The practical purpose of a navigable discourse structure is to represent the discourse structure actually implemented in a document. Without such a representation it is impossible to compare actual with intended structure and, from this comparison, to decide how production might proceed. Like the representation which may be used for intended discourse structure, discussed above, this one will consist of a hierarchy of components.

### 4.1. *Definition of NDS*

Any practical techniques for determining navigable discourse structure will hang on the framework provided by a general definition. The purpose of this definition is to fix the essential elements of the structure and to show how it will fit within any model of production model. An NDS is defined as:

*A set of the necessary order relationships, between abstract media elements, required to support effective discourse, combined with the mechanism for the user to move between the associated physical elements according to these discourse relationships.*

This definition builds on the following assumptions derived from the preceding discussions:

~ specific communicative purpose requires an associated discourse structure;

~ components of a discourse structure will be related to each other by at least one order that renders the structure effective;

~ abstract media components of the structure will be carried by elements belonging to a physical medium, in the case of multimedia the digital medium exclusively ;

~ maintenance of discoursive relationships requires mechanisms which are both appropriate to the structure and operational within the physical medium.

Construction of an NDS for any communicative object requires four major elements:

~ a set of components which belong to an intended discourse structure, and are realisable as abstract media elements,

~ the relationships linking the abstract media elements with their physical media representations,

~ the operations available for moving between elements in the physical medium, and,

~ the application of operations to spreads generating a sequence from opening to closure of discourse and to any other permissible sequences via the controlled addition of linkages.

A set of elements belonging to a physical medium, although not a part of the discourse structure in itself, is an essential prerequisite for an NDS, in fact its *raison d'être*. This is true both in traditional communicative artefacts and in multimedia. In a multimedia artefact the basic elements will possess a set of minimum properties reflecting the flexibility of the digital medium. The set of abstract media elements belonging to a multimedia artefact will serve the same functions as the set in multiple media but within a single physical carrier, restricted only by available storage capacity and chosen display size and without fixed relationships of position.

The effectiveness of any multimedia artefact as a communicative object will depend on the effectiveness of its structure as discourse. The smallest defined units of discourse in speech (and text) comprise pairs of clauses, following a temporal sequence for a narrative (Labov, 1972) or fixed in the relationship of datum and claim for an argument (Toulmin, 1958). The largest units are the proposed 'superstructures' used pragmatically as schemata for discourse with particular practical purposes. Such schemata, as already discussed, employ composite general components in an hierarchical relationship. All units comprise sequences of utterances and have internal inter-relationships fixed by their order and by the discourse structure, in whatever manner defined.

For a multimedia artefact these relationships will define firstly how each component of the discourse structure finds a representation in an abstract media element, or elements, and secondly how these abstract media are linked to the corresponding physical components of some spread presented to the user. Final presentation must accurately reflect the components of desired discourse structure via the representation of suitable abstract media elements. Operations available as part of the presentation must correspond with those required for representation of the internal relationships of a discourse structure. The form of this structure, an emergent property of the presentation, will be the NDS of the multimedia artefact.

### 4.2. *NDS within production*

Accommodation of the relationships discussed above allows an essential extension of an interim model of document production (Morris & Finkelstein, 1996; Morris, 1998). This representation incorporates three levels of generality (or abstraction) at which production takes place, with the addition of a level for the NDS derived from the final presentation. An internal discourse structure may be ‘composed of’ components, each of which is ‘represented in’ a single abstract medium element or composite of abstract media elements. The abstract elements will themselves be 'presented individually' or 'presented jointly' as part of the physical media element that presents the audience with a final display to the audience, the 'spread' defined earlier. Finally the components of the NDS are represented in the spreads and have internal compositional relationships of the same type as those in the initial discourse structure. With the addition of this fourth level showing the actual NDS as opposed to the originally intended discourse structure, this structure is open to an alternative interpretation, as a representation of the stages of production rather than their levels of abstraction. It is this interpretation which is developed in the next section, which offers a comprehensive model of the production process.

Realisation of any discourse structure requires physical display of abstract media elements in a sequence that conforms with the required sets of discourse relationships. Traversal of the sequence by an audience depends on the availability to users of suitable operations that will facilitate progress in accordance with the same relationships. Implementation of these operations provides an access mechanism, and the association of the operations with particular media elements, their application to spreads, will result in the generation of the NDS.

The standard operations associated with traditional objects such as the book and film have already been discussed in Section 2.3. In the case of a multimedia artefact the nature of the operations is limited only by the functionality of the programming language available to the artefact, either as one of its own media elements or external to its hardware platform. It is possible, however, to distinguish between two general groups, static and progressive. All operations are related directly to elements of display and only indirectly to the components of the discourse structure that the elements may represent.

Each abstract medium requires a static operation to display or play a basic element. Static images and graphics only need an operation to show a basic element in its entirety for an unlimited period of time. In the case of moving images the display period is fixed and is complemented by an operation of continual replacement at uniform intervals. Any progressive operation allowing the display of another segment of images is separate from the replacement activity within the play operation. In the case of text the progressive operations of moving forward or back mirror those of the book, except that the ‘page’ or spread size is not fixed and there is no equivalent to *recto* and *verso*. Movement to an arbitrary approximate position is not essential for sequential discourse, although it may be considered a useful adjunct, as it is in the case of the book.

The set of twelve operations incorporated into the viewer?emo includes examples of both groups:

*static:*

- hold display only for specified period,
- hold display until movement operation initiated,

~ move forward through a segment of moving images,
~ combinations (4), at the discretion of the audience,
of text or speech elements, or both, with each still image.

*progressive*:

~ step forward in sequence of spreads,
~ step back in sequence of spreads,
~ move back to the initial frame of a segment of moving images,
~ movement to designated points including 'start',
'finish' and initial frames of 'sections',
~ switch to the 'start' of subsidiary modules
(with same set of operations) and return to the position
where the 'main' module was left.

