1 Four Axes of Rhetorical Convergence

1.1.1 Background

Rhetorical convergence is the coexistence in one text of means of expression that we are used to regard as belonging to different media. It is a perception of rhetorical heritage in the encounter of new media forms. The term *rhetorical convergence* is new, but the perspective is not. Studies of the appearance of new media are often conducted by comparing the forms in a new medium to those in earlier media. Raymond Williams compared television to radio, newspapers, theatre, past-time games and vaudeville in *Television: Technology and Cultural Form.* Cinema historians such as Charles Musser or David Robinson note the various influences that led to the development of narrative cinema. Indeed, Jay David Bolter and Richard Grusin argued in *Remediation* that comparison to earlier media is the only way we are able to perceive a new medium as a medium.

The new medium in question in this thesis is the World Wide Web. It is commonly described as a medium that will incorporate (even surpass!) all earlier media, in a process that is described as convergence. The U.S. media conglomerate AOL Time Warner may serve as an illustration of this.

In January 2001, the media conglomerate Time Warner merged with the Internet company America Online. This was not only larger than any previous merger in the United States; it was also a seen as a logical outcome of *convergence.* The merger was a

A shorter version of this essay is previously published in *Dichtung Digital* 5.30 (2003): <http://www.dichtung-digital.org/2003/issue/4/fagerjord/index.htm>.

consequence of capitalistic logic, as technologies of production, distribution and consumption of what we have known as different media had come together. After the merger, a person in New York City could call up his AOL Internet account through Time Warner Telecom's telephone service or the Time Warner Cable network, and read news stories on the AOL home page or in *Time Magazine,* watch *CNN* or Warner Brothers Television's *Friends*, view a trailer for the Warner Brothers' movie *Harry Potter* or *The Lord of the Rings* from New Line Cinema. Both movies would soon be on HBO or Time Warner Cable. Later, he could or listen to Madonna and Faith Hill's latest albums on the Warner Bros. label being played on AOL Web Radio. This would all be available not only with the same computer, but with the same computer program, Netscape Navigator, also made by AOL Time Warner.•

This novel mix of media within one computer screen is the focus of this thesis. AOL Time Warner is a particularly telling example, but neither the first nor the only one. Although by far the largest such company, AOL faced competition from several other alliances between the software industry, television and telecommunications networks, publishing, and broadcasting.

As these different media companies were all distributing different kinds of media through different kinds of channels, cross-media experiments also flourished: combinations of print stories from different sources with parts of radio and television programs, often with associated online shops. Such experiments show an apparent belief in, or at least willingness to try out, the possibilities of successfully combining video and radio recordings with news copy, photographs and "interactive" features. Multimedia Web sites drawing on many sources seemed to confirm the view that in the future we will have one medium that is the sum of all previous media. The logical outcome of convergence, it seemed to many, is that all earlier media will converge into one.

• Information about the various AOL Time Warner companies taken from "Companies." *AOL Time Warner.* (2001, AOL Time Warner, 24 March, 2003 .<http://www.aoltimewarner.com/companies/index.adp>).

1.1.2 Aim

"Most current media formats will die and be replaced with an integrated Web medium in five to ten years" wrote the influential Web design consultant and writer Jakob Nielsen in 1998 (par. 1). The same year, MIT Media Lab researcher Andrew Lippman was less conservative in his prediction: "Forget television. In three years, there won't be any." The Media Lab's founder, Nicholas Negroponte asserted three years earlier that "the future of the [...] TV is the PC. Period" (47).

Scholars of new media also noted this current of thought: "The ultimate ambition of the Web designer seems to be to integrate and absorb all other media." Jay David Bolter and Richard Grusin write (208). Janet Murray holds that "[o]ne of the clearest trends determining the immediate future of the digital narrative is the marriage between the television set and the computer" (253).

This thesis grew from a curiosity towards these predictions. More precisely, a curiosity towards what such a convergent Web medium would look like. How would it use images, text, sound, movement?

It seemed to me that these predictions contained an inherent contradiction. The technological possibility of mixing the media together could not alter the fact that different media have very different and long-standing traditions of writing, photography, editing. Different genres and their corresponding writing styles have been thoroughly studied in media studies, as have the "languages" of cinema, of photography, and lately, of new media. Such studies point to similarities, but also to considerable differences between media.

Following from these observations, the main aim of the present work is *to investigate whether the "languages" of earlier media are converging on the Web, and if so, how?* To answer this question, we need to know these "languages" of earlier media. As the term *language* is too closely tied with natural, spoken language and its corresponding field of linguistics, I will in this thesis use *rhetorics* instead of *"languages."*

Four Axes of Rhetorical Convergence

Rhetoric should be understood broadly as "means of expression." In Antiquity, these were restricted to the use of words, and yet the classical tradition has delivered a huge catalogue of rhetorical devices to us. In modern media, the means of expression are expanded to include images, sounds, moving images, and texts that respond to user input. Being a network of multimedia computers, the Web has an unprecedented range of rhetorical means at its disposal: not only the means known from earlier media, but also new combinations of rhetorics from different media.

Media rhetorics are determined by many factors. Each medium uses a technology that can carry certain kinds of signs for communication. Television technology, for example, can broadcast moving images and sound of a certain quality in a limited number of channels via air or cable. Each technology further has a place in the society it is used in, and a cost of use. In the case of television, domestic receivers have become the norm. Costs involve equipment and personnel for recording and broadcasting events, as well as the staging of the events themselves. The economy of the media production and consumption is both determined by and determines the medium's place in society. Television, for example, has relied either on a licence fee or inserted commercial messages, with a schedule that is repeated on a daily or weekly basis and its "prime time" in the evening.

Within the limits and possibilities of each medium's technology, economy and position in society, a number of conventional forms have established themselves. *Convention* is a key term in this respect. This thesis uses a semiotic framework where human communication is seen to operate with *signs,* which are tied to meanings by a shared convention or *code.* Codes or convention also determine how signs are combined, as in the rules of syntax and grammar that make up natural languages, or as in practises of editing evolved over a century of film-making. Furthermore, many combinations of words, images or ideas are conventionally understood as carrying an additional meaning, what is known as *connotations* or *overcoding.* Finally, it is hard to say

anything comprehensible – by speech, writing film, or otherwise – without following the conventions of a certain *text type, genre* or *register*, a pattern of meanings and elements that fits the situation, and helps the listener, reader, or viewer predict much of what will be said, order the information according to the pre-established pattern, and know how to react (whether to answer, to laugh, to cry, to be angry, to think, to feel informed…). Uses of sign systems, editing and syntax, connotations, overcodings, genres and text-types are all aspects of what we as a whole will call *rhetoric*, a notion of rhetoric we will return to shortly. All these aspects of different media's forms have been the attention of scholars, some of them for millennia, so there is a rich literature to draw on. Much less studied are the forms of computer media, and how these are similar or different to the forms of older media. This thesis is a contribution in that direction.

Earlier media are understood as relative to the World Wide Web, which debuted in 1991. Print media, radio, and television are obvious predecessors, but it should be remembered that in 1991, computers had already been used for text processing for decades.• This means that there exists a tradition of computer rhetoric – in user interfaces, games, hypertext, MUDs, CD-ROMs, Internet protocols preceding the Web (such as Gopher), and experimental interactive media that runs before the development of the Web site as a genre or medium. Although this has received much less attention from scholars than other media, there is a considerable literature on interaction design and literary forms on computers.

If we can form knowledge of what the rhetorics of earlier media are, what, then, would a *convergence* of these mean? To *converge* is "to tend to one point from different places," according to the Oxford English Dictionary. The term is used to describe many developments. To converge may simply mean to join, to meet in a point. In mathematics, however, a convergent function may approach a limit, but never reach it. When linguists say that two languages converge, one of the

• The actual year would depend on what "text processing" is taken to mean. The British calculator "Colossus" was designed in 1943 to break German codes during World War II, thus turning cipher into text ("Computers"). Experiments with word processing and hypertext started in the mid-1960s (Engelbart).

two will dominate the other, as the weaker language will take over forms from the dominant one; and in biology, two convergent species may look the same, or otherwise have the same characteristics without being related (as fish and whales). Two cities may even be said to converge without moving at all; geographers talk of a space-time convergence, if better communications reduce the travel time between to places. The fact that I can both read a book and watch a movie on my computer may very well be called convergence. What used to be on paper and silver screen has converged on my monitor's liquid crystals. But this is not a convergence of *rhetoric*. The book is written in the same way as before, and the film's editing does not change. A convergence of means of expression, a *rhetorical convergence* would require that rhetorics from print tended to join rhetorics from cinema. If this is happening – and we will soon see that there is good reason to argue that it is – then *how* it is happening? Do approaching rhetorics meet, or do they just approach each other? Are there certain forms that may be combined, while others are like oil and water?

This thesis is an attempt to answer these questions, using a certain theoretical perspective and research strategy, which now will be outlined.

1.1.3 Rhetoric

What is studied in this thesis is the use of rhetorical forms or devices in media texts. These are the communicative means used, be it written, spoken, pictorial or a combination of these. The means found in a text are the results of the choices of the author(s), choices that could have been made differently. Such choices are made within a range of possibilities that is limited not only by the capabilities and imagination of the author, but also by the communication technology used, the situations of the author (sender) and the reader (receiver), the relations between author and reader, and the history or tradition of earlier similar texts the author and reader know. Understanding *rhetoric* in this way might contrast with the conception many have of *rhetoric.* Arguably, we might have chosen another word

to avoid this initial confusion, but the candidates are not very attractive. Concepts such as "communicative strategies," "discursive practises," "formal elements," or "language" are equally ambiguous, and could all be called synonyms to *rhetoric*.

Two major objections to this use of the term *rhetoric* are easy to conceive. The first is that *rhetoric* is the art of persuasive oratory, and thus not applicable to computer multimedia at all. The second is that there hardly is any use of the massive terminology of rhetoric in this thesis, and that instead a semiotic terminology is used. One may indeed ask if I need the concept of *rhetoric* at all, if not *semiotic* would be better.

I will refute the first objection with an argument by authority. While there are many scholars who wish to reserve the label *rhetoric* for spoken argumentative discourse, historian of rhetoric George A. Kennedy explains that such a position is not supported by tradition.

> All communication involves rhetoric. A speaker or writer has some kind of purpose, and rhetoric includes the ways of accomplishing, or attempting to accomplish, that purpose within a given culture. The Greeks and Romans thought of this purpose as persuasion, but by that they meant something more general than persuasion as understood by a modern social scientist. Purposes cover a spectrum from converting hearers to a view opposed to that previously held, to implanting a conviction or belief not otherwise entertained, to teaching or exposition, to entertainment and demonstration of the cleverness of the speaker. (1–2)

To clarify, Kennedy discerns between *primary rhetoric* as the act of speaking, while *secondary rhetoric* "refers to rhetorical techniques as found in discourse, literature, and art forms when those techniques are not being used for an oral, persuasive purpose" (3). Both primary and secondary rhetoric has been present at all times, Kennedy claims.

Four Axes of Rhetorical Convergence

> It has been a persistent characteristic of classical rhetoric in almost every stage of its history to move from primary to secondary forms, occasionally then reversing the pattern. [...] There is also a secondary rhetoric in arts other than literature. In antiquity, the analogy between rhetoric and painting or sculpture was repeatedly noticed – by Aristotle, Cicero, and Quintillian among others – and an analogy to architecture is occasionally mentioned as well. Writers on the arts sometimes borrowed terminology from rhetoric. In the Renaissance and later, treatises on music, painting, and other arts borrowed the structure and categories of classical rhetoric. (3–4)

The twentieth century saw a renewed interest in rhetoric, both the classical tradition, and the "new rhetoric" in theories by writers such as Ivor Armstrong Richards, Kenneth Burke, Stephen Toulmin or Chaim Perelman and Lucie Olbrechts-Tycheca (see Foss, Foss and Trapp for an overview). These writers were primarily concerned with questions of communication processes in general, persuasion through identification, or logic and argument from probability. Another twentieth-century incarnation of rhetoric was initiated by Roland Barthes in his essays on rhetorical analysis itself, and as practised on images and fashion. Barthes' interest in rhetoric lies not with formal (enthymemical) arguments, but with metaphors, metonymies and other tropes. Just as in *Mythologies,* Barthes focuses in his essays on rhetoric on connotational meanings. In "The Rhetoric of the Image," his famous analysis of an ad for Panzani pasta, he shows how the image denotes merely dinner ingredients, but carries connotations to Italy, homemade cooking, and art. Two decades later, semiotician and narratologist Seymour Chatman, discusses in *Coming to Terms* the use of rhetoric both in film and in literature. He also launches the term *suasion* as an alternative term for *persuasion*. *Suasion* carries the extended understanding of a message's effects on audience members that Kennedy in the quote above explains also the Greek rhetors held.

It is thus in line with both the old and the modern tradition of rhetoric to include written language and

images, and discourses not primarily meant to change the audience's opinions.

The second objection: why "rhetorical convergence," why not "semiotic convergence"? Aristotle defines rhetoric as "the faculty of discovering the possible means of persuasion in reference to any subject whatever" (*Rhet. 1355b*). The rhetorical tradition is an enormous storehouse of such means, of different ways of communicating that have proven themselves to be effective in different circumstances. It is from this perspective we will view converging media in this thesis: that there is possible to isolate and list devices of speaking (and further of writing, of drawing, of making photographs, of editing) that may be applied relatively independent of the subject treated.

The semiotic (or semiological) theories of Peirce, Saussure, and Eco attempt first and foremost to explain how communication and understanding is at all possible. These questions are almost entirely left out of this thesis. Instead we ask what the power or effectiveness of the rhetorical (or semiotic) means chosen is, and what their combined powers are. Questions of what constitutes effective communication are the classical focus of rhetorical study.

A further reason for choosing the term rhetoric is the tradition's focus on the whole communication process. Typically, Aristotle's *Rhetoric* is divided into three books: on the speaker, the audience, and the speech itself respectively. Greek and Roman rhetors understood that a speech on a public square would be different than one in a private dinner. Their topics and aims would be different, as the speech on the square would probably be political (*deliberative*) or part of a lawsuit• (*forensic*), while the dinner speech might be a speech of praise (an *epideictic* speech). The size of the audiences and the acoustics on the scenes would be different, and there might also be considerable differences in education and ideology between different audiences. These aspects should be considered before one opens the mouth to speak. Again, such a communicative focus is not

• A Greek jury would have several hundred jurors, according to Kennedy (20).

outside of semiotics, but the central issue of semiotics has always been the text, the signs.