A distinction can also be drawn between these operations, which define which media elements are displayed and when, and the operations (and actions) that allow the user to activate them, the user interface. In the case of the viewer?emo this interface involved a combination of keyboard and menu actions.

An NDS is completed by the application of suitable operations to every possible spread. Implementation of all the basic operations should facilitate the traversal of any intended discourse structure. There must be elements presenting every discourse component, and suitable sets of operations must be attached to every element of the artefact. The definition of these combinations of elements and operations will control the possible flow of the discourse. Suitable sets of operations will allow the user to traverse the intended discourse structure. The default sequence will be a depth first traversal of the discourse structure, if it takes one of the basic hierarchical form discussed earlier.

### 4.3. *Identification of NDS*

A framework for NDS identification must be able to serve two basic functions, recording and analysis. The recording function can be broken down into three tasks, identifying the media elements included in a document, the combination of these media elements in spreads, and the association of operations with spreads. Analysis using a form of spreadsheet serves this function. The analytical function demands an initial recognition of which spreads act together as discoursive components, followed by the combination of these components in a recognisable discourse structure. Figure 3 shows a section of the spread analysis sheet designed to fulfil this function. The general purpose of the spread analysis sheet is to aid production. In particular it provides a means for showing how media elements are associated with software operations. Thus it facilitates examination of the structure of the document as a communicative object. This examination will produce, by the means suggested below, a representation of the actual structure to be compared with that intended when design and development commenced.

The fourth column of the spread analysis sheet contains symbols representing the discourse components with which each spread is associated. The first step in their derivation is the identification, by examination of the attached operations, of those individual spreads or groups of spreads that are not accessible separately and therefore cannot independently form components or sub-components. Inaccessible spreads are denoted by a pair of lines, ||. Those that are directly reachable at some point in the discourse are denoted by a pair of square brackets, []. This square bracket notation may apply to a single spread, for example spread 14 in Figure 3, or may be used to enclose a composite component of any size, for example spreads 7-11 or the substantial “designer” section of the document beginning at spread 7. A component comprises a separately accessible combination of media elements, or components, which have independent meaning within the discoursive context.

| **Characteristic media element (*sound*, {text ref #}, G *or I thumbnail)*** | **Other media elements** | **Generic and specific operations** | **Discourse components [name ], [], \|\|** | **Spread (def. seq.)** |
|---|---|---|---|---|
| *The start-up window of the viewer defines the scope of the tool. {t4}* | T: t4<br>I: D11 | g 1-5<br>•<br>"start"<br>408 | [intro [[] | 4 |
| I:  | T: t4 | g 1-5<br>° | \|\| ] | 5 |
| *The user can either be a Method Designer, who chooses and constructs the templates that make up a software engineering method, or a Method User who ...t6}* | T: t6<br>I: D12 | g 1-5<br>•<br>120 | [] ] | 6 |
| *If you click on the Method Designer button, you are presented with a Template Browser. Essentially this allows you to ...{t7}* | T: t7<br>I: D13 | g 1-5<br>•<br>"designer"<br>612 | [designer<br>[methods<br>[[] | 7 |
| V:  | T: t7 | g 1-5<br>° | \|\|] | 8-11 |
| *Clicking on Load Template Database shows the available methods, or collections of templates.{t12}* | T: t12<br>I: D14 | g 1-5<br>•<br>348 | [[] | 12 |
| *I:*  | T: 12 | g 1-5<br>° | \|\|] | 13 |
| *The method called My Method is for demonstration purposes. It contains three templates or ... {t14)* | T: t14<br>I: D15 | g 1-5<br>•<br>696 | [] | 14 |

**Figure 3 :** Section of Spread Analysis Sheet for viewer?emo v 2.0

On the largest scale a whole document is a single discoursive component. Applying the notation to a traditional film, it would appear simply as [ || .. || ], a single accessible component made up of a sequence of frames each inaccessible individually. Any document with a degree of 'restricted access' can be notated by the same means, but the complete representation, the entire contents of column 4 from the spread analysis sheet, may appear complex. The notational symbols used, the lines and brackets, were chosen both because they are simple and because they are available on all standard keyboards, allowing them to be typed as well as drawn.

A linear representation of nested discoursive components and sub-components can subsequently be rearranged to show a complete discoursive hierarchy and thence a simplified structure (Morris & Finkelstein, 1996). Any one of these forms represents, at different levels of detail, the navigable discourse structure that has been generated in the document via its combination of media elements and operations. Generation of at least one comprehensive NDS within a completed multimedia document is an essential prerequisite for successful achievement of its communicative purpose. This section has provided a definition of the components required to initiate construction of an NDS, a framework for recording and analysing the instantiated components and a technique for representing the NDS as actually generated. The concept of an

NDS, with its intended and realised forms, will allow the accommodation of dynamic design activities within the production of a computer based system. The manner of this integration is the subject of the next section.

## 5. Process model and production method

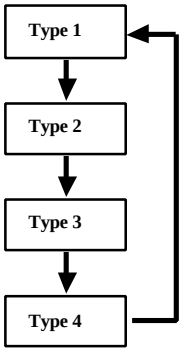
In the absence of any appropriate production method, major practical difficulties arise because of recurring uncertainty about the order in which activities should take place. Production of the viewer?emo showed specifically that at any particular time there would be a choice for the participants from a limited set of alternatives. The first alternative consists of work on a particular media element, either in isolation or in the particular context of elements in other media. Both possibilities fall within the scope of the designer role. The second alternative is the development of some part of the access mechanism that will allow a user to inspect a spread at an appropriate moment. This is part of the software development role. The third major alternative is the full, or partial, integration of media elements with the access mechanism.

Uncertainty about the ordering of production activities raises two issues: the need for methodical, stage by stage guidance and the absence of a dynamic model representing the complete process. Production activities require a clear ordering which safeguards the differing approaches of the participant roles, distinguishes between their individual products and facilitates their integration. This section presents a model for multimedia document production, a process driven by the need to generate a discourse with a specific communicative purpose. Such a model shows its practical value by providing a framework for a production method which offers detailed guidance. The section concludes with a brief overview of the production method tested during the project (Morris, 1996).

Any comprehensive description or model of multimedia document production, or any kind of relevant guidance, should differentiate between and accommodate three distinct roles (design, development and editing) and the activities associated with each, and must reconcile their differing approaches (dynamic, structured or mediating). More specifically, production activities require a clear ordering which safeguards these differing approaches and guarantees their integration. As a guide to meeting these requirements, the differentiation between model and method offered by Boehm provides a useful starting point (Boehm, 1988). A process model should provide a generalised representation of how development takes place and explain the purpose and significance of its constituent stages, both individually and in relation to each other; it should define the order of stages and establish the criteria for transition between them. A development method should provide detailed directions which respect the framework imposed by a particular process model. Consequently it needs to show how to proceed within each stage and how to represent all its products.

The roles of designer and developer raise general questions about the nature of their working processes, in particular the contrast between the ‘dynamic’ approach of the former and the ‘structured’ approach of the latter. The difference between these approaches is characterised by contrasting relationships with the end product. The ‘dynamic’ views the final nature of the product as uncertain until it is deemed complete by the designer; the ‘structured’ requires a much greater degree of predetermination and external verification.

The model has four essential characteristics: a basic four stage cycle, repetitions of this cycle, inter-cycle relationships of comparison, transformation and reconstruction, and an interlacing of cycles depending on these relationships. These characteristic are shown in diagrammatic form in the four sections of Figure 4. A complete view of the model appears as Figure 5.



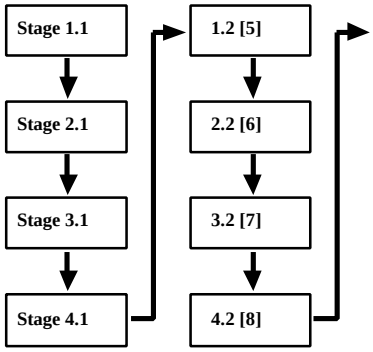
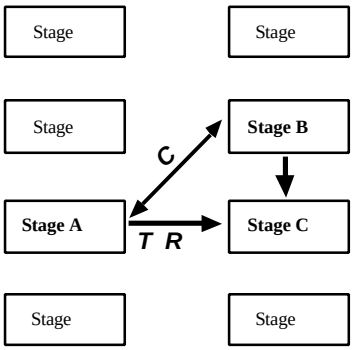




**Figure 4a** (left) : Basic four stage production cycle

**Figure 4b** (right) : Repetition of basic cycle as production progresses



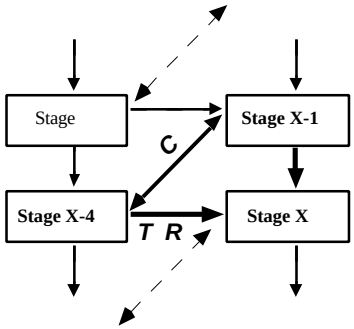




**Figure 4c** (left) : Inter-cycle relationships between stages involving **C**omparison, **T**ransformation or Reconstruction

**Figure 4d** (right) : Interlacing of cycles as stages are repeated

The basic cycle (Figure 4a) consists of four stages, each of a different type. The stages correspond to the levels in the four level representation of production outlined at the start of Sec. 4.2. The vertical arrows indicate the flow of stages and stand for the intra-cycle relationships between stages. Cycles are repeated in such a way that the fifth stage, for example, is the second of the first type (Figure 4b). A square bracket notation provides the means of identifying all the stages of the same type in successive cycles. Thus [6] stands for stages of the second type, numbered 6, 10, ... , N-3, where N is the number of the final stage. This terminating stage is a special variant of a stage of the first type. There are also important inter-cycle relationships (Figure 4c). A comparison (C) takes place between Stages A and B as the basis either for a transformation (T), or for a reconstruction (R) of Stage A, thus forming Stage C. Stages A and C are always of the same type, but B belongs to the type that precedes the type of A and C. This combination of characteristics results in a complex interlacing of stages and cycles (Figure 4d) in which any stage numbered in order as X, whatever its position in its cycle, is related to its predecessor, Stage X-1, and to the last stage of its own type, Stage X-4.

The objectives of the model are as follows:

***(i)*** to incorporate discourse structure as a major determinant of content and system structure,

***(ii)*** to recognise the nature of the multimedia document as an artefact disassociated from traditional physical media and not bound to culturally defined structures,

***(iii)*** to provide basic ordering of activities that will support detailed direction of design and development activities to suit particular production circumstances,

***(iv)*** to ensure that production may be reviewed at regular intervals with specific editorial objectives in mind, and,

***(v)*** to integrate dynamic design of content, and a means of representing it, into a staged production process.

The process model proposed here provides a generalised representation of how production takes place. This representation employs components, stages, cycles and activities. The four components are: ***1)*** discourse structure, ***2)*** abstract media elements, ***3)*** presentation spreads and operations and, ***4)*** navigable discourse structure.

Each component has a corresponding stage during which it changes state and gives its name to one of the four types of stage on the basic cycle described above. Generation of the four components, in sequence in their stages, completes one cycle of production. A combination of three activities, comparison, reconstruction and transformation, causes components to change state. A diagram showing all the elements of the model together appears as Figure 5. In this diagram the rows represent the stages and the states of the components, named on the left; the columns contain alternately the stage numbers and the production activities that cause components to change between cycles.