These are subtle differences within perspective and questions asked, however. There are considerable overlaps between the two traditions (indeed, both traditions claim to be valid for all human communication (Kennedy 1; Eco, *A Theory of Semiotics* 3)). This is to be expected. As Barthes reminds us in "The Old Rhetoric" (92), the rhetorical tradition is an integral part not only of Western language philosophy, but also of the languages themselves. Neither is thinkable in their present forms without rhetoric. In *Iconology* W. J. T. Mitchell asks what is new about semiotics, as most of its concepts can be traced back centuries in time (59–62). Umberto Eco also openly credits this heritage in *Semiotics and the Philosophy of Language* and *Kant and the Platypus.* That semioticians of the twentieth century have been able to reinterpret rhetoric in a semiotic vocabulary is not necessarily an example of a theoretical colonialism, but rather an example of a healthy offspring.

Within semiotics, the word *rhetoric* most often shows up when discussing secondary meanings, as in Barthes' analysis of the Panzani ad. Barthes calls the secondary meanings *connotations*, Eco labels them *overcoding* (*Theory* 278–279). Both call this the rhetorical level of communication.

In choosing *rhetoric*, I thus want to focus on the isolation of textual strategies that are tied to certain media and genres, but relatively independent of subject matter, mainly looking at a rhetorical level of connotations; above the denotational, and above the individual sign. If we find that these are combined on the Web in previously neglected or impossible ways, we will call this rhetorical convergence. Parallels in media studies are easy to find. Historians of cinema (Bordwell and Thompson; Musser; Robinson) often point to how the short films of the first decade or so were various kinds of spectacle, including (1) films of everyday events; (2) rarities and vaudeville numbers, including some short narrative episodes, mainly comical, and, (3) in the case of Meliés' work, impossible, fictive scenes. Longer,

dramatic narratives were a later development, where the scene as the dramatic unit slowly was replaced with the shot, and greater flexibility in the interplay of camera angles, simultaneous scenes, different locales and jumps in time. This development is normally described as the "invention" or "discovery" of the "cinematic language," but it could also be viewed as a gradual fitting of rhetorical elements from various kinds of spectacle, staged drama and the novel.

Another parallel may be found in Raymond Williams' work *Television*, where Williams views many television genres (news, debates, education, drama, films, variety shows, sports, advertising, and pastimes) as derived from earlier media or social events, brought together in the television schedule, and adapted to the new technology.

1.1.4 Semiotic description

In this thesis, we will use a semiotic vocabulary to describe the rhetoric of various media genres and their convergence. As a science of all kinds of human communication with signs, semiotics is a theory that makes us able to describe in one consistent vocabulary all the different rhetorics of different media we will discuss in the following essays. To be more specific, I am using Umberto Eco's semiotics as a theoretical basis, but also relying heavily on the sensitive analyses of Roland Barthes, especially in *S/Z*. I agree with Jacques Derrida (51), and later Eco (*Theory* 172 ff.), that Barthes must have been wrong in "Eléments de sémiologie," where he takes language as a model of all semiotic systems. This does not falsify the many brilliant insights in Barthes' later works, which I believe easily can be aligned with the framework of Eco's semiotic.

Narratology is another important source of theory in this thesis, especially that of Gérard Genette's *Narrative Discourse*, but also Seymour Chatman's discussion of rhetoric within narrative and Edward Branigan's combination of theories of film narrative and cognitive psychology. Several narratological studies of computer media have already been undertaken, suggesting that

expansions of the narrative terminology are necessary (Liestøl, "Wittgenstein"; Aarseth, *Cybertext;* Walker; Eskelinen, "Cybertext Narratology"). I believe these studies demonstrate the fruitfulness of applying narratology to computer media. The merit of semiotics may seem a little more debated, however. In *Cybertext*, Aarseth (24–41) has expressed doubts that a semiotic approach is able to shed any light on computer media, as it posits a "lower threshold" between signs – which stand for something else according to some convention – and *signals,* "units of transmission which can be computed quantitatively irrespective of their possible meaning […]" (Eco, *Theory of Semiotics,* 20). In line with this, Aarseth finds inconsistencies and problems in semiotic writings of Jay David Bolter, Jens F. Jensen, and Peter Bøgh Andersen. Although I share Aarseth's view that in order to understand computer texts we need to understand the cybernetic principles they incorporate, I still find a semiotic perspective useful, even necessary, to describe not only the combinations of different sign systems such as moving images and writing, but also how users come to understand how to read computer texts with links, input fields, rollovers and other input mechanisms that influence on the text while it is read.•

• This is dealt with in more detail in the chapter "Semiotics of the Web" (chapter 5).

1.1.5 Medium, Genre, Architext, Text-type

There is another aspect of both Andersen's *Theory of Computer Semiotics* and Jensen's "Formatering af forskningsfeltet" we need to consider. It is a point not discussed by Aarseth: Andersen's and Jensen's decisions to view the computer as a medium. The term *medium* is here used in the narrow sense, as a technology enabling signs to be communicated. Within the fine arts, this is the normal use of the term, where "oil on canvas" or "marble" are different media.

Media studies, on the other hand, understands the term differently. Across the essays that make up this thesis, I have used a more complex sense of the term *medium*, where technology is only one of several parts. In this view a medium is a complex social institution where a particular use of a certain technology has its

own economy, legal status, and function within society and our private lives. This is the sense of the word most used in media studies, and I believe also what most people mean when they speak of media. If we want to understand media's communicative role, it will not do simply to regard them as technology. Print technology is used for many media: books, newspapers, magazines, leaflets, billboards, et cetera. As Raymond Williams, for example, is quick to point out, media technologies such as television could be imagined taking many other forms than they have. In Williams' words, media are technologies and cultural forms.•

Following this argument, I will argue that the Web might best be described not as a medium, but as a shared technology or technological framework encompassing several media. The online store *Amazon* has a very different economy and function in society than *The New York Times on the Web,* Microsoft's corporate home page, the page created by the Norwegian Tax Authorities for submission of tax reports, or my private homepage with family pictures. I think these pages are so different from each other, not only textually, but also economically and in terms of role in society that they can be treated as different media from a media studies perspective.

This opens a related problem, however. If the Web is seen to enable several media, how are we to separate a Web medium from a Web genre?

The distinction cannot be sorted out without some consideration of the nature of genres. Only a short glance at the large literature on genres in different media makes it very clear that both the concept of genre and each individual genre are difficult to define; most theorists still seem to believe that genres are very important for readers' attitudes towards texts.

A text belongs to a genre when there is a clear relation to a family of other texts. Ludwig Wittgenstein pointed out that family resemblance is a network of shared traits; although no trait is found in all the members of a family, we are still able to see that they look alike (*P.U.* §67). Wittgenstein did not write about genres, but

• The issue of Web media is also discussed in the essay "Rhetorical Convergence" (chapter 3).

his discussion of family resemblance seems to fit media genres and may bring some clarity to the discussion.

Rick Altman has observed that genres within film studies tend to be defined either "as a list of common traits" or by "certain constitutive relationships between undesignated and variable placeholders," or in his terms: *semantically* and *syntactically* (30). He notes that the thematical (syntactic) definition will tend to be more narrow and allow fewer films as really belonging to a genre than the formal (semantic) definition. For Altman, both criteria should be used when discussing a genre. Altman uses the Western movie as an example; a Western will be set in the American West in the 19th century, and contain horses, cowboys, gunslingers, lone sheriffs, Indians, and so on. These are the semantic elements. Syntactically, a Western will have as its thematics the conflict between nature and culture, law and lawlessness, and using the frontier as a liminal space (30–31).

Jerry Palmer reminds us that we should further realise that a text may belong to several genres at the same time, and genres may be recognised within genres. Think of the movie *Star Wars,* which deserve many labels: science fiction, action, melodrama and narrative cinema (and it has many syntactic similarities to a Western at the same time). Genres evolve in history, and the same text may be seen as a representative of different genres to different readers at different times.

A decade before Altman's article appeared, Gérard Genette published an even more nuanced treatment of the issue. In *Architext: an Introduction,* Genette analyses the classic genre division of drama, narrative and poetry. Where Altman finds two different aspects necessary to recognise a genre, Genette finds at least three: theme, mode of enunciation (narrative is recounted, drama is "shown"; a text may be in first person, second person, or third person and so on), and what he describes as "the formal 'medium' of imitation (in what language, in what meter, etc.)" (78). Realising that any text may belong to several genres, and that genres overlap and include each others, Genette prefers to describe these relations as a

network. His name for this network is *"the architext,"* and he describes it as

> that relationship of inclusion that links each text to the various types of discourse it belongs to. Here we have the genres, with their determinations that we've already glimpsed: thematic, modal, formal, and other (?). [...] So let's call architextuality the relationship between the text and its architext. [...] The architext is, then, everywhere – above, beneath, around the text, which spins its web only by hooking it here and there onto that network of architexture. (82–83)

In order to understand, explore, or navigate this network, Genette proposes a long list of necessary disciplines "whose object, let us firmly state, is not the text, but the architext [...]": theory of genres (genology), theory of modes (modistics, including narratology), theory of figures (figurology), theory of styles (transcendent stylistics), theory of forms (morphology, including metrics), theory of themes (thematics) (83–84).

Approaching genres and language use in general from different angles, several authors arrive at conclusions very similar to Genette's claim of a network of affinities. Authors inspired by cognitive psychology, such as Edward Branigan, talk about *cognitive schema* we apply to understand anything, also texts, by patterning them to our earlier experiences; Branigan argues that narrative should be reckoned as one of those schemas. This does not seem far related from the concept of *frame analysis* Jerry Palmer borrows from Goffman's sociology, when Palmer describes genres as similar to the frames with which we interpret social events such as weddings or funerals. In M.A.K. Halliday's social semiotics, language is always used in a certain *register* that fits the type of situation (*Language as a Social Semiotic* 111). Genres are special kinds of registers in Halliday's system (134).

I believe it is fruitful then, to view genres as more or less loosely defined groups within the architextual network. As patterns of clear affinities between texts, genres function for readers as frames of understanding, as they help the readers to predict what will happen and

how they should react. Even if these predictions turn out to be false, and the reader thus is surprised, the very expectation assisted the reading process. Similarly, authors position their texts within an architextural network of genres to ensure a "correct," "intended," or "effective" reading of their works.

A medium is different from a genre in that it is more stable, and defined in technological, economic, and sociological terms, rather than by thematics, form, mode and style. A medium is dependent on one particular technology, or more precisely, on one particular use of a technology. This use is well established within a society, where it can be seen as performing a certain function in economy and politics as well as in the relation between people, and it is regulated by law. Media do not exist within each other as genres may, and the perception of what medium a text is made within does change slowly in time.

If we are right in concluding that the Web should be seen as comprising both a number of media and a (much larger) number of genres, it may still be difficult to empirically draw the border lines both between different Web media and between Web media and Web genres. It further complicates matters that the Web is used for forms of communication that is not just the one-to-many mass communication pattern we usually have studied within media studies. As for our examples above, I suggest we would have to speak of something along the lines of "Web purchase" and include *Amazon*, "Web governmental services," including tax form submission, "commercial public relations on the Web" for *Microsoft.com,* and "private Web utterances" including my daughter's home page. Are all of these media in the complex sense? They certainly are stable patterns of use of a technology with a recognisable economy, legal status and societal function, but at least two of them (the book store and the tax form page) allow for the exchange of more than ideas or stories.

The distinction between medium and genre is a conceptual one, and dependent one one's perspective and objectives. It may benefit different studies to draw the

border lines differently and thus highlighting different aspects of texts. For these essays, I have chosen to limit myself to Web sites that are very similar to traditional media both in form and function, most of them even made by companies with a long history in print or broadcasting (a decision discussed under 1.1.6 below). Thus, they are all examples of the same Web medium, which we might call "mass communication." I have also stuck to one particular genre that is thematically defined, namely the longer non-fiction documentary or feature. As a consequence, I have not had the need for a further discussion of Web media and Web genres in the analytic essays.

A few words are also in order about a third distinction (in addition to media and genres) I have made use of in some of the essays collected here, namely that of text-types. Seymour Chatman asserts that it is "customary" to list four text-types: *Narrative, Description, Argument,* and *Exposition* (6).

I have found the terms useful, although I realise that the status of the text-types is a little ambiguous. Chatman himself rejects *Exposition* as a useful category in itself, as "some rhetoricians believe, there can be no expounding or explaining that does not entail a degree of description and of argumentation" (6). This is the first hint that the list of four text-types will not hold as a typology; it is clearly not the same criteria that divide between the different text-types. *Exposition* is a purport, a text made to help readers understand something complex or new to them. *Argument* is also a purport, aimed at persuading an audience and possibly change their minds: "Arguments are texts that attempt to persuade an audience of the validity of some proposition […]" (9). *Description* and *Narrative* are not purports, but semantic structures: "[W]hat makes Narrative unique among the text-types is its 'chrono-logic', its doubly temporal logic. […] Descriptions render the properties of things […]" (9). From these definitions, it is possible to imagine a text that would belong to all four categories. Chatman is in fact not opposed to this, and does state that a text very well may contain description, narrative,

and argument, as they "routinely operate at each other's service" (10). He is still able to classify a text as one of the three text-types by claiming that each text sets up a hierarchy, where one or two text-types are "at the service" of another "overriding" text-type.•

Another source of ambiguity is the similarity of the text-type to genre systems. In classic treatments of genres, (epic) narrative is often seen as a large, overarching genre like drama and poetry. Rhetoric tradition is primarily about arguments, with the classic genre division into forensic, deliberative and epideictic. If we allow *Exposition* to be a distinctive category, we see that also this text-type is a genre or a range of genres, treated in didactics and pedagogy as well as journalism, and blossoming in modern genres such as popular science and "edutainment." Chatman is unclear on the relations between genres and text-types: "By 'text-types,' further, I mean something other than genres. Genres are – at least in one sense of the word – special *subclasses or combinations* of text-types" (10, my emphasis). The peculiar phrase "subclasses or combinations" indicates that genres may exist both below and above the division of text-types, and must imply that text-types divide along other lines than genres altogether.

I do not wish here to resolve these matters, or to save or condemn the concept of text-types forever. Instead, I will try to explain how I use the terms *Exposition, Description, Argument* and *Narrative* in these essays.

Like Chatman, I treat *Exposition* as different from the others. I will not argue that the text-type is nonexistent, as Chatman seems to think, but I see it as a purport, a pragmatic concept where the other three text-types are semantic. *Description, Argument,* and *Narrative* are different ways in which signifiers construct their signifieds, or in Charles Sanders Peirce's terms, how a sign determines its interpretant.