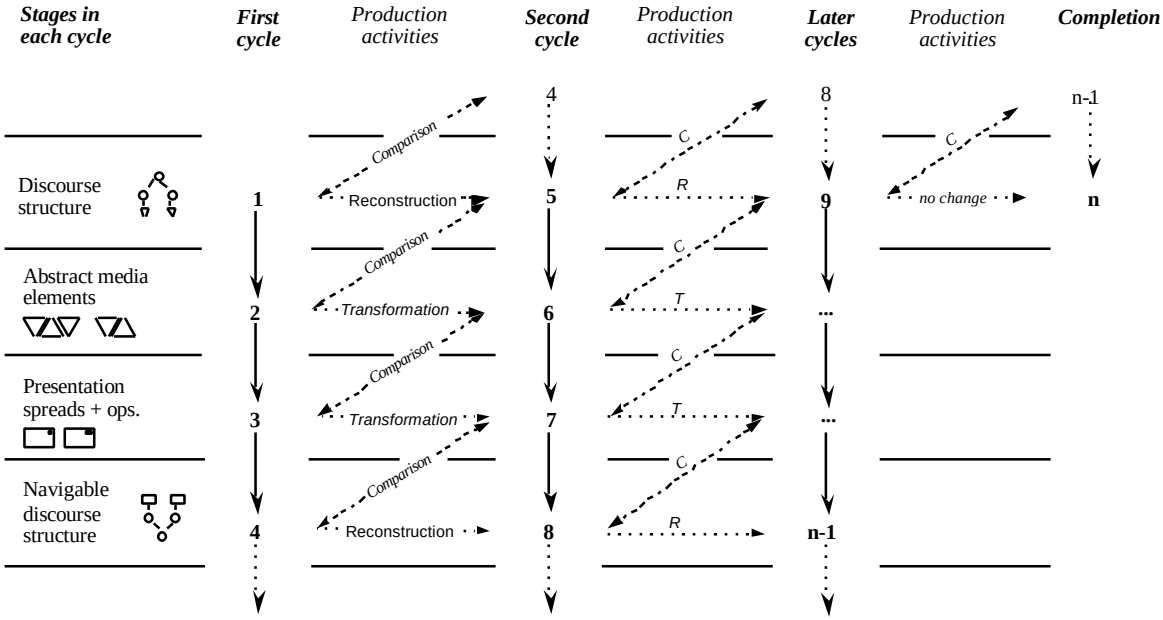



**Figure 5 :** Discourse driven process model for multimedia document production

The initial cycle of the model contains four initial stages of different types, shown in the first column of Figure 5. The discourse structure guides the intended structure of content. It has an hierarchical form and functional organisation and, as such, may be represented as a tree. The leaves of this tree become the instantiated abstract media elements that control realised content. Together these form a resource base which also includes supportive interim elements. The presentation spreads and operations control the disposition and juxtaposition of media elements and the application of operations derived from the intended discourse structure. The navigable discourse structure is the consequence of relationships generated within the presentation component, and can be related to the originally intended discourse structure.

The stages of the initial cycle involve the construction of an initial version of each component. During subsequent cycles the discourse structure and NDS are reconstructed and the media elements and presentation spreads are developed by transformation, the production activities shown as horizontal broken lines in Figure 5. The required activities at each stage x (when x>4) are dictated by comparison of the outcomes of stages x-1 and x-4, the activity shown as a diagonal line. Termination occurs when comparison of the intended discourse structure and

the NDS shows no necessity for further change, as shown in the last two columns on the right of the figure.

In theory a single pass through Stages 1, 2 and 3 of the initial cycle of four stages might produce a satisfactory document. The potential complexity of the document, both in terms of its structure and its variety of producers, suggests that this is an unrealistic prospect. Such a truncated cycle would take no account of the essential requirement for a dynamic approach to the design of content and it would not accommodate any form of evaluation of the final product. The process model must, therefore, accommodate repeated cycles. The iteration is controlled by two simple rules:

***a)*** In any stage x of the subsequent cycles, a comparison between the results of the preceding stage, x-1, and the results of the last stage of the same type as x, i.e. x-4, will guide production activities during stage x.

***b)*** Cycles continue until comparison between the constructed NDS and the required discourse structure indicates that no further changes are needed.

Each of the four inter-stage comparisons has its own purpose. The comparison between discourse structure and abstract media elements will identify those components of the structure that are not 'represented' in any element and therefore need to be instantiated. Likewise the comparison between elements and presentation spreads will show the elements that are not 'presented' in existing or additional spreads. In the case of spreads and the NDS, the comparison will again identify spreads not represented and thus which parts of the NDS need to be reconstructed. The final comparison between NDS and intended discourse structure is the means of determining termination. This comparison is between hierarchical structures, as discussed earlier.

Incorporation of these regular comparisons in the process facilitates orderly production. At the same time it allows the two activities of element and spread creation to 'run ahead' of the results of preceding stages in a controlled manner. An example would be the creation by a designer of additional media elements in expectation, or hope, of a change of discourse structure three stages later.

The products of comparison motivate the activities that change component states. These products include the unrepresented or now redundant discourse components, media elements not presented or spreads no longer required, spreads not included or to be deleted from the current NDS, and the component disparities between intended discourse structure and NDS.

The notion of 'media transformations' provides the necessary conceptualisation for the variety of activities involved in changing collections of media elements from one state to the next. Media transformations provide the means of describing the production of all types of media elements as their properties and relationships vary dynamically. As discussed elsewhere (Morris, 1998), transformations involve the generation or regeneration of one or more abstract media elements of the same or different types and fall into two categories, constructive and supportive. The former includes all processes that produce elements of the final artefact directly (for example the origination or initial creation of a single media element, or the amalgamation or combination of two or more elements of the same or different media types, or the creation of a substitute element in one medium to stand in place of an element in another medium by representing some essential characteristic of the original), the latter those processes involving the use of elements in some type of supporting or subsidiary role (for example the use of an element as the basis for the comparison between it and another element with a view to some constructive transformation of the latter). New individual elements may be the result of some combination of transformations that can be recorded by applying the defined rules (*ibid*). As part of a detailed production method transformations will be useful both organisationally and prescriptively for planning production, and descriptively for design rationale.