Peirce's typology of signs in regard to their representative power – icon, index, and symbol – is the most famous, but not the only typology of signs. Among his many trichotomies are also the three kinds of interpretative power a sign may have: *rheme, dicent,* and

• It is hard to see how this notion of hierarchy solves problems of classification. Reading other treatments of the difference between narrative and description such as, for example, Genette's "Frontiers of Narrative" or Kittay's "Descriptive Limits" clearly demonstrates how difficult it is to argue with authority which text-type overrides the others. Chatman, Genette and Kittay all provide many examples where the borders seem to blur into invisibility.

argument.• A *rheme* is a possible quality, such as 'blue', or 'tall'. A rheme is a predicate that may be coupled with a subject to form a proposition: 'the sky is blue', 'John is tall'. Such a complete proposition is a *dicent sign* ("Sundry Logical Conceptions" 275 ff.; "Nomenclature and Divisions of Triadic Relations" 299). When I use the term *Description* in textual analysis, I use it as a more common term for dicent signs or combinations of dicent signs, forming complex propositions of the qualities of a subject such as a person, a place, a thing, or a period of time.

An *argument* for Peirce is "a sign whose Interpretant represents its Object as being an ulterior sign through a law, namely, the law that the passage from all such premisses [sic] to such conclusions tend to the truth" ("Nomenclature" 296). James Jakób Liszka's way of explaining the argument sign is perhaps more readable: the interpretant of an argument is not only its conclusion (the product), but also "the rule of inference, or leading principle, in the argument" (the process), and thus the argument sign also has an additional power (the effect) to change the thoughts of the interpreter (42).

When I speak of the text-type *Argument* in these essays, I mean the parts of texts where meaning is produced by the structured patterns of argument signs. Similarly, I view *Narrative* as a certain pattern of signification, giving rise to meaning not only from its parts, but also from its structure. Chatman points to narrative's double temporal logic of *story* and *discourse* (the retelling of the story as a narrative) in the definition cited above, but although necessary, this temporal construction does not seem sufficient to create a narrative. Edward Branigan has argued that a narrative also needs a continuous centre of the events and causal chain of development (19–20; See also Hayden White for a similar position). To Branigan, Narrative is a cognitive schema, a significant structure we apply in order to understand texts (13).

My use of the terms *medium, genre* and *text-type* may thus be summarised as a set of inclusions. *Text-types* are different relations between a sign's interpretant and object. *Genres* contain text-types, and a text-type may be

• Peirce notoriously uses many different terms for the same concepts. I am here using the terms recommended by James Jakób Liszka in *A General Introduction to the Semeiotic of Charles Sanders Peirce.*

one of the defining characteristics of a genre: a detective novel is a narrative, a sermon is an argument. A text-type is not enough to define a genre however, genres are also recognised by their architextual relations such as mode and medium (Genette) or syntactic and semantic content (Altman). *Media* contain genres, and are usually associated with certain genres, but architextual genre characteristics are not enough to describe a medium. Media are social institutions that also are defined by their technology, economy, legal status and use in society.

We will now look more specifically into how these perspectives were applied.

1.1.6 Research Strategies

Two connected, but different strategies are followed in the collected essays. The first strategy is a theoretical discussion about the semiotics of computers and the Web as what is called an "interactive medium." It is an attempt to understand the basic building blocks of the rhetorics (the "languages") that are supposed to converge, and to develop a vocabulary to describe them in. This discussion also helped formulating more detailed subordinate research questions for the second research strategy, a comparative analysis of Web sites and older media.

Convergence is a word describing a process over time. To grasp this process within the format of a thesis may seem futile. In the period of 1998–2002, when this thesis was written, the World Wide Web passed its first ten-year mark. In these first ten years, it has grown to immense proportions, and no individual can claim to have an overview over what it comprises. Any attempt to map stylistic developments would have huge difficulties in selecting a corpus to draw general conclusions from. The Web is also poorly documented, most sites are not archived at all, the archives that exist are generally incomplete and not available to the public. In 2002, the Internet Archive developed a interface to the Web, and made its contents available.• Maybe it now will be possible to write such a history. Then again, although

• The interface is called "The Way Back Machine," 23 Sep. 2002 <http://webdev.archive.org/>

the eleven years that has passed since the Web was born has spawned an impressive growth, its history is still so short it may be difficult to establish the necessary distance to write a stylistic history of this many-faced medium.

A conclusion to this could be that this study was undertaken too early. I believe it was not. The texts I have studied both show a rich variety of forms on the Web, and that some of the more advanced forms already had existed for several years when this research was begun. The development of the Internet is not always as fast as some would like to have us believe. My examples may perhaps later serve as documentation for a Web history, but they do not form clear historical patterns to me.

Another approach, tempting to many, is to muse on the process of convergence in the future. With some basis in what we have observed and observe now, we might project our hypotheses beyond the horizon. I have not wanted to do so. Above, I cited a few similar and presumably well-based predictions that have been made of media's convergence in the past. In 2002, we are still waiting for them to come true.•

My choice has been to study the Web as it appeared in the period when research was undertaken, that is from September 1998 through September 2002. Some Web sites from before 1998 are included in the study: as they were still online, they were still a part of the World Wide Web at the time. When asking whether the means of expression of earlier media are converging on the Web, my attempts of providing an answer have been to look for evidence of some convergence that has already taken place.

I have not attempted to study the Web or any part of it as a whole, trying to cover, for instance, the diversity of genres or kinds of content on the Web compared to earlier media. Again, an overview of the Web universe is too hard to establish. Earlier media exist within national or regional borders, and legal and economic constraints have resulted in relatively few publishers or broadcasters within each region or nation. As the Web both is truly

• This point is developed further in the essay "Reading-View(s)ing the Über-box" (chapter 2 in this volume).

world wide (although the world's richer regions are far better represented on the Web), and the threshold for new publishers is low, it is virtually impossible to establish a meaningful body for comparison. Any selection of Web sites will be a compromise and hardly representative of the Web as a whole.

My very limited operationalization is thus to select some small samples and study them in detail. The only finding that could be generalised from such a study is that of possibility. If I find convergence, then convergence is possible in this form at least. To discuss the mix of forms, I have compared Web sites with the older media print, radio, and television. To establish the grounds for such comparison, I have chosen Web sites of established media institutions. Sites of broadcasters and publishers are compared to their output in broadcasting and publishing. Furthermore, I have limited myself to one family of genres: longer non-fiction reportage, which is known in print as the feature, and in film and television as documentary. Non-fiction is found widely in all the media I study, while fiction is rare on radio and the Web. Longer reportage was chosen, as I believed it to give more interesting material for analysis. A longer story can use a wider range of styles, and provide more challenges for creating unity and storyline.

The limited approach of studying a small sample of texts has been the choice in studies of the aesthetics of film, literature, or music, and is thus in line with the rhetorical and semiotic traditions this thesis refers to. But again, the decision of what work is worthy of the investment of analysis is troublesome. The most common solution is to study canonical works, works that are seen to have an aesthetic value in themselves. Other works are not found worth the effort. While *Citizen Kane* is widely studied, I believe there have been few detailed readings of the movie *Police Academy 5.* A more sociological approach is to study a work known to be popular or typical. However, to know what is typical brings us back to establishing a universe, which is a problem on the Web. Even knowing what is popular is a

challenge. Audience measurement is done on the Web, but the methodological problems abound.•

To an extent, I have tried to make a commonsense choice combining the two approaches (the canonical and the sociological). I have looked for sites listed as among the popular by audience measurement firms, in awards for excellence in design such as the Webby Awards, and I have looked in the "showcase" listings published for promotion by makers of multimedia software, such as Macromedia's Flash and Apple's QuickTime. I believe I have found sites that are widely read, and that the designers could be proud of at the time they were published. All the analysed texts are referenced in detail in each essay. I now turn briefly to how they were selected.

The core set of examples has been texts published by The National Geographic Society. Drawing on a tradition spanning more than a century, the Society has an ambitious output of books and print magazines, taking pride in the products of their writers, photographers and graphic artists alike. In addition, they make radio programs for NPR and television documentaries featured on television networks all over the world as well as their own National Geographic Channel. Since 1996, the National Geographic Society has also had an ambitious Web presence, regularly publishing new, extensive Web "features" which are all archived. As the Society not only has experience and quality material available from a wide range of earlier media, but also the will and the financial muscle to use the Web to full effect, I believed it would be a good source for interesting material. During September 1999, I recorded all the television documentaries aired on the European National Geographic Channel. From these I selected three documentaries on topics that also had been featured in *National Geographic Magazine* and in multimedia Web sites.•• This selection was initially made only for the study reported in the essay "Linearity and Multicursality" (the present chapter 6). The material gave rise to many new questions, however, and I decided to continue using National Geographic

• Jakob Nielsen has, for example, criticised the measurement of "unique visitors" to a site as close to useless ("Loyalty on the Web," *Alertbox* 1 Aug. 1997. 23 Sep. 2002 <http://www.useit.com/alertbox/9708a.html>)

•• A total number of six documentaries had "sisters" in print and on the Web. To be able to cope with the material, I limited myself to three, chosen to have the widest range in rhetorical techniques, combinations of sign systems, and time of publication of the Web site.

as my main source of examples. I continued to follow Nationalgeographic.com, and sixteen more Web features were added as large and small examples of new and interesting genres.

In addition to the National Geographic Society material, I have used examples from three of the most popular international news sites: *MSNBC, BBC News,* and *CNN.com,* two important Norwegian sites, *VG Nett* and *TV2.no,* and the Web site of the major Spanish newspaper *El Mundo.* These sites are typical of news sites throughout the Western world.

The last major example used the winner of the Webby Award for best science site 2001, *Becoming Human,* which is also praised by many in the Web industry, as it is an example of a site using a wide range of rhetorical techniques and was the state of the art of Flash technology in 2001.

In addition, a number of other Web sites are drawn in to illustrate smaller points. The total list of Web sites is found in the bibliography at the back of the thesis. Whether the sampling strategy was useful or not, should really be judged from the results of the investigation and the use of these. The argument put forward is not one about the nature of the Web, but about the details of signification in Web sites.

1.1.7 A Few Practical Definitions

For this thesis, I have defined the Web as what can be accessed with a standard Web browser. To me this is the most commonsense approach, the one I believe to be most similar to what most users think of as the Web. The borders are not clear-cut, but all the pages I have studied use the URI• system of the Hypertext Transfer Protocol (http). The bulk of these pages consist of HTML-coded text and inserted image files, but many also have included video clips or animations in different formats. Many of these require that the user have certain "plug-ins" installed, but I have treated these as part of the "standard Web browser." Pages often enable the user to download files that will open in another application that the browser. A typical example are streaming video

• In technical specifications of Internet architecture, URI is short for Uniform Resource Identifier, covering what in specifications written before 1998 are known as Uniform Resource Locator (URL) and Uniform Resource Name (URN).

opening in the video application RealPlayer. As long as the download process was activated in the Web browser, I have treated this as part of the Web. RealPlayer and other video applications also enable access to a number of radio and television channels via the Internet. These are not included in the study, as they primarily are accessed with the video application. Likewise, a few pages with links that open the user's e-mail application are studied, but e-mail messages that open the Web browser are not. This way of separating may not make sense technically, but I believe most users think of the Web as that which appears in Netscape or Internet Explorer, not considering technical protocols and file formats.

I have also tried to describe Web sites in widely used terms, mostly derived from popular discussions of Web design in books and news sites, although a few additions have been necessary. Thus, a HTML document (with inline images and other files included) as displayed in a browser is called a *page.* In hypertext literature, this unit is normally called a *node,* or (less fortunate) a *lexia.* In preferring *page*, I want to keep the vocabulary less exclusive. Whether a page is composed of one HTML file, or is a composite of several (using "frames," for instance) is not of importance; again, it is the appearance to the user that matters. A page may be taller than what is visible on the screen, I will thus call the visible part a *screen* or *screenful*. The act of moving down a tall page screenful by screenful is called *scrolling.* A (rectangular) part of a page, set apart visually is called a *pane*. The concept of *pane* is useful for discussing pages where only parts are changed when activating a link (for example, using "frames"), or where video is included. A page is displayed in a *window.* Windows are made using the computer's operative system, and are not treated as parts of the Web pages. However, a link in a page may open a new window with a new page in it. When links (link anchors, to be specific) are placed in text, I call the part of the text that is the link (which typically would be rendered in blue or another colour that stands out, and often is underlined) the *link text.*

Four Axes of Rhetorical Convergence

How to name a collection of Web pages is often difficult. I have tried to be consistent in using the term *a site* for a set of related pages under one Internet domain name with a shared title, a similar design, published by the same institution, and using coordinated navigation links. Closed texts consisting of several pages about a shared topic within a site, I have called a *section.* Thus I call the whole of *Nationalgeographic.com* a site, containing, for example, the sections "Return to Midway," "Photography," and "National Geographic Magazine." I have tried to avoid further divisions, but some times it has been necessary to talk of subdivisions of sections as *sub-sections* when describing the objects of analysis.•

What is a section and what is a site is often open to judgement. It could clearly be objected that when I describe *BBC News* as a site containing sections, it would be more strictly consistent to call it a section within the *BBCi* site. The truth of the matter is that both is correct, many sites are "related" in what we might describe as "families" or "cooperatives." One example is the many "Web rings," found on the net, where a number of sites link to each other, encouraging readers to visit all the sites in the ring. It will probably vary from reader to reader how a page is perceived to be related to a larger whole. Luckily, this complexity is outside of the scope of this thesis.

I have set the titles of sites and sections in italics and quotation marks respectively, mirroring how titles of books and articles conventionally are set.

Written language, still images, moving images and spoken language are examples of sign systems. In chapter 2, "Reading-View(s)ing the Über-Box," I called these *information types,* following Liestøl. In the rest of this thesis, I have preferred the semiotic concept of *sign system.* Any use of sign systems in communication will be called a *text* in this thesis, whether it is written, spoken, pictorial, a game, or a piece of music. Correspondingly, I will use *author* and *reader*•• for makers and receivers (consumers) of a text. To avoid confusion, written alphabetical text will be called *writing* rather than *text,*

• In chapter 5, "Linearity and Multicursality," I used the term *area* for what I later chose to call *section*. The two terms should be considered as synonymous.