The document development model proposed here consists essentially of repeated cycles, each containing four discrete stages. In a subsequent stage of the same type, the preceding stage of that type is the subject of a reconstruction or media transformation. Each stage employs its immediate predecessor both to provide source material and to make a comparison with the last

stage of the same type, that in the previous cycle. These comparisons sustain the consistent development of discourse structure, a fundamental objective. From this follows the choice of the term 'discourse driven' to characterise this process model. Omission of both the discourse-related stages from the model, reducing it to Stages [2] and [3], would render it similar to an 'ad hoc' approach. This new model accommodates different approaches in a general description; it gives an ordering to production activities; it incorporates a means of representing and recording activities and it ensures the establishment of a discourse structure.

The process model provides the backbone for a staged method of production. The stages of the initial cycle and subsequent cycles, and their relationships, correspond directly to the model. The production method provides specific directions, organised around a work plan, which respect the constraints imposed by the process model. The work plan provides guidance in three parts, representing consecutive sets of tasks:-

- establishment of context, purpose and work strategy in an initial specification,
- definition and creation of the components in four initial stages of production, and,
- definition and creation of components in subsequent cycles that repeat the initial stages.

The names of the method stages are:

**0** Preliminary investigations,
**1** Initial discourse structure,
**2** Initial abstract media elements,
**3** Presentation spreads,
**4** Navigable discourse structure,
**[5]** Reconstruction of discourse structure,
**[6]** Transformations of media elements,
**[7]** Transformations of spreads,
**[8]** Reconstruction of NDS,
**N** Final comparison of NDS and discourse structure.

The full details of the method, illustrated with examples from the second case study, production of the SAA Demonstration are available elsewhere (Morris, 1996).

## 6. Related work

The process model presented here exhibits a structure similar to that of the standard 'spiral model' for software development (Boehm, 1988). That model is termed 'risk-driven' because an analysis of risk is an integral part of each cycle, prior to the planning and implementation of each new phase of work. In a similar way, the comparison between NDS and discourse structure is the preliminary to any further development in each cycle of this model. However the specificity of the comparison makes the model presented here less general. There are few other direct comparisons to be made with existing software engineering techniques. Other authors (Hix & Hartson, 1993) have provided an alternative extension of the standard software system development cycle. In their case the emphasis is on the construction, integration and testing of the user interface. Noninterface and interface development activities are connected in a way which emphasises the distinction between interaction design and interface software design. Noninterface development involves both the design and implementation of application software and problem domain design. Problem domain design provides a formal model of the application, produced by software engineers with inputs from system analysis. The work reported here essentially complements and extends this approach by redefining this part of the overall process. This extension facilitates the integration of content design, executed by subject specialists, with the work of engineers responsible for the formalisation of system structure. This extension is essential for the purposes of multimedia document production because of the fundamental role of content designers.

There has also been a considerable amount of work done whose basic theme is the basic structure or architecture of a multimedia system or artefact. The starting point of such work is the manipulation of the many digital formats which make up multimedia from the non-abstract point

of view. Architectures for multimedia systems have taken a variety of forms, often layered. Microcosm (Hall et al., 1996) is a representative example. It is an open architecture for hypermedia systems handling large multimedia information collections. One of the most important features of the architecture is that link information is held in link databases, separately from the documents being linked. The architecture is three layered: an application layer in which general and special purpose media viewers provide the user interface to the multimedia collection, a link service layer in which any number of communicating processes provide the hypermedia functionality, and the storage layer where the documents, link information and other databases are maintained. Systems such as this are clearly, in discoursive terms, closely associated with purely descriptive and expository activities.

Ideas of 'structured documents' (André *et al.*, 1989) and of mark-up languages such as SGML (Goldfarb, 1990) have contributed most to developing notions of the structure of multimedia artefacts, documents in the general sense used here. Such concepts underlie 'HyTime', which is a 'hypermedia/time-based document structuring language', based on SGML and on work originating in a standard music description language (Newcomb et al., 1991; ISO 10743, 1991). It is now an additional standard (IS 10744, 1992; DeRose & Durand, 1994). It aims to represent the structures of all the following: multimedia documents, hypertext documents, time-based documents which specify one or more time schedules to which its rendering is meant to conform, and space-based documents which specify a two- or three-dimensional finite co-ordinates and the relative positions at which the objects in the document are to be rendered.

An alternative approach aimed at the specific problem of interchange concentrates on the final form of a multimedia composition rather than on a generalised multimedia document type. The work of the Multimedia and Hypermedia Information Encoding Expert Group (MHEG) exemplifies the final-form approach (Buford & Brennan, 1994). Its format retains spatial and temporal relationships, but omits any logical model of content. The principal objective of the MHEG standard (Effelsburg & Boudnik, 1995) is to address the coded representation of final-form multimedia and hypermedia objects that will be interchanged across services and applications, by means, for example, wide area telecommunication or broadcast networks. MHEG defines an interchange format, including specifications regarding how to process the data being interchanged, within the context of a client server architecture.

All these products of this artefact driven approach will act as essential complements to design techniques, together giving multimedia documents their final presentational form. However no type of 'structured document' recognises the discoursive structures already discussed. For this reason they cannot contribute to any integration between design and development techniques. However the design of systems, and particularly 'authoring systems' does however often involve some implicit or explicit assumptions about process.