•• In "Reading-View(s)ing the Über-Box" (chapter 2), I used the term *user*, which is more common within computer science. Later I decided on *reader*, as a term more in contact with media studies and literary theory.

unless it is clear from the context that *text* is used in the meaning *writing.*

1.1.8 Overview of the Essays

Above, I described the research process as a number of "probes" into the research area.

The first probe resulted in the two essays presented here as chapters 2 and 3, and are two contemplations on the possibilities of a total convergence of media, as predicted by Nielsen, Lippman, Negroponte, and Murray, and cited in 1.1.2 above. "Reading-View(s)ing the Über-box: A Critical View on a Popular Prediction" (chapter 2 of this thesis) took a closer look of the idea of a convergence of media in itself. If the predictions are right, that all media will converge into one Web medium, what would the texts of this medium look like? It was found that the thought of one technology taking care of all our media is at least as old as the computer, and that the dream has not yet become reality not so much because of technological constraints, but because media forms cannot be understood from technology alone.

Four aspects came up in the discussion that were expanded on in later essays: 1) The difference between static and dynamic sign systems; 2) The difference in presenting alphabetic text and images; 3) The importance of devices of the text enabling the reader to control the reading process; 4) The incompatibility of certain forms of earlier media. It was demonstrated that all four aspects represent obstacles for a total convergence of media, and that ways around them are few yet, if at all conceivable.

Chapter 3 originated in my experience of discussing my project at several workshops. I was often met with little understanding of my approach to convergence. Within media studies, convergence is most often studied as the convergence of different industry sectors, as a consequence of a technological convergence with economical and legislative results. To clarify my research aims, I started to use the term *rhetorical convergence* as opposed to aspects such as *network convergence, terminal convergence* or *corporate convergence.* I developed this

concept further in "Rhetorical Convergence: Studying Web Media" (chapter 3). The essay asked how existing theories in media studies can shed light on the convergent forms of many popular Web sites. Through a sketchy analysis of *VG Nett*'s "Trippeldrapet" section, It was shown that these sites are complex texts that cannot be accounted for using theories of older media alone. Earlier conceptualisations of kinds of convergence and Bolter and Grusin's concept of *remediation* also fall short of explaining the mingle of earlier forms in mainstream Web pages. As a solution to this, the concept of *rhetorical convergence* was introduced, meaning a process were formal elements – rhetorics – from the traditions of earlier media are combined. It was further pointed out that computer media is one of these traditions that needs to be considered. Realising that this might be an area of study in its own led me to formulate a program for further research. Picking up the thread from the earlier paper (chapter 2), it was noted that many earlier forms cannot be combined. To study rhetorical convergence is thus to map both the forms that are combined, and those that are not. A conceptual convergence would be necessary in order to describe these combinations. A discussion of whether the Web is properly described as one or several media and the relation between genre and medium was initiated, to be followed up later.

The second probe was into the growing body of theory on computer literature. It had two aims: to look for the applicability of the different theories on the Web, as they all were developed for other computer technologies; and then to investigate what the rhetorical tradition (the "means of expression") of the computer is. Hypertext theory is very much focused on issues of *nonlinearity* (see, for example, Landow, *Hypertext*; Aarseth, "Nonlinearity"; Harpold; Douglas; and Liestøl, "Wittgenstein"), or *multicursality*, as it is called in Aarseth's *Cybertext*. I wanted to see whether this could be found on the Web, and how it differed to supposedly linear media such as print or television. Media output from the National Geographic Society was chosen as a research object, I selected three TV documentaries on topics that also

were reported in *National Geographic Magazine* and on the Web site *Nationalgeographic.com.* Issues of sequence or choice of sequence were mapped in great detail, and the results were reported in the paper "Linearity and Multicursality in World Wide Web documentaries" (chapter 5). In the course of the analysis, a framework for study was developed, identifying four principles of order in linear texts, three classes of links in Web sites, and three different linking practices. Only one of these practises deserved the name *multicursal*, showing that claims that the Web is inherently multicursal or nonlinear due to its hypertextual technology are unfocused, if at all true. Sequence was found to be constructed differently in television and print, and the Web sites studied used few of the devices identified in either of the two older media. Neither the "nonlinear" stance of hypertext theory or simple convergence hypotheses could thus be sustained. However, this analysis cannot claim to explain why the Web pages look as they do. A more "balanced" rhetorical convergence, borrowing more forms from more media would certainly have been possible, as is pointed out several times in the essay.•

While studying linearity and multicursality, I started thinking about the semiotics of links. An apparent problem was that while it seemed to me that semiotics is a fruitful perspective to describe both multimedia and linking and Web sites, it was equally clear that Aarseth's concepts of ergodics and cybertext were necessary. To align the two perspectives seemed difficult all the while Aarseth has openly criticised semiotic approaches to computer media in general, and those of Jensen and Andersen in particular. The present chapter 4, "Semiotics and the World Wide Web" undertook a close reading of Aarseth, Jensen and Andersen. Where Andersen uses Hjelmslev's glossematics as his starting point, I attempted to be consistent in using Eco's semiotics, which seems to escape much of Aarseth's criticism. Furthermore, I claimed that Aarseth's argument is insufficient at places. The conclusion in the chapter was that an adaptation of Eco's semiotics in no way

• The recommendations for hypertext design found in the essay should be read with the fact in mind that the paper initially was written for the *Hypertext* conference, which focuses much on hypertext design.

contradicts the theory of cybertext, and that semiotics is a powerful perspective for explaining both the workings of computer interfaces and Web linking. Leaving Andersen's typology of computer signs, and instead viewing links as signs of available processes, seems to add a rhetorical perspective that perfectly balances the material and mechanical perspective of cybertext.

The third probe in this research project was initiated the autumn of 2001. As I realised that although multimedia was a central issue in discussing computer media in general and rhetorical convergence in particular, there is very little theory on the subject. I started to collect a catalogue of conventional combinations of different rhetorics as seen on the Web (a research strategy also outlined in the essay "Rhetorical Convergence," chapter 3). A pattern slowly emerged, showing that some convergent forms begin to establish themselves as conventional, some even as proper Web genres. The last four chapters investigated some of these genres and forms, keeping the comparative approach that was used in "Linearity and Multicursality" (chapter 5).

"Interactive Graphics" is the title of the essay that here is chapter 6. With a basis in my National Geographic Society material, explanatory diagrams in print, television and Web sites was compared, to see if (and possibly how) their different uses of graphics were converging. The comparison made the difference between static and dynamic sign systems come out clearly. Static images allow for comparison, dynamic images may show processes unfolding, and the dynamic also attracts and directs attention. Four different kinds of interaction were identified: (1) Alternating what is visible in the display, (2) controlling the display of sequences of images, (3) using images as maps of hypertexts, and (4) the direct input of data to be displayed. The first two of these figures were conceptualised through the term *canvas*, the physical appearance of the computer display, and were viewed as means to overcome the poor resolution of computer screens.

Online photo collections was the theme of the essay "Presenting Photographs" (chapter 7). How did Web

sites present collections of photographs, and was this different from how earlier media use sets of multiple photographs? Three different genres of display in Web pages were compared: image collections, image exhibitions, and slide-motion movies. These different ways of presenting photographs are not neutral, they have different implications for the reading process. When several images are laid out on one page, their relative size and placement will assign a different priority to each image. When images are strongly prioritised, the reader is less encouraged to select among them. Even less empowering is the choice of presenting images one after another, as in slide-motion movies. The chapter also discussed how different forms of presentation tend to align themselves with different photographic genres. The power relations of images and words and relations of anchoring and description are discussed with reference to Roland Barthes' work on photography. These issues are further complicated when still images are "made moving" in slide-motion films. Movement alters the original composition of photographs, and tends to narrativise them, it was argued.

The relation between images and language was the topic also in "Combining Video and Writing," the present chapter 8, but this time moving images were discussed, trying to find established conventional ways of combining video and writing, and comparing these with earlier media. Video is not a part of standard HTML, and this puts limits to video use on the Web. Writing and moving images cannot merely be aligned next to each other; one has to be contained within the frame of the other. The contained sign system will then be constrained under the rhetoric of the container sign system. Different uses of video and text combinations were shown, to varying degrees using moving images' abilities of description. Different kinds of controlling playback and their effect on the video were discussed. Devices such as chaptering or video-on-demand tend to take authority and the delay of closure away from the video, leaving many television forms irrelevant on the Web.

Throughout this thesis, I acknowledge that there are existing traditions of computer rhetoric. By far the most commercially successful is computer games, and it seems reasonable to ask how aspects of computer gaming may have converged with forms from other media. The essay "Quests and Worlds: Two Devices of Computer Game Rhetoric" (chapter 9), discussed game-like texts and navigable spaces together because of their affinity to computer games. This was shown to be an artificial combination, as the two are really different forms. Game-like texts are linear texts, but with delays in the line. Navigable space is not used for gaming, but as an interface or a "table of contents," or alternatively, as a description or simulation of an environment. A third use of 3D space – rarely implemented – is as a container for a collection of texts using many sign systems. Typical for both these approaches is the creation of a claustrophobic text, where "inside view" replaces overview.

Each of these chapters discusses some themes related to the overall research aim, that is to understand how the rhetorics of earlier media converge on the Web. The insights from the different chapters will now be viewed together in the second half of this essay.

1.2 Understanding Rhetorical Convergence

Rhetorical Convergence is the combination in one medium of rhetorical forms or devices that were earlier only seen in separate media. This thesis documents and studies rhetorical convergence in Web media, as rhetorics from television, print, photography, radio, and computer media are combined in Web sites.

In my essay "Rhetorical Convergence" (chapter 3), two tasks for the study of rhetorical convergence were identified. The first is to collect a catalogue of forms of rhetorical convergence; this task is begun in the essays assembled here, especially in chapters 6 through 9. As we have seen, the number of Web sites studied is small, and the sample strategic, but I believe a number of convergent rhetorical forms are found and identified. As demonstrations of rhetorical convergence, I believe the essays speak for themselves, so I will not repeat these findings here.

The second task of the program was to chart the ease or difficulty with which rhetorics in different media combine. Which forms may be merged, which have to be altered, and which are mutually exclusive? I have begun to answer this in the following essays, and the insights from these attempts are collected and developed into a theoretical framework in the remainder of this "Final Contribution." In the following sections, I will propose a *multidimensional model of media rhetoric, collected in four "axes"* that enable an overview of the different convergences of media forms.

The essays that make up this thesis span a multitude of examples, all perceived to be mixing rhetorics of several established media. Each example shows a form very similar in one respect to what we are used to see in an earlier medium, but in another respect just like another earlier medium; just as a daughter may be the spitting image of her mother, yet no one can deny she has her father's eyes. I have collected all such similarities "in one respect" and "in another respect," and tried to reduce the number of "respects" by assembling them into groups and constructing a theoretical model.

The purpose of the model is to enable us to view all my examples together, and the ambition has been to construct a model so simple it gives an overview, so general that all my examples can be described according to it, and so complete that all my detailed findings can be positioned within the model.

This is done regularly within statistics; one may for example measure a set of objects in many ways with many different instruments. For the statistician, each reading would then be one dimension. Using mathematics, the number of dimensions may then be reduced, for example to three, by projecting the measurements down to three dimensions. The objects may then be compared using only these three numbers, a much more manageable task.

When Web sites in this thesis are compared to texts in earlier media, the many aspects that are discussed cannot always be expressed quantitatively. I am not proposing a mathematical or statistical model, I am

using statistical methods as an analogy for the process of collecting related aspects into larger categories.

This is not without precedence in studies of texts. Barthes' method of textual analysis, developed over a number of essays and perfected in *S/Z*, consisted of dividing the text into small units, and registering the presence of five co-present "codes." There are several different sub-types under each code. Both chronology, descriptions of time and place, and common sayings are, for example, coded as the "referential code," as they all refer to common knowledge within a culture. The code of actions consists of all the diverse actions performed by the characters in a story. To explain this, Barthes used the analogy of a polyphonic musical score, where several voices develop relatively independently in the same piece. Genette, on the other hand, used the term "dimensions" in *Architext* when arguing that genres are recognised by (at least) three aspects (which we listed under 1.1.5 above), aspects that in their turn are made of different parts or dimensions, such as different meters or the enunciative modes (of which all of narratology is just one part).

I am using the analogy of "axes" for this way of ordering the large number of different qualities a text has, so it is easier to get an overview over all the details. Still, it is not my ambition to *reduce* the complexities of architextual references we here call rhetorical convergence. Rather, I want to present a perspective that enables us to see the full range of complex relations, just as Barthes did in *S/Z* and Genette in *Architext*.

The inductive process of constructing the model has not been entirely data-driven; earlier literature has provided much insight and inspiration. One cannot read much into the topic of computer media without noticing the two aspects we might sum up under the headings *multimedia* and *interactivity*. Under the heading *multimedia* I put texts about the computer's ability to combine different sign systems such as writing, images, video and music. These include works by Nelson (*Dream Machines*), Negroponte, Liestøl ("Essays"), Bolter and Grusin, Manovich, and others.

Although the term *interactivity* is contested, I use it here as a heading for literature concerned with how computer texts respond to user input, and include texts by Nelson (*Dream Machines*), Negroponte, Bolter, Lanham, Joyce (*Of Two Minds*), Bernstein ("Patterns of Hypertext"; "Hypertext Gardens"), Landow (*Hypertext*), Aarseth ("Nonlinearity"; *Cybertext*), Liestøl ("Essays"), and others.

Several writers, including Bolter, Joyce (*Of Two Minds*), Aarseth (*Cybertext*), Landow (*Hypertext*), Greenspun and Veen (*Hotwired Style*), have also commented on the fact that hypertext systems such as the World Wide Web are simultaneous media that allows for continuous updates and alterations, and may allow users to comment or modify texts. Early on I realised that this could be treated in relation to the stress on live broadcasting within radio and television theorised by Ellis, Hjarvard, Eco (*Kant and the Platypus*), and others. A Web site may be live and archive at the same time.

Writers on Web design (of which there are many, including Sano, Siegel, Veen (*Hotwired Style; Art and Science*), Black, Greenspun and Nielsen (*Designing*)) never fail to stress how Web authors are constrained by technical limitations such as bandwidth, screen resolution and colour depth. My analyses of the selected Web sites showed that much of the "interactivity" in the texts actually were remedies that compensated for low screen resolution and bandwidth. This led to the formulation of the concept of *canvas,* which will be introduced under 1.2.2.1 below.