The dominant themes of investigation have been design at a conceptual level, sometimes termed 'authoring-in-the-large' (Schiff, 1992), and structured domains of content, often using hierarchically organised databases. The former approach has its origins in software engineering and echo its terminology (DeKremer & Kron, 1976). The alternative assumption that an artefact will be built around an hierarchy in which all nodes of the tree have a similar presentation structure and all are potentially visible, forms the basis for other work. The implication of this assumption is that a user will always pass through the document by making a choice wherever the tree branches. It is possible to capitalise on such inherent modularity by a top-down approach to 'authoring' which starts at the root level, decomposes to the next level and then moves progressively downwards. This is the basis for the approach of the CMIF (Centrum voor Wiskunde en Informatica Multimedia Interchange Format) authoring approach (Hardman *et al.*, 1993; 1994; Rutledge *et al.*, 1997).

Work taking the hierarchical or structured domain approaches usually assumes that a potential user will wish to search in, or browse through, a database modularised to maximise information content and accessibility. A much more directed, and discourse orientated approach requires the document to be unified by some specific communicative structure or purpose, other

than the provision of a comprehensive description. This is partly the basis for one approach (Thüring *et al*., 1991) which takes as its starting point the idea from discourse analysis of local and global coherence and combines it with a tripartite structure from classical rhetoric. The three parts are 'content', 'structure' and 'organisation', corresponding to 'invention', 'disposition' and 'presentation' (apparently a kind of amalgamation of 'style' and 'delivery'). This approach is the foundation for SEPIA, a 'co-operative hypermedia authoring environment' (Streitz *et al*., 1992). SEPIA offers a flexible environment for authoring with support for some discourse structures, but without any recognition of the different roles of participants or how the products of their work will be integrated in a purposive manner. Its schemata is a close relative of the general graphical representations of knowledge structures which already provide the basis of systems such as gIBIS (Conklin & Begeman, 1988).

Detailed studies of automated argument or narrative construction via AI or other techniques are outside the scope of our current work. However it should be noted that detailed multimedia design, in the sense defined in Section 1, is now the subject of serious study, both in terms of guidelines and heuristics for design and of user performance. An example is the study of 'attentional design' and an associated prototype tool embedding guidelines in an expert system (Faraday & Sutcliffe, 1997).

## 7. Conclusions

Multimedia production introduces an important new dimension to the practical responsibilities of software engineering. The integration of the design of content with the development of software access mechanisms is the key to successful production. Introduction of this new dimension of engineering activity requires a radical widening of perspective in order to provide a framework suitable for new participants and activities. These new activities demand reconsideration of basic factors such as the nature of media and the relationship of multimedia production to traditional publishing and other activities.

The research reported here makes a basic differentiation between abstract, physical, and virtual physical or digital media, thus bringing within the ambit of software engineering the sign systems used in all forms of human communication. On this basis it is possible to exploit a wide range of work related to content structure, ranging from formal linguistics and traditional discourse structures to recent work on digital genres. The discourse structures available from work in these other fields of research provide the essential basic framework of content for a multimedia production.

In order to integrate content generation into software engineering practice, it is necessary to have a general notion of discourse that can form part of an overall production process. The concept of navigable discourse structure provides this notion. Our process model incorporates, crucially, the ideas of a required discourse structure which at the end of each cycle is compared with the actual NDS created by media elements and access mechanisms. This process model provides the framework for a detailed production method which has been developed and tested via two case studies involving the production of multimedia demonstrations of software engineering tools.

This approach, although not directed to multimedia systems in the usual sense of the means of manipulating representation media or formats, should in the future aid both practical development with existing tools and the design of new tools and multimedia systems. Moreover it offers a general theoretical framework for considering all aspects of multimedia production. Future work will proceed along three avenues, method refinement using further case studies, application of techniques to related fields, in particular to requirements engineering, and development of the concept of abstract media as it might apply to software engineering modelling languages.

## References

(André *et al.*, 1989) André, J., Furuta, R., and Quint, V. (Eds.) *Structured Documents*. Cambridge: Cambridge University Press, 1989.

(Andersen, 1990) Andersen, P. Bøgh *A theory of computer semiotics: Semiotic approaches to construction and assessment of computer systems*. Cambridge : Cambridge University Press, 1990.

(Bertin, 1983) Bertin, J. *Semiology of graphics*; translated by W.J.Berg. Madison, Wis: University of Wisconsin Press, 1983.

(Blair & Stefani, 1997) Blair, G. and Stefani, J-B. *Open distributed processing and multimedia*. Harlow : Addison-Wesley, 1997

(Blakemore, 1988) Blakemore, D. The organisation of discourse. In Newmeyer, F. (ed.) *Linguistics : The Cambridge survey, IV Language: the socio-cultural context*. Cambridge: Cambridge University Press, 1988

(Blattner & Dannenberg, 1992) Blattner, M. M. and Dannenberg, R. B. (Eds.) *Multimedia interface design*. New York : ACM Press, 1992.

(Boehm, 1988) Boehm, B. A spiral model of software development and enhancement. *IEEE Computer*, Vol 21 No 5, pp 61-72, May 1988

(Borenstein & Freed, 1993) Borenstein, N. and Freed, N. *MIME (Multipurpose Internet Mail Extensions) Part One: Mechanisms for specifying and describing the format of Internet message bodies*, RFC-1521 [On-line], September 1993. Available: FTP nic.ja.net/ pub/ newsfiles/ JIPS/ rfc/ rfc.1500-1599/ rfc1521.

(Buford & Brennan, 1994) Buford, J.F.K. and Brennan, R. Multimedia interchange. In J.F.K. Buford (Ed.) *Multimedia systems*. New York : ACM Press, 1994.

(Chatman, 1990) Chatman, S.B. *Coming to terms : the rhetoric of narrative in fiction and film*. Ithaca, NY: Cornell University Press, 1990.

(Checkland, 1981) Checkland, P. *Systems thinking, systems practice*. Chichester: Wiley, 1981.

(Checkland, P. & Scholes, 1990) Checkland, P. and Scholes, J. *Soft systems methodology in action*. Chichester: Wiley, 1990.