With these texts in mind, I chose to project the many dimensions of Web rhetoric on four axes. Multimedia, interaction, "liveness" and technical limitations on the Web were refined into four axes of rhetorical convergence and labelled s*ignification, acquisition, distribution* and *restrictions.* I have chosen to call them "axes" for several reasons; the first is that they are co-present. All kinds of media communication I have studied in these essays are characterised by how their authors have chosen to use these four aspects. They are different sides of a text, different dimensions. Furthermore, each of these four

aspects are ranges of choices on which texts will position themselves differently. A broadcast is either live or not. A text may be written, spoken, or both. When discussing the four axes below, I will discuss different modalities, or positions, along each axis, and a text's position will thus be called its "mode of" distribution, acquisition, restrictions, and signification.

We will also realise that the axes are projections of aspects of rhetoric that in themselves are multidimensional. The axis of distribution, for example, contains bandwidth, creation latency, transmission latency, and permanence. The axis of restrictions contains range and detail in time and space.

I propose these four axes as a projection, a way of getting an overview over the many dimensions of rhetoric. I think of rhetorical convergence as taking place in a four-dimensional space (to the extent I am able to visualise such a construct). When we discuss popular media what we tend to discuss is really their typical genres. Examples from television might be the news show, the serial, the soap opera and the talk show, from newspapers, the report, the feature, and the op-ed. Each of these media-specific realisations of different genres would occupy a unique space in my four-dimensional space if their modalities along the four axes were mapped. Media texts that combine aspects of different genres would then position themselves *between* the positions of established media, similar to one genre along one axis, similar to another genre along another axis.

We may thus rephrase rhetorical convergence as being a form from one medium or genre that is copied in another medium, but where the copy differs from the original by substituting its mode on one or more of the axes with the mode of a form we recognise from a different medium's rhetoric.

Before the computer, the *mode of distribution, mode of acquisition, mode of restrictions,* and *mode of signification* were usually given by technological constraints or conventions of the medium. What makes rhetorical convergence an interesting study object is that in a

computer medium, few of these aspects of rhetoric are given. Instead they are ranges of choices to the author.

At the same time, the four axes also spell out constraints to rhetorical convergence. Rhetorical convergence is understood as a rhetorical form that is a blend of two well-known earlier forms. Some traits from one or more axes are recognised from one earlier form, some traits from one or more other axes are from another. But, as the modalities along each axis are distinct, the new form cannot inherit two different modes from the same axis. It has to be a blend of modes on different axes, not of modes from just one axis.

After these initial remarks, we will proceed to describe the modalities of the four axes.

1.2.1.1 Mode of Distribution: Bandwidth

Let us use the terms *creation* and *distribution* to describe the process of getting the messages from author to audience, a process that takes time. Authors are aware of the time it takes, and this influences on media rhetoric. Newspapers are printed during the night and distributed to subscribers and newsstands in the early morning. It is a carefully timed process, and all journalists have to deliver within deadline or the newspaper is delayed. In broadcast media, the distribution from sender to receiver is almost instantaneous, but the creation of the broadcast takes a lot more time. Recording and editing sound or video may take a lot of time, and the raw material has to be transported to the editing facility, and then to the sender facility. For fast evolving news, the first reports are likely to be read from studio or reported on telephone from a reporter in the field, and only in later newscasts does edited video material appear. That creation and distribution times determine the rhetoric is true for all media. But the Web is fundamentally different from earlier media as its distribution time is variable.

A Web page is not loaded before the reader asks for it, by typing in a Web address (a URI) or following a link. Following Jan L. Bordewijk and Ben van Kam's typology from the article "Towards a New Classification of Tele-Information Services" the Web's traffic pattern is

consultation, as it is a relation with one sender (the server of the page being read) to (potentially) many receivers, and the receivers initiate the transaction. In contrast, the traffic pattern of earlier media is *transmission* in this model. In order to see how this difference makes Web rhetoric different, it is useful to look at three different aspects of distribution which we will call bandwidth, latency, and permanence.

As the Internet is a packet-switching network of many different networks with different standards, the capacity of transmission, known as bandwidth, varies enormously. I am writing this on a computer with a connection of 100 Mb/s, which is several hundreds, or even a couple of thousands times faster than the 56Kb/s dialup modems still used by most home users, or the bandwidth of mobile phones or PDAs with Internet connection.• If the files that are transported over the net are small, however, the user may not necessarily notice the bandwidth difference; research in Human-Computer Interaction has shown that we experience everything that takes less than about a second as instantaneous (Nielsen 42–44). It does not matter if it actually takes three-quarter of a second or a thousandth, we can't tell the difference. And for static pages (without sound or moving parts), Nielsen's studies conclude that we normally do not care much about differences of a couple of seconds either (44). But some sign systems are slower to distribute than others. When different sign systems are coded in computers, the file formats takes up different amounts of space. Text and numbers are coded extremely economically. Images are coded less economically, sound takes up more space than images, and video takes the most space. Images, sound, and video may be compressed to be less voluminous, at the expense of image or sound quality. Heavy compression also requires more computer power, both to "wrap for shipping" through the net and to "unwrap" or "deflate" at the receiving end. A slow computer will thus slow down the display of compressed media, but again, as long as the decompression takes less than a second, the reader won't notice.

• This is just an example of the variations. The actual bandwidth any one user achieves will vary from site to site, depending on the speed of the server, the total of the network (servers, routers, lines...) between the user's client machine and the server, and the speed of the client machine.

A Web designer who is aware of this fact will thus face choices that are unique to computer communication. Many pictures in a magazine or quality footage in television may slow down the production process and bring up production cost, but the images always arrive in the newsstand with the rest of the magazine. Television would be a very different experience indeed if we had to wait for a video segment in television news, while a talking head was instantaneous. In Web television, this may very well be the case, so in Web design, "heavy" or "slow" elements such as big colour pictures or video have to be used sparingly if the pages are meant to be appreciated by any Web surfer.

The "interactive feature" "Sights and Sounds of the Way West" in *Nationalgeographic.com* provides an example of an aesthetic grown out of bandwidth concerns. "The Way West" is a movie created solely of still photographs that are "made moving" as if filmed by a movie camera that moved over each image, zooming in or out. The projection of the show is timed to a soundtrack with music, sound effects and narration.• Animating still images by moving would only be used in mainstream television when there is no video footage available, as often is the case in history documentaries. In "The Way West," on the other hand, the whole feature is made in this manner. In the introduction, for example, photographer Jim Richardson who also is the narrator introduces himself in voice-over, illustrated with two shots of him doing field work. In television, it would be rare indeed not to show the moving image of Richardson speaking. On the Web, on the other hand, it makes more sense to do it in this way, as the still images made moving can be large, fine-grained and have rich colour (high resolution and colour depth) while keeping the movie files relatively small. Video would take up much more computer memory, and thus require more bandwidth (or patience during download) to be played. In addition, still images look less out of place on the Web than on television. On TV an image with no movement is extremely rare, while most Web pages have still images.

• "Sights and Sounds of the Way West" is studied in detail in the essay "Presenting Photographs" (chapter 7).

Four Axes of Rhetorical Convergence

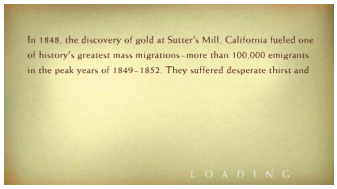
• Flash is a file format for animation films made by Macromedia, that over the years has been expanded so Flash movies may contain photographs, written text, video in other formats, and simple computer programs. Flash files are stored as vector graphics, allowing them to be small and bandwidth efficient.

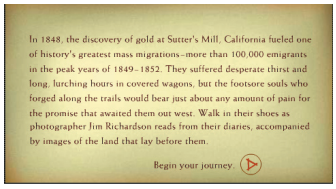
The introduction sequence of "The Way West" shows the influence of bandwidth concerns on rhetoric in more detail. "The Way West" is made in Flash,• and requires that a fairly large file be downloaded before one can start watching. To cover the download time, the authors have taken advantage of the fact that writing loads faster than recorded sound and colour photographs. While download still is in progress, an animated title sequence appears.

http://www.nationalgeographic.com/ngm/0009/feature2/media2.html

The title screen shown at left appears one word at a time. After a while (how long depends on the download), the letters are animated as "blown away," accompanied by wind sounds.

A written introduction to the sequence then appears three lines at a time. The appearance is timed to the download, and on a fast connection, the next three lines will appear a little before one is finished reading the previous at normal reading speed.





A little slower than a television introduction, this "splash sequence" as it is called in industry lingo cleverly fills the download time with interest, and at the same time makes a smooth transition to the moving images that are to follow.••

This is just one example of how slide-motion films such as "The Way West" stand midway between television and print rhetoric. The introduction contains more writing than is normal on television, but it is dynamic, it appears over time. In the body of the feature, narration resembles radio (as it with few exceptions is totally

•• Splash sequences are treated in more detail in "Combining Video and Writing," chapter 8.

comprehensible without the images), but is illustrated; the illustrations are not video but still images made for print; yet they are not completely still, as they are viewed through a moving frame. "The Way West" is not exactly either print, radio, nor television, but borrows a little from each; and it makes a reasonable compromise of bandwidth concern and television rhetoric.

The concern for bandwidth is principally about reducing the time readers have to wait for a page to load. I call this waiting time *latency*, and it is worth considering in itself.

1.2.1.2 Mode of Distribution: Latency

Broadcast media – radio and television – eliminated distribution time when they were introduced. Radio or television signals travel from sender to receiver in almost no time, and in the years before storage media such as magnetic tape were introduced to broadcasting production, there was no delay between creation and broadcast either. Nowadays, most television programs are broadcast from an edited tape, but the possibility (and the history) of broadcasting live is still what gives television both its particular aesthetic, and its particular place in our lives, according to Raymond Williams, John Ellis, and Umberto Eco (*Kant and the Platypus*), to mention just a few.

We saw in 1.2.1.1 that bandwidth is a concern to all Web designers, as the Web is slow for many users. Still, live broadcast is an absolute possibility on the Web. In order to see the differences and similarities between broadcasting and Internet media, we have to consider two different measures of network speed. Computer scientists speak of *bandwidth* and *latency* as two related, but not identical features. *Bandwidth*, as we saw above, is a line's capacity to carry large amounts information. *Latency* is the speed with which a single packet can be transmitted over a certain line. A simplified explanation could be that *latency* measures the time it takes before the user gets the very first part of a message, while *bandwidth* measures the time it takes before the user gets the very last part, and the message is complete. A

physical line has the same bandwidth all the way, while latency increases the longer the line is.• An example may clarify the distinction: With popular "streaming media" formats such as Real Media, Windows Media or QuickTime, it is possible to broadcast live video even over narrowband modem connections. The image will be smaller and with less detail than what would be possible on a broadband connection, but the latency is short enough that it is reasonable to call it live.

Both latency and bandwidth may determine the rhetoric chosen, as authors strive to match their texts to the expected bandwidth and the latency they wish to obtain. The "Web TV" made by *Yahoo! FinanceVision* is an excellent example of this.

Until 2001, *FinanceVision* was a live "Webcast." It was a page where a video pane showed live streaming video of a talk show about finance and stock markets in fairly traditional television style, while stock quotes and other investor information appeared in another pane whenever a company was mentioned. *FinanceVision*'s creators chose to make it a live show, maybe to capture a fast-moving market, maybe to obtain connotations of "being there when it happens," maybe both. The visual style and the semi-scripted conversation closely resembled what can be seen on television every day in a lot of countries. At the same time, keeping the clothes of the presenters in few, clear colours and having a similarly clean and monochrome studio made video compression more effective, as there was less information in the image. Few camera movements and edits, and programme hosts sitting fairly still added to this. By ensuring there is as little difference as possible from one frame of the video to another, compression algorithms work more effectively to get high quality video over low bandwidth, while ensuring a short enough latency to call it live.

Stig Hjarvard defines *live* as the near simultaneous events of media creation, broadcasting and consumption. *Near* simultaneous, for in fact, there will always be some latency – for satellite broadcasts, it may approach a second. Following this definition, Web cameras are also

• I am indebted to Bjørn Remseth for pointing out this distinction to me.

"live." Web cameras take snapshots of a scene at regular intervals, for example, every minute, and the latest image is always available from the server connected to the camera. Many Web camera pages have text next to the camera pane reading "watch place x now!," and the parallel to television cameras used for surveillance is easily drawn, as Bolter and Grusin have treated in depth. Bolter and Grusin also point out how similar the promises of Web cameras are to television's aura of being live.

But if a Web camera is "live," then any Web page is live, as it is available from the moment it is put on a server. During the high-profile Orderud murder trial in Norway in 2001, tape recorders were prohibited in the courtroom. Bypassing the ban, a journalist of the Web newspaper *VG Nett* did his best to type everything that was said throughout the trial. Every time he finished a couple of sentences, he pushed a button, and the new lines were sent via a wireless Internet connection and added to the transcript on *VG Nett*'s Web page.• After Hjarvard's definition, this is live, as creation, distribution and consumption may be almost simultaneous. What differs is that the events and what is said is mediated by a journalist, and written by him on a keyboard. This is a slower and more filtering process than a video camera, and outside what is normally thought of as live.

• This example is also discussed in chapter 3, which contains illustrations from the Web site

These examples show that latency is a variable on an ungraded continuum. The shortest latencies will be felt as "live," but what the border values are cannot be pinpointed exactly. Towards the other end of the continuum, we find delayed release. Up to this point we have discussed as if media texts are always distributed as soon as they are ready, which of course is not always the case. Most earlier media are periodic, and a story has to wait to be published in one "issue," whether it is a printing or a special segment in a broadcasting schedule. On the Web, a story that is kept from immediate publication can be released at any point in time; whether to publish at the same time every day or week, whenever material is ready, or even "live" is up to the author to choose.

Latency is thus the sum of two parts: the time from event to broadcast, which in this example of a live show was (almost) none; and the time from the user enters the page to the video appears in the video pane. The first, which we may call *creation latency,* is a rhetorical choice the authors made. *Transmission latency*, the other part, is a physical limit of the line, the file sizes, and the speed of the computers involved.

Like bandwidth, the transmission latency needs to be considered by Web authors if they want to achieve a low latency overall, such as live or almost live distribution of an event. The (rhetorical) choice of whether to transmit live or not may also be a consequence of the third aspect of distribution: *permanence*.

1.2.1.3 Mode of Distribution: Permanence

A certain kind of live broadcast that is very widespread on television becomes very peculiar on the Web: the practise of interviewing someone live during a newscast.• What is interesting for our present discussion is not the interview as such, but the fact that it is conducted live during the newscast, while all the information normally has been known for several hours. Very little news happen to take place in front of a filming camera team during the time the newscast's program slot. Live interviews in television newscasts appear to be very current, bringing the very latest information, but in fact they only bring information that has been known for quite a while, at least by the news desk. More important than the actual freshness of what is said is the *image* of freshness. Live interviews project an image of the television news desk as "standing in the middle of the stream of current affairs," as Hjarvard puts it.

• The live interview example is used several times in this thesis, as it is one of the most illustrative examples of rhetoric that does not easily move from one medium to another.

A live broadcast needs an audience to be worthwhile, however. It does not seem likely that a Web news site would put up a Web camera and post a journalist in front of it, so he could be interviewed live whenever someone logged on to the server. Live streaming on the Web is thus normally reserved for special, heavily marketed events such as sports events or important speeches. Another example is the "Web meeting," a popular genre of live

transmission in which a Web site invites an interesting person to sit for an hour or so and answer questions from audience members. The questions and answers are published on the Web page in real time as writing, sound or video and sound combined. Unlike television news, such meetings have to be advertised over a time period in order to attract an audience. On television, in contrast, the audience shows up every night to watch the news without being asked. The evening news is a performance, a spectacle of announcing the events of the day, heightened by live reports from studio and locations around the world, performed for the audience that has gathered. This audience has gathered because it has little choice; if they turn on the television half an hour later, the news is over.

It is this lack of *permanence* that is the source of television's peculiar rhetoric of live interviews. After the interview is finished, it is to late. The live interview is a way to make up for the lack of permanence by reducing latency. But this is a trick, a sliding in story time, what Genette (in *Narrative Discourse*) would call an *anachrony*. What actually (in (hi)story) happened earlier is turned into an event that happens now (in discourse); or rather, an event is staged, which sole content is an earlier event. If the event reported is story, the discourse is the report of the event. But by stressing that the interview is live, this discourse is turned into the story of a new discourse (the newscast). Genette has no name for this anachrony, we might call it *currentisation* for its effect or *staged narration* for its way of doubling discourses.

On the Web the recording of an interview may be stored and offered as video-on-demand to readers. It may become a (relatively) permanent offer. To stage a live conversation at a point in time with only a random relation to the time of the actual event makes no sense in a technology allowing for permanence.

What technologies allow for permanence, then? Print is in principle permanent, at lest until the paper withers away. In practise, however, only libraries are able to keep complete collections of periodic publications. It seems that in earlier media, permanence goes hand in

hand with latency, both being results of the distribution technology. The shorter the latency from sender to receiver is, the shorter the permanence. A live broadcast is gone when the broadcast is over, while books are kept in the bookshelves. For print products, the user decides whether an item should be saved or thrown away; but few people collect magazines, even fewer save newspapers longer than about a week, while bound books rarely are thrown away. For broadcasting, saving material not only depends on active action by the viewer or listener, he or she will also have to have extra recording equipment.

In Web sites, permanence is just as much a choice of the author. A Web server is a public repository of files, and any new file may be archived forever if the author wants it. There are no longer connections between latency and permanence, as a live video feed may be available from an archive immediately after the live event has ended. Most periodic Web sites make use of this possibility and keep extensive archives of older material.

When whether to keep an archive becomes a rhetorical choice, a rhetoric of the archive becomes a possibility, although not all sites take the opportunity. In newspapers, few news reports stand very well by themselves, as most news stories are follow-ups of earlier stories, bringing new developments, comments or reactions to issues already reported on. When an article is pulled from an archive, this context is often lost, and it requires more thinking and guessing from the reader to understand what is reported. The requirement of background knowledge is well known within journalism, however, so many stories will contain a summary of earlier events, typically at the end of the article. Thus, for readers of archived reports, the problem of lack of background may be reversed: it is common to find several articles with almost identical summaries of the background. What is regarded as significant also changes over time. Not everything deemed newsworthy in the day-to-day running of a news desk stands the test of time, and in hindsight, it is easy to see that some

developments are more important than others. Writing history means to summarise, and in an archive, the lines of developments may be lost in the details. (Not to mention the many "related articles" that are found in some automated archives that turn out not to be related at all.)

In contrast stand those periodic Web sites that keep background material, overview articles, biographies and timelines in one place, and consistently link to this material whenever these issues are reported on. *MSNBC*'s "interactive" on Palestine, titled "Searching for Peace" is an example of this.• From a central timeline, a large number of maps and overview articles about Israel and Palestine is available, as well as biographies of key players in international Middle East politics. Another example that stands out (despite being old by Web standards) is *Nationalgeographic.com*'s 1996 feature "Gaza," where an article about Gaza is linked to parts of earlier articles from several years of *National Geographic Magazine* with great effect. What makes both these sites different from mere archives of earlier articles is that they are edited as overview sites, and care is taken to provide overview and understanding and avoiding redundancy or excessive detail.

Another, simpler approach may be found in most "blogs" or "Web logs," personal sites where new material is added frequently, sometimes even several times a day. Blogs became popular around 2001 due to inexpensive publication tools such as Blogger, which made it possible for anyone with a Internet access to become writer and publisher of her or his own publication. Blog writers frequently link to earlier articles (or "posts" as they are commonly known), not bothering to write anything twice.••

These few examples demonstrate the whole new range of rhetorical possibilities that opens with a technology that enables the author to choose how long a text should be available. Permanence, together with bandwidth and latency form the axis of distribution.

• "Searching for Peace" is also analysed in chapter 7, see that chapter for illustrations.

•• To be fair, it should be remembered that most blogs are written by one person. It is easier for the author to remember what has been written before than it would be for a site written by a group of fifty journalists all contributing daily.

1.2.1.4 Distribution and Rhetorical Convergence

A Web form is perceived to be the result of a rhetorical convergence if it blends rhetorical properties of two or more genres and media. The axis of distribution is one set of rhetorical properties: it describes how a text is transported from author to reader. It has two related dimensions which find a balance in each text: the kind and amount of material to be transported; and the time between writing and reading. (In the case of reports of real events a third dimension is added: the time between event and writing.) Above, we used the concepts of *bandwidth, latency,* and *permanence* to describe this balance. The axis of distribution is useful for describing certain kinds of rhetorical convergence, as may be expressed in this way:

> *A Web form with a balance of amount of material and time between authoring and reading similar to one genre in one medium, but in other respects similar to another genre in another medium is perceived as a result of rhetorical convergence between the two.*

We see this in many popular Web media: continually updated Web newspapers do not change all articles from one edition to the other, but add new stories whenever they are ready, in a way we recognise from news channels on TV or radio. When recorded newscasts are offered as "video-on-demand," they appear more like print newspapers, as they may be "read" at any time. Bandwidth concerns tend to find compromises between video and print practises, as when moving images are replaced by stills in a documentary, or when a small video pane is placed on a page surrounded by paragraphs of writing, as an illustrative image would be.

In 1.1.5 above when discussing Genette's concept of the *architext*, we argued that any text is understood or recognised via its affinities with texts we have encountered before. Rhetorical convergence is when we

recognise a single text as being similar to other typical texts of two or more different media. A Web newspaper is an example of rhetorical convergence, as both writing and lay-out is very similar to print, but its mode of distribution is broadcasting-like.

News in newspapers and television is compared to Web newspapers and video-on-demand in the table below. By registering aspects from the axes of distribution and signification, we see clearly how Web newspapers and video-on-demand may be seen as results of rhetorical convergences of print and television news.•

• I have entered typical values in the table, assuming that the newspaper would arrive at the newsstand eight hours after deadline, and that the Web sites would keep written articles for three years and video for three months.

Web sites are registered with a transmission latency close to nil, as they have the possibility to transmit live.

	Mode of Distribution		Mode of Signification
	Transmission latency	Permanence	Dominant sign systems
Print newspaper	8 hours	approx. 1 week	Writing, still images
Web newspaper	≈ 0	3 years	Writing, still images
Web video	≈ 0	3 months	Moving images, speech
Television	0	0	Moving images, speech

1.2.2.1 Mode of Restrictions: Canvas

The second of the four axes is the *mode of restrictions,* the limits of a text's "canvas." We will use the term *canvas* for the material that the signifiers are made of, or in the words of Eco's *Theory,* the continuum within which the signs are shaped (217). Let us initially use the term for the paper in print, the silver screen in cinema, the loudspeakers of a radio, or the screens of television sets and computers. When these membranes are written upon or vibrated to produce sound, they have different

physical qualities. Wide-screen cinema has far better image resolution and colour range than television. This has implications for the rhetoric; cinematographers may photograph subtle images that would be wasted on TV, as the subtlety would be invisible.

The properties of the canvas are based in technical limitations and possibilities. Why introduce a special term, *canvas*, for what seems to be merely technology? The answer is that the canvas is not just the technology in use. In using a painter's canvas as a metaphor, I want to direct attention both to the physical properties of a technology, as in the way painting on canvas is different from painting on paper or a dry wall, and to the choices the author makes, as when a painter cuts the canvas to the desired proportions before painting, thus setting the maximum size of the image.

These choices of the author are restrictions. Like a painter's canvas is a reduction from the original piece of canvas, any text is realised within a subset of the possibilities the technology offers. Each medium has some fixed standards, some flexible ones. The frame in cinema is normally limited to a few standard ratios of height and width (American, wide-screen, et cetera), while the length of a feature is much less standardised; most feature films are roughly between 90 and 150 minutes long, but there are many exceptions. Television screen ratios are even more fixed than cinema's, while books can be of virtually any size a person can handle without special tools.

Thus, the author sets many restrictions for the text. It appears to me that we form an overview of the many properties of the physical appearance of a text by seeing it as existing within a limited range and having a limited number of details. The limits of *range* and *detail* operate both in space and time. *Range* is the distance between the outer limits of a continuum for sign-production. A visible text is restricted within a vertical and a horizontal range. In addition, it has a range of contrast, the span between the colours that contrast the most. Sound has little spatial existence, but is nonetheless restricted at any point in time by a frequency range (the span

between the highest and lowest frequencies possible) and a dynamic range (the span between the softest and the loudest sounds).

The range used in a text is often less than the largest range technology allows. In most of the Web sites studied in this thesis, the area used in any page is significantly smaller than my computer screen. This is normal practice, in order not to annoy those readers who have small screens. Just as Web designers need to design for different bandwidths, they also need to allow for the different screens used by their target audience. Cameramen in television have also followed this principle throughout the tradition of their industry; most home television receivers have much poorer image quality than the monitors used in television production studios. A skilled television producer knows that the images must be bold enough for their messages to come through even on a small old receiver.

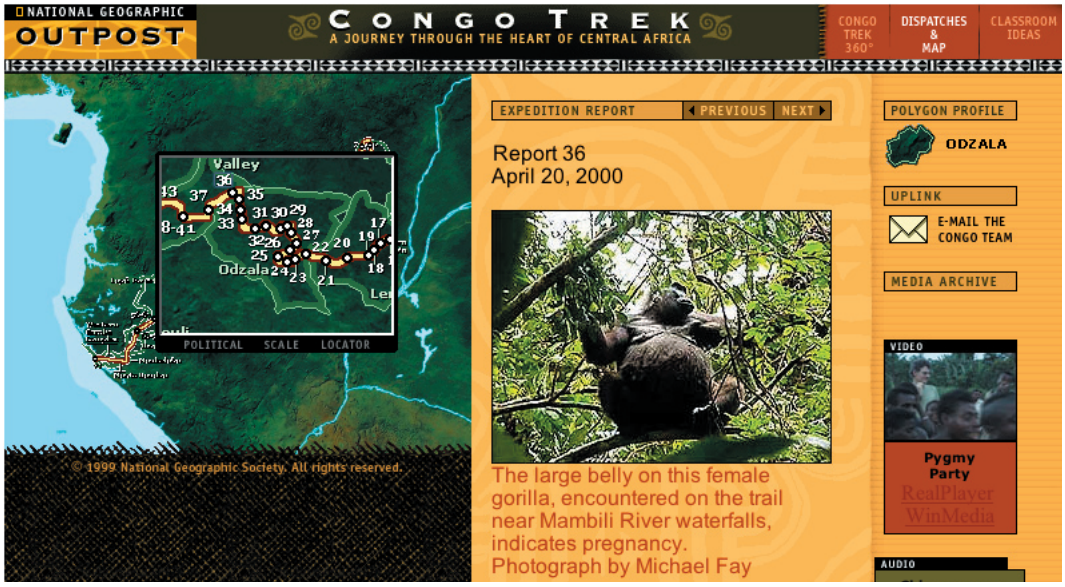
Within the range(s), there is a lowest level of *detail*, that is, there is a lower limit to how small a part can be, which also sets an upper limit to the number of physical parts. In computer images, these are called *resolution* (the largest possible number of details within a range) and *colour depth* (the number of possible colours). Also in painting – which does not have any graded differences in the same manner as computer graphics – there is a limit to how fine a line is, set by the paint, the surface, the brushes and the painter's technique. These limits are also rhetorical choices: if the painting is a roof decoration, for example, the smaller details are not called for, as all viewers will stand several meters away from the artwork.

Computer screens have very low resolution compared to cinema and print,• and this puts a limit on the use of detailed graphics. Much of what is called "interactivity" is rhetorics devised to reduce or bypass this disadvantage. One example is the "Congo Trek" site made by *Nationalgeographic.com*. The site is an archive of more than seventy-five letters from conservationist Michael Fay, who spent fifteen months walking across the rainforest of central Africa. A map of the route Fay

• At the same time, the range of contrast in images is much better on computer screens than in print, it is almost as good as in positive film projection (slides or cinema).

•• "Congo Trek" and the concept of *canvas* in relation to the site is also discussed in chapter 6, "Interactive Graphics."

followed contains links to all the seventy-five letters, but the details needed of the map are not visible on a computer monitor. To remedy this problem, the authors programmed a "magnifying glass" that the reader can direct over the map to reveal the details.••



http://www.nationalgeographic.com/congotrek

Range and detail also have a temporal element. A dynamic text has a range (time span), and a number of possible changes within this range (the frame rate in film, sample rate in recording, screen refresh rate for computer monitors, polygons per second in graphic computer games).

The temporal range is often a clear indicator of genre. A feature film or novel is supposed to be of a certain length, while an American television sitcom like *Friends* is adapted to a half-hour playing time, made to be interrupted by commercial messages. In computer media, temporal detail – frame rate, for example – is often a trade-off between technical quality and bandwidth concerns, a part of what we called *mode of distribution* above. A high frame rate is necessary to show fluent motion, but it also drives up the file size. It

is thus a necessary decision for a Web author, that will have rhetorical consequences.