(Conklin & Begeman, 1988) Conklin, C. and Begeman, M.L. gIBIS : A hypertext tool for exploratory policy discussion. *ACM Transactions on Office Information Systems*, Vol 6 No 4, pp 303-331, October 1988.

(Crawston & Williams, 1997) Crawston, K. and Williams, M. Reproduced and emergent genres of communication on the WWW. In *Proceedings of the Thirtieth Hawaii International Conference on System Sciences* Vol 6 , Los Alamitos, Ca.:IEEE Press, 1997

(Crocker, 1982) Crocker, D.H. *Standard for the format of ARPA Internet messages*, RFC-822 [On-line], August 1982. Available: FTP nic.ja.net/pub/newsfiles/JIPS/rfc/rfc.500-999/rfc822.

(DeRemer & Kron, 1976) DeRemer, F. and Kron, H.H. Programming in the large versus programming in the small. *IEEE Transactions on Software Engineering* ,Vol SE-2 No 2, 1976

(DeRose & Durand, 1994) DeRose, S. and Durand, D. *Making hypermedia work: A user's guide to HyTime*. Boston Kluwer Press, 1994

(van Dijk, 1977) van Dijk, T.A. *Text and context*. London: Longmans, 1977

(van Dijk, 1997a) van Dijk, T.A. The study of discourse. In van Dijk, T.A. (ed.) *Discourse as structure and process, Discourse studies: A multidisciplinary introduction Volume 1*. London: Sage, 1997

(van Dijk, 1997b) van Dijk, T.A. (ed.) *Discourse as structure and process, Discourse studies: A multidisciplinary introduction Volume 1*. London: Sage, 1997

(van Dijk, 1997c) van Dijk, T.A. (ed.) *Discourse as social interaction, Discourse studies: A multidisciplinary introduction Volume 2*. London: Sage, 1997

(Effelsburg & Boudnik, 1995) Effelsburg, W. and Boudnik, T.M. MHEG explained. *IEEE Multimedia*, Vol 2, No 1, Spring 1995

(Eggins & Martin, 1997) Eggins, S. and Martin, J.R. Genres and registers of discourse. In van Dijk, T.A. (ed.) *Discourse as structure and process, Discourse studies: A multidisciplinary introduction Volume 1*. London: Sage, 1997

(Faraday & Sutcliffe, 1997) Faraday, P. and Sutcliffe, A. Multimedia: Design for the 'Moment' In *Proceedings of ACM Multimedia 97*, ACM Press, 1997

(Finkelstein *et al.*, 1992) Finkelstein, A., Kramer, J., Nuseibeh, B., Finkelstein, L., and Goedicke, M. Viewpoints: A framework for integrating multiple perspectives in system development. *International Journal of Software Engineering and Knowledge Engineering*, Vol 2 No 1, pp 31-57, March 1992.

(Goldfarb, 1990) Goldfarb, C.F. *The SGML Handbook*, Oxford: Clarendon Press, 1990.

(Hall et al., 1996) Hall, W., Davis, H.C. and Hutchings, G.A. *Rethinking hypermedia: the Microcosm approach*. Amsterdam, Kluwer Academic Press, 1996.

(Hardman *et al*., 1993) Hardman, L., van Rossum, G. and Bulterman, D.C.A. Structured multimedia authoring In *Proceedings of ACM Multimedia '93, First ACM International Conference on Multimedia,* Anaheim, CA., August 1993. New York: ACM Press, 1993.

(Hardman *et al*., 1994) Hardman, L., van Rossum, G., Jansen, J. and Mullender, S. CMIFed: A transportable hypermedia authoring system. In *Proceedings of ACM Multimedia '94, Second ACM International Conference on Multimedia,* San Francisco, October 1994. New York: ACM Press, 1994.

(Hix & Hartson, 1993) Hix, D. and Hartson, H.R. *Developing User Interfaces: Ensuring Usability Through Product and Process*. New York: Wiley, 1993

(Horn, 1988) Horn L. Pragmatic theory. In Newmeyer, F. (ed.) *Linguistics : The Cambridge survey, I Linguistic Theory: Foundations*. Cambridge: Cambridge University Press, 1988

(Hovy & McCoy, 1989) Hovy, E.H. and McCoy, K.F. Focusing your RST: A step toward generating coherent multisentential text. In *Program of the 11th American conference of the Cognitive Science Society*, Ann Arbor, Michigan pp 667-674. Hillsdale, NJ : Erlbaum, August 1989.

(ISO 10743, 1991) International Organisation for Standardisation *Information technology - Standard music description language (SMDL)* ISO/IEC CD 10743, April 1991.

(ISO 10744, 1992) International Organisation for Standardisation *Information technology - Hypermedia/time-based structuring language (HyTime)* ISO/IEC 10744:1992(E), 1992.

(Johnson-Laird, 1983) Johnson-Laird, P.N. *Mental models*. Cambridge: Cambridge University Press, 1983

(Kramer *et al*., 1993) Kramer, J., Magee, J., Ng, K. and Sloman, S. The System Architect's Assistant for design and construction of distributed systems. In *Proceedings of Fourth IEEE Workshop on Future Trends of Distributed Computing Systems*, Lisbon, September 1993. Los Alamitos, Ca: IEEE Computer Society Press, 1993.

(Kress *et al.*, 1997) Kress, G., Leite-Garcia, R. and van Leeuwen, T. Discourse semiotics. In van Dijk, T.A. (ed.) *Discourse as structure and process, Discourse studies: A multidisciplinary introduction Volume 1*. London: Sage, 1997

(Kronlöf *et al*., 1993) Kronlöf, K., Sheeham, A. and Hallman, M. The concept of method integration. In K. Kronlöf (Ed.) *Method integration: Concepts and case studies*. Chichester: Wiley, 1993.

(Labov, 1972) Labov, W. The transformation of experience in narrative syntax. In W. Labov (Ed.) *Sociolinguistic patterns*. Philadelphia : University of Philadephia Press, 1972.