The table below places some technical terms that may describe a text's mode of restrictions in relation to range and detail in space and time.

	Range	Detail
Spatial	Size	Resolution
	Aspect ratio	Colour Depth
	Contrast range	Sample depth
Temporal	Duration	Frame rate
		Polygons/sec
		Refresh rate
		Sample rate

1.2.2.2 Canvas and Rhetorical Convergence

A Web form is perceived to be the result of a rhetorical convergence if it blends rhetorical properties of two or more genres and media. The axis of distribution is one set of rhetorical properties. It describes the canvas: the range and detail in time and space available for sign production in a text. Rhetorical convergence may be perceived by a shared canvas:

> *A Web form with a canvas similar to one genre in one medium, but in other respects similar to a genre in another medium is perceived as a result of rhetorical convergence between the two.*

A unique feature of computer texts is that they are flexible, a fact that has been deliberated at some length by, for example, J. David Bolter in *Writing Space* and Nicholas Negroponte in *Being Digital.* A consequence of this is that a text may have a flexible canvas, it may vary between different modes of restrictions. Most file formats have not set all the restrictions we have discussed, but let the author specify them for each new text. While, for example, the frame rate of television is a given, it may

be different in two subsequent computer videos, or even change during the course of a film. Some restrictions may be left for the reader to decide; so-called "fluid" designs will fit the computer window whatever width or height the user has set it to. The possibilities of combining modes of restrictions from different media's rhetorics thus seem virtually unlimited. A few examples are given here only to demonstrate rhetorical convergence along the axis of restrictions.

In 2002, the visual quality of computer displays is so poor compared to paper that there are very few visual forms from these earlier media that can be reused on the Web without changes. The only typography from paper that can be recreated on the Web is that of small formats: newspaper layout would be impractical on the Web, as it would require a lot of scrolling both horizontally and vertically.• Almost any Web form derived from print would be perceived as converging with computer rhetoric (or computer interface conventions, which is an alternative term) because of the difference in canvas. The "magnifying glass" from "Congo Trek" discussed above is clearly related to the "locator" interface in the popular image editor application Adobe Photoshop or the electronic document viewer Acrobat Reader.

• The consequences of this for the "usability" of Web sites are discussed by Nielsen in "Print Design vs. Web Design." Nielsen also uses the metaphor of the *canvas*, but in a more restricted sense than it is used here.

The table below renders aspects of three of the four axes when comparing the map in the "Congo Trek" Web site with the similar maps in the *National Geographic Magazine* and the television program *Africa Extreme.* Although the Web site has a canvas with even less detail than a television screen, it uses a map almost as detailed as the one in print. To make the map readable, a mode of acquisition ("display control," that is, the magnifying glass) from computer media is used, exemplified in the table with Acrobat Reader, an application used to view on a computer documents designed for print that often are too large for the user's screen. (In fact, a fourth map is available in the "National Geographic Magazine" Web site, to be viewed with Acrobat Reader.) "Congo Trek" is thus a convergence of three different forms.

	Canvas	Mode of Signification	Mode of Acquisition
Television	720x576 pixels	Small maps, few details, animation	Watching/listening
Web site	351x280 pixels	Large map, many details; writing	Reading + Display control
Magazine	12000x16200 dots•	Large map, many details; writing	Reading
Acrobat	1280x1024 pixels	Large map, many details, writing	Reading + Display control

The flexibility of computer technology opens for a multitude of positions between established media's modes of restrictions. Web camera pages are good examples, as they are midway between films and still images, technologically speaking. A Web camera is in fact a video camera, but it is connected to a computer that is programmed to copy the camera's image at a much lower temporal detail (frequency or frame rate) than video has, so it will not be perceived as moving. Still, it is a dynamic image, it does change over time. All earlier media are either dynamic to the degree that they have enough temporal detail to make us perceive continuous movement or sound, or else they are still. The computer has opened up a whole range of rhetorical possibilities between these two perceptual poles.

• Resolution for the printed map is calculated as if printed on a 1200dpi printer. Actual resolution is probably higher. For Adobe Acrobat, it is given as the full viewable area of what in 2003 is a fairly large monitor.

1.2.3.1 Mode of Acquisition

Among the established media, there are marked differences in how they are read. Cinema is viewed collectively in one sitting from beginning to end, in a darkened, public theatre. A newspaper is read privately, can be taken anywhere, put down and taken up again, flipped through and read in any sequence. Few readers read the whole paper, instead most readers rely on headlines, leads, images and captions to decide what to read. We will discuss such differences here under the heading *mode of acquisition*.

Mode of acquisition is how the reader accesses the signs of a text. Again, this notion would often seem superfluous in earlier media. In computer media, on the other hand, the kind of reading process the text enables and encourages is a rhetorical choice, and encompasses

the different devices often contained under the label *interactivity*.

While most films and novels are told in one sequence which the reader is expected to follow, Web sites rarely are. Most Web sites offer a partial list of their contents as a menu of choices for the reader. The reader chooses what links to follow, and only when a link is activated (typically clicked on) by a user does new page appear.

The reader's choose-and-click reading activity is named *interaction* by some writers, while others reserve this term for other kinds of reading and writing activities, with computers or between human beings only.• A less disputed term for the reading activity involved in Web reading would be *ergodic*, coined by Espen Aarseth in *Cybertext*. Aarseth defines *ergodic literature* as literature in which "nontrivial effort is required to allow the reader to traverse the text" (1), and having made this effort, the reader "will have effectuated a semiotic sequence" (ibid.). Reading traditional novels, watching films or television of course requires cognitive effort, but apart from interpretation•• the effort is considered trivial, as the reader only has to turn pages, the viewer to watch. In hypertext or computer games, however, the reader/player constantly has to make decisions as to what to do with the text.•••

These "non-trivial efforts" are manipulations of the text as a material and mechanic object. Manipulations may be throwing dice or coins, ordering pieces of paper, or controlling a computer interface with mouse, keyboard or other input devices. To focus in this way on the material structure of a text and the ergodic effort involved in reading is what Aarseth terms a *cybertextual perspective* (22). To view a text as a cybertext is to view it as textual machine "for the production of a variety of expression" (3), a machine that according to some principle combines pieces of text into the text the reader reads. The machine thus has three important sets of parts: the *textons*, the pieces of text that may be combined; the *scriptons*, which are the pieces of text the reader is expected to read; and the *traversal function*, "the mechanism by which scriptons are revealed or

• See Jens F. Jensen's "The Concept of 'Interactivity'" for a collection and comparison of many of the contradictory uses of the terms *interaction* and *interactivity*.

•• Aarseth uses the Husserlian formulation "in an extranoematic sense" (1).

••• As Aarseth shows, ergodic texts are not limited to computers. He lists a large number of examples, from Egyptian hieroglyphic murals and the Chinese *I Ching* to twentieth-century experiments in books that may be read in many different ways. In the same way, computer texts are not automatically ergodic texts or cybertexts (ergodic texts that involve calculation, so they will change each time they are read). A computer text may be perfectly linear and non-changing, even more so than a book: a book can always be opened on any page, while a computer system may block access a page until the reader has read the preceding page, Aarseth argues.

generated from textons and presented to the user of the text" (62). Examples of traversal functions are the links in a hypertext, the simulation and representation engines in an adventure game, or throwing coins in the case of *I Ching*.

We can in principle treat any text from this cybertextual perspective. In a traditional printed novel the textons and the scriptons would then be identical, and the traversal function nothing. By stating that, we have said very little about the novel, however, the cybertextual perspective is only interesting when studying ergodic texts. In a static Web site, the pages stored at the server are the textons, the links between them the traversal function, and the pages displayed to the reader when links are followed are the scriptons. When developing his typology of ergodic texts, however, Aarseth also uses *cybertext* as the name of a genre. Understood in this way, a cybertext (perhaps we might call it a cybertext proper) is a text where the traversal function involves some principle of calculation (75), so the text will not look the same in two readings. In this typology the static Web site would not be a cybertext proper, while a computer game would.

Viewed from the cybertextual perspective, most Web media are constructed differently from earlier media. The Web technology was initially designed to support distributed hypertext, and the basic structure of saving Web pages as separate files makes it a manifestation of the cybertext model. Links, and more advanced techniques such as javascripts and server-side scripting, allow for the construction of a wide range of traversal functions, opening for an even wider range of modes of acquisition. It is rare to see forms from other media being adapted to Web media without adding cybertextual features such as links and search routines.

What I call *mode of acquisition* should then be understood as the reading process that results from the mechanical (cybertextual) construction. Modes of acquisition in computer media have been classified by different theories.

Based on a text's mechanical (or cybertextual) construction, Aarseth is able to classify it as inviting

the user to perform one of four *user functions*. The *interpretative* function is to read and comprehend; the *explorative* also requires the user to decide where to read next. These two user-functions are required by *unicursal* and *multicursal* texts respectively.•

• The terms *unicursal* and *multicursal*, taken from a study of labyrinths, are Aarseth's alternative to terms such as *sequential* and *non-sequential; linear* and *non-linear* or *multilinear* (44).

Dynamic texts, that is, texts where the number of scriptons is not constant, open for two other user functions. The *configurative* function is to let the reader "configure their scriptons by rearranging textons or changing variables" (64), while allowing the reader to add textons is a *textonic* user function.

Mode of acquisition is also related to the sign system(s) used. Sound and moving images exist over time, and have to limit the reader at least for a moment for at all to exist as texts, while all writing in principle is open to be read in any sequence. Still, a looping video of wind in trees, or a long sound recording of waves may not be very different from a still image in the way it is read.

Following Kant's distinction of objective and subjective sequences, Gunnar Liestøl has discerned between different kinds of activity in reading texts. "With the consumption of dynamic information (audio/video) the dominant activity is located in the textual object itself, as object-action; with the consumption of static information, however, the dominant activity is located with the user-subject, as subject-action" (45). Of *subject-action,* there are two kinds, *intervening* and *non-intervening. Non-intervening subject-action* is the activity of reading a static text, while *intervening subject-action* is to actively choose where to read on. Reading hypermedia is to alternate between parts of the text dominated by object-action and the different kinds of subject-action, according to Liestøl.

Yet another way of discerning between different modes of acquisition is proposed by Jens F. Jensen in "The Concept of 'Interactivity'." Jensen proposes a typology of twelve different media positions, depending on whether the media allow for *registrational* interactivity and *conversational* interactivity, and which of three kinds of *selective* interactivity is offered (*none, transmissional*

or *consultational*). Based on Bordewijk and van Kaam's typology of "traffic patterns," Jensen's typology orders media after the power relations they set up between providers and consumers. *Registrational interactivity* is "a measure of a medium's potential ability to register information from and thereby also adapt and or respond to a given user's needs and actions [...]" (60), or in other terms of the flow of information from consumer to provider. *Conversational interactivity* is information exchange between consumers "in a two-way media system" (60). *Selective interactivity* is the possibility for consumers to select between different available programs or texts made by providers. If the providers control the distribution and consumers only select what to read, it is of the transmissional kind, if consumers initiate distribution, it is of the consultational kind.

Aarseth, Liestøl and Jensen all provide perspectives from which the modes of acquisition may be described and assessed; perspectives of material construction, of sign system, and of power relations respectively; perspectives that are not necessarily contradictory, although they divide the range of texts differently.

In the essays "Linearity and Multicursality" and "Semiotics and the World Wide Web,"• I have argued that the selection of links by readers is governed very much by how the links are signified. I am thus able to see a range of different kinds of what we may call "explorative user functions," "intervening subject-actions," or "consultational selective interactivity" in the different vocabularies reproduced above.

• Chapters 5 and 4 in this thesis.

Imagine a continuum of how much influence and control a reader has of the sequence of parts in a Web site. At one end we find movies: videos, animations and narrated slide shows (what I in this thesis call slide-motion films). The most television– or film-like of these run in one sequence by themselves; I have called this *cinematic mode*. Most movie streams, however, have a pause button and perhaps VCR-style controls for fast forward and "rewind." A similar control is handed over to the user when Web pages are linked in a chain with a "next" link on every page. Similar to reading a novel,

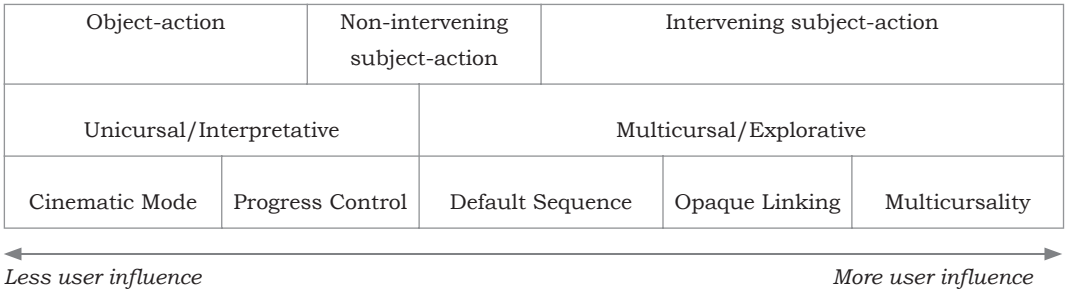
there is only one next page, but the reader chooses when to go to it. Although these two modes have different dominant actions (static Web pages with "next" links are dominated by subject-action, while VCR-controls introduce subject-action to texts dominated by object-action), I have chosen to group them under the term *progress control.* All these texts would be what we call linear, sequential, or unicursal.

At the other side of the continuum's middle are the multicursal sites; sites where most pages have more than one link. In several of them, there are links, but there is either so little information of what will be at the other end of each link, or the text of each page depends so much on what was written on a previous one that one sequence is strongly prioritised before all others.• I have called this *default sequence,* using a term inherited from Jane Yellowlees Douglas' and Jill Walker's discussions of the novel *Afternoon* by Michael Joyce. Further along the continuum, we find texts where there are more possible or probable sequences, but the reader still has limited control. These are sites where links load random pages, and sites without a prioritised sequence, but also so little information about the links' destination that the reader is navigating blindfolded through a labyrinth. In my paper "Linearity and Multicursality," I called this *opaque linking.* An example of this kind of text would be Michael Joyce's canonical hypertext novel *Afternoon.* At the end of the continuum furthest removed from films are Web sites where not only there are many links, but where the structure of the linking is made so explicit that the reader can navigate freely, as in a building she knows well. I will argue it is only this mode of acquisition deserves the name *multicursal* in a strict sense. While other forms of linking may open the possibility of many courses through the text, only an explicit linking practice makes it likely that the reader experiences the text as making multiple courses possible.