(Landow, 1992) Landow, George P. *Hypertext : the convergence of contemporary critical theory and technology*. Baltimore : Johns Hopkins University Press, 1992.

(Macromedia, 1994) Macromedia Inc. *Macromedia Director Version 4 - Using Director*. San Francisco: Macromedia, 1994.

(MacroMind, 1991) MacroMind Inc. *MacroMind Director Version 3.0 - Overview Manual*. San Francisco: MacroMind, 1991.

(Mann & Thompson, 1988) Mann, W.C. and Thompson, S.A. Rhetorical structure theory: Toward a functional theory of text organisation. *Text* ,Vol 8 No 3, pp 243-281, 1988.

(Metz, 1974) Metz, C. *Film language : A semiotics of the cinema*; translated by M. Taylor. New York: Oxford University Press, 1974.

(Morris, 1938) Morris, C.W. *Foundations of the theory of signs*. Chicago: University of Chicago Press, 1938

(Morris, 1996) Morris, S.J. *A method for the design and development of multimedia documents, Chapter 8 -Production method*. Ph.D. Thesis , Dept. of Computing, Imperial College of Science, Technology & Medicine, London University, January 1996. (Available: WWW URL http://www.soi.city.ac.uk/~sjm/Summary.html).

(Morris, 1998) Morris, S.J. Media transformations for the representation and communication of multimedia production activities. In Sutcliffe, A. et al. (Eds) *Designing Effective and Usable Multimedia Systems.* Norwell, Mass.: Kluwer Academic Publishers, 1998

(Morris & Finkelstein, 1993) Morris, S.J. and Finkelstein, A.C.W. Development of Multiple Media Documents. In *COOCS 93, Proceedings of ACM Conference on Organizational Computing Systems*, Milpitas Ca., November 1993, pp 69-77. New York: ACM Press, 1993

(Morris & Finkelstein, 1996) Morris, S.J. and Finkelstein, A.C.W. Integrating design and development in the production of multimedia documents. In *Proceedings of MMSD 96, International Workshop on Multimedia Software Development*, Berlin, Los Alamitos, Ca.: IEEE Press, 1996

(Newcomb *et al*., 1991) Newcomb, S.R., Kipp, N.A. and Newcomb, V.T. "Hytime" The hypermedia/time-based document structure language. *Communications of the ACM*, Vol 34 No 11, pp 67-83, November 1991.

(Ng, 1992) Ng, K. *Visual support for distributed programming*. PhD Thesis, Department of Computing, Imperial College, August 1992.

(Nunberg, 1997) Nunberg, G. Genres in digital documents: Introduction. In *Proceedings of the Thirtieth Hawaii International Conference on System Sciences Vol 6* , Los Alamitos, Ca.: IEEE Press, 1997

(Nuseibeh & Finkelstein, 1992) Nuseibeh, B.A. and Finkelstein, A.C.W. Viewpoints: A vehicle for method and tool integration. In *Proceedings of CASE '92*, Montreal, Canada, July 1992.

(Perelman & Olbrechts - Tyteca, 1968) Perelman, C., and Olbrechts - Tyteca, L. *The new rhetoric: A treatise on argumentation*; translated by J. Wilkinson and P. Weaver. Notre Dame: University of Notre Dame Press, 1968.

(Prince, 1988) Prince, E.F. Discourse analysis: a part of the study of linguistic competence. In Newmeyer, F. (ed.) Linguistics : *The Cambridge survey, II Linguistic Theory: extensions and implications*. Cambridge: Cambridge University Press, 1988

(Rutledge *et al.*, 1997) Rutledge, L., van Ossenbruggen, J., Hardman, L. and Bulterman, D.C.A. A Framework for generating adaptable hypermedia documents. In *Proceedings of ACM Multimedia 97*, ACM Press, 1997

(Saussure, 1983) Saussure, F. de *Course in general linguistics*; translated by R. Harris. London : Duckworth, 1983.

(Schiff, 1992) Schiff, J. (Ed.) *Esprit Project 5252, HYTEA - Hypertext Authoring*, Public report, Munich: Siemens-Nixdorf Informationssyteme AG (NSI), November 1992.

(Stalnacker, 1972) Stalnacker, R. Pragmatics. In Davidson, D. & Harman, G. (eds.) *Semantics of natural language*. Cambridge: Cambridge University Press, 1972

(Streitz *et al*., 1992) Streitz, N., Haake, J., Hanneman, J., Lemke, A., Schider, W., Schütt, H. and Thüring, M. SEPIA: A cooperative hypermedia authoring environment. In *ECHT '92, Proceedings of Fourth ACM Conference on Hypertext*, Milan, December 1992. New York: ACM Press, 1992.

(Thüring *et al.*, 1991) Thüring, M., Haake, J.M. and Hanneman, J. What's Eliza doing in the Chinese room? Incoherent hyperdocuments - and how to avoid them. In *Hypertext '91, Proceedings of Third ACM Conference on Hypertext*, San Antonio, Texas, December, 1991. New York: ACM Press, 1991.

(Toulmin, 1958) Toulmin, S. *The uses of argument*. Cambridge: Cambridge University Press, 1958.

(Winograd & Flores, 1986) Winograd, T. and Flores, C. F. *Understanding computers and cognition : A new foundation for design*. Norwood, NJ: Ablex Publishing Corp., 1986.

(Wray *et al*., 1994) Wray, S., Glauert, T. and Hopper, A. The Medusa applications environment. *IEEE Multimedia*. Vol 1 No4, pp 54-63, Winter 1994.

(Yates & Sumner, 1997) Yates, S.J. and Sumner, T.R. Digital genres and the burden of fixity. In *Proceedings of the Thirtieth Hawaii International Conference on System Sciences Vol 6* , Los Alamitos, Ca.: IEEE Press, 1997