In the figure below, my five modes of acquisition are related to Espen Aarseth's user functions and textual positions, and to Gunnar Liestøl's kinds of activity.

• Examples of this are found in "Linearity and Multicursality" (chapter 5).

As these five modes only describe texts with static textons, the figure only maps what Jens F. Jensen calls the dimension of "selective interactivity," and is not able to grasp the "conversational" and "registrational" dimensions, or Aarseth's "configurative" or "textonic" user functions.



This way of mapping modes of acquisition implies that hypertext or computer rhetoric has two levels: a cybertextual level and a rhetorical level. *Default sequence, opaque linking* and *multicursality* are all results of similar cybertextual constructions, it is the writing on the pages and links that separate them. Clearly, link-node hypertext is a certain figure of cybertextual construction that may give rise to many different rhetorical figures.

Espen Aarseth touched upon this separation in his essay "Nonlinearity and Literary Theory," and later in *Cybertext*. Following Pierre Fontainier's nineteenth-century rhetoric, he separates between tropes and "le figures non-tropes," or syntactical and semantical figures, as Aarseth terms the two kinds (*Cybertext* 91). In "Nonlinearity" Aarseth lists some "figures of nonlinearity" as syntactical figures, or figures; static Web sites and other link-node hypertexts typically make use of the figure linking/jumping. In *Cybertext* (91), he addresses the other side of the pair, the "tropes" or "semantic figures," in the analysis of Michael Joyce's *Afternoon*.

As this particular use of the traditional concept pair *trope* and *figure* for semantic and syntactic figures

respectively is far from universal in the rhetorical tradition, I will prefer the more common terms *figure of thought* and *figure of diction,* leaving *trope* for figures that transform the meaning of words or expressions such as metaphors and metonymies (compare Plett 309). But Aarseth's point, that the various cybertextual mechanisms for textual production are different figures of diction rather than figures of thought is an important one.

Many cybertextual figures are neutral techniques that can be used with different effects in different texts. Static Web sites are examples of "link-node hypertext," a cybertextual figure where stable scriptons are bound together with stable links, so the scriptons in the text always have the same relation to each other. Aarseth views "link-node hypertext" as a topological structure. I will also call this structure a figure of diction, a way of constructing a text. As a structure of scriptons and their relations, which may be filled with any semantic content or mere gibberish, the structure will remain the same.

Aarseth's two hypertextual figures of thought, *aporia* and *epiphany* describe the users frustration when lost in the textual labyrinth of Michael Joyce's *Afternoon*, and then the bliss when a new direction of reading, is found. These two figures of thought are put on top of a certain structure, a use of the link-node figure and a figure of restricted access, both figures of diction. The very same structure, with the same conditional links could have been made explicit, with every link and block explained to the reader. In that case, the figures of thought would have been different.

As with all other rhetorical figures, the hypertextual figures we have discussed here, both figures of diction and figures of thought, may be combined with most other rhetorical or poetic devices. A worked through system of navigation links, for example, is usually associated with professional business sites, while hypertext novels of the "Eastgate school" usually employ relation links, but this need not always be the case. Bobby Rabyd's Web novel *Sunshine '69,* is an example of a serious hypertext fiction using a clear and understandable set of navigation

links, resulting in a truly multicursal work of fiction. In *Hamlet on the Holodeck,* Janet Murray lists numerous experiments in immersive computer fiction, among them multiform stories, stories that present "a single situation or plotline in multiple versions, versions that would be mutually exclusive in our ordinary experience" (30). When giving multiform stories as a writing assignment to her students, many respond with a "violence hub" story, Murray reports (135). These are stories where something violent or traumatic has taken place, and the reader is invited to follow links to explore how a number of characters react to the event. In these texts, many rhetorical figures are at work simultaneously. The "violence hub" is a certain realisation of the more general multiform story. None of these need to be multicursal works; in fact, most of Murray's examples of multiform stories are Hollywood movies. But some are multicursal computer text, and these need to communicate this structure to the reader, using certain linking figures. All these devices, multiform, violence hub, and linking figures, are figures of thought, used within a certain figure of diction, the link-node structure. Similarly, the decision to posit the reader as a character in the diegesis as in many computer games, is a figure of thought not concerning the cybertextual construction of a text, a figure that also has been used in codex novels, such as Italo Calvino's *If on a Winter Night a Traveler.*

Mode of acquisition should thus be understood as the reading effort and experience the text invites and expects from its readers. It consists of both the required handling of the text's mechanical structure, how this requirement is communicated to the reader, and how it is aligned with the text's message, be it narrative, argumentative, descriptive, or explanatory.

1.2.3.2 Mode of Acquisition and Rhetorical Convergence

As mentioned, a Web form is perceived to be the result of a rhetorical convergence if it blends rhetorical properties of two or more genres and media. The mode of acquisition

is one set of rhetorical properties: it describes the reading process required of the reader to read a text.

> *A Web form with a required reading process similar to one genre in one medium, but in other respects similar to a genre in another medium is perceived as a result of rhetorical convergence of the two.*

The mode of acquisition is an integral part of all earlier media, something which is witnessed by the large literature on the difference between the "linearity" of print and the "nonlinearity" of hypertext. As with mode of distribution and mode of restriction, a change in the mode of acquisition will be perceived as changing the rhetoric towards the rhetoric of a different medium. It further seems rare to recreate any rhetoric on the Web without introducing some linking beyond the imitation of page turning or the controls of a VCR. Introducing a cybertextual figure is a convergence towards computer rhetoric, but as we have seen, this may result in a wide range of different rhetorical forms. A simple example may show how the convergence of video and ergodic texts may take place:

Many television networks offer video recordings of their news on the Web. Many such newscasts are chaptered, so the individual reports can be selected from a list, and viewed out of sequence.• Viewing television in this form becomes more like reading a newspaper. The reader may select what to view when, and is not likely to have the patience or interest in following the original sequence of the newscast, a sequence that probably was made with care. Following her own priorities, the reader will continually consider whether a story is worth her attention, and be ready to break it off at any point by selecting another item from the menu. And it is likely that she will skip some items entirely, as they do not interest her. Changing the mode of acquisition in this way is a profound change from the kind of ceremony the evening news is on broadcast television. The next table renders the rhetorical convergence along two axes:

• In "Combining Video and Writing" (chapter 9), the example of the Norwegian video-on-demand site *Frihet* is analysed. On *Frihet*, the newscasts from the television channel TV2 is available on demand as chaptered streaming video.

	Mode of Acquisition	Dominant sign systems
Print newspaper	Multicursal/progress control	Writing, still images
Chaptered video	Multicursal/progress control	Moving images, speech
Television news	Cinematic	Moving images, speech

1.2.4.1 Mode of Signification: Sign systems

What I here will call *mode of signification* is not difficult to grasp, it is simply the sign systems used in a text; for example, video, writing, spoken language, or still images. The similarities and differences between the sign systems are complex, however. As with the other three axes of rhetorical convergence, the mode of signification encompasses many modalities a text can occupy. And as with some of the other axes, the modalities are intertwined.

Any printed matter, sound recording, video, film or broadcast may be reproduced, or re-represented, or copied in the computer. It will be stored as one of four classes of file formats: text (mainly ASCII, but other formats exist), images (formats such as JPEG, GIF, TIFF, PICT, BMP, PNG), sound (formats such as AIFF, WAV, MPEG, AAD), and video (formats such as AVI, QuickTime, Real, Windows Media, MPEG). To the human ear and eye, however, the distinctions blur. A photograph of a poster would be stored in the computer as an image, but we read it as writing nevertheless. Similarly, we may film a still image, or draw an image with ASCII signs. Audio may contain recognisable sounds of all kinds, including music and spoken language. To form an understanding of multimedia as communication between humans, we must instead distinguish between the different kinds of signification these formats store.

There are many ways of relating the four classes of sign systems listed above. We will in the following list several distinctions between characteristics of sign systems, in order to show similarities and differences between different modes of signification.

Philosophers and scholars of rhetoric, poetics, aesthetics, linguistics, and semiotics have discussed at length the differences between different communication

systems such as language and image, spoken and written language, the eye and the ear, the spatial and the temporal. These dichotomies are all intertwined, as we may illustrate by putting the four sign systems into a table:

Writing	Speech
Pictures	Video

Of the sign systems in this table, only speech is perceived with the ear, the other three by the eye. The top two, writing and speech are based on natural language, the bottom two are kinds of imagery. The right half, speech and video exist in time, they are dynamic or temporal, while the left half are fixed, spatial. One way of labelling the rows and columns is thus:

	Fixed/Spatial:	Dynamic/Temporal
Language:	Writing	Speech
Image:	Pictures	Video

Many of the differences between popular genres in different media can be sorted along these two dichotomies. Alphabetic writing is abstract, for example, and the length of lines and pages do not matter in many writing styles, while alterations of proportions or size in images do change the impact of the image. When writing and images are combined, writing will have to give up some flexibility to fit with the pictures. Adding moving images to explanatory graphics is another example that makes the difference stand out. Using animation, processes and time relations may be rendered more effectively, but at the same time the ability a reader has to scan and compare the different parts of a still image is taken away.•

Within each row in the table, the differences are also much discussed in literature. The difference between spoken and written language has been a topic

• Divisions of static writing and moving images are discussed further in chapter 2. Movement in explanatory graphics is treated at length in chapter 6, while chapter 7 discusses effects of moving the frame over still images.

for language philosophy since Socrates, renewed by present-day linguistics. The film theories of, for example, Jean Mitry or Gilles Deleuze are concerned with the specificities of the moving image as opposed to the still.

What is lost in our labelling of the categories in this manner is the distinction between eye and ear mentioned above. This is an important distinction, however, when discussing the combination of forms. It is often easier to comprehend combinations of sign systems involving two different senses, as we will return to in 1.2.4.2 below.

Sound is always temporal, always dynamic.• But all sound is not speech. Music and sound effects are obvious examples. To fit this distinction in, we might have to reduce the language/image dichotomy to language and non-linguistic signs. Thus:

• At least, sound is always temporal. It might be said that some sounds are static, for instant a ship's engine noise. A mechanic would take the sound as a static sign that the engine is working steadily. A change in the sound would be a sign that the speed is changed, for example, and/or that something is wrong.

	Fixed/Spatial:	Dynamic/Temporal	
	Eye:		Ear:
Linguistic:	Writing	Animated writing	Speech
Non-linguistic:	Pictures	Video	Music, Sound effects

This division also makes us aware of animated writing as a distinct form. It occupies a middle position between writing on one hand and video and sound on the other, and Gunnar Liestøl has demonstrated in "Aesthetic and Rhetorical Aspects" how animated writing may smoothen the transition from writing to video (a point we also touched upon while discussing bandwidth above). Moving writing looses one of the powerful aspects of alphabetic writing however: the ability to read at different speeds. Reading more than a few words of moving writing is thus likely to annoy many readers.

One might very well suspect that the "non-linguistic" category is too broad, and this is brought to the surface if we consider Web pages combining photographs and diagrams, two very different forms of images.

Four Axes of Rhetorical Convergence

In the Spanish newspaper *El Mundo*'s "Oleada de atentados en Estados Unidos," an "interactive graphic" of the September 11th disaster in 2001, schematic drawings and photographs of the World Trade Center are combined with great effect.

The drawing explains what is happening (a logos appeal), while the photograph is a witness, and much more emotional (a pathos appeal).• In *Languages of Art,* Nelson Goodman clarifies the distinction between the two, as well as the distinction between language and image. In his vocabulary, languages are *differentiated*, while images are *dense*.•• Like in Saussure's semiology, words are seen as disjunct, differentiated signs with differentiated meanings in Goodman's theory.••• Dense sign systems (or dense *notational schemes* in Goodman's vocabulary) do not have differentiated positions. Thus, between two signs, there is a possible infinite number of signs. His example is a mercury thermometer without a grid marking the temperature scale. In such a thermometer, any position of the mercury scale would be meaningful. Dense sign systems may further be either relatively *attenuated* or relatively *replete*. In a photograph or a painting any aspect of the image is

• This Web site is also discussed in chapter 7, using many illustrations from the site.

•• Goodman also uses the synonyms *digital and analog*, which are practical, but dismissed in this dissertation to avoid confusion with other uses of the word *digital.*

••• Goodman distinguishes further between disjunct and differentiated both on the syntactic and the semantic plane, in order to define proper *notational schemes* (such as music notation) but this is beyond the needs for the present discussion.

http://www.elmundo.es/elmundo/2001/graficos/septiembre/semana2/atentados/atentados2.html

potentially meaningful, so the image would be another if any detail was changed. In a map, on the other hand, choices of colour or thickness of line are relatively arbitrary. On a world map, it has little importance if Zimbabwe is coloured green or pink as long as its colour is different from the colours of neighbouring Zambia or South Africa. Maps and diagrams are relatively more attenuated than photographs, as some dimensions of the visuals carry meaning while others do not. Below, Goodman's distinctions are drawn into our diagram.

		Fixed/Spatial:	Dynamic/Temporal	
		Eye:		Ear:
Differentiated (digital):		Writing	Animated writing	Speech
Dense (analog):	Attenuated:	Diagrams, Typography	Moving diagrams	Music
	Replete:	Pictures	Video	Sound effects

This further subdivision not only helps us discern diagrams from other images, it also tempts us to place of other sign systems used in Web sites, such as typography and lay-out. Music may also be distinguished from sound effects in this manner.

When I claim that photographic images of the World Trade Center catastrophe are "witnesses," it is based in the knowledge that photography is a process of chemistry and optics. In Peirce's terms, the photograph is both *iconic*, as it resembles its motive, and *indexical,* as it is a physical imprint caused by another physical imprint of the light rays that were reflected off the motive. Now that we have introduced Charles Sanders Peirce's canonical trichotomy of signs: *iconic sign, index* and *symbol,* we see that in the above table, all the differentiated sign systems are symbolic, while all the replete are iconic. The attenuated, however will have aspects of both, while there is no separate place in the table for the indexical. Furthermore, a shot in a film of a landscape with a column of smoke rising from a distant hill would be iconic first, in that the image resembles an actual scene, then indexical second, as the smoke is a sign that there is a campfire burning. Our table

cannot capture Peirce's typology of signification while maintaining the differences we have charted so far. To map all dimensions of signification in one diagram would be overly complex.

Only a little reflection on how music communicates will complicate this yet further. Music may be seen as a system in which some parts (rhythm, scales, harmonies) are parts of a differentiated system, while others (timbre, volume, pulse, phrasing) are dense. In addition, music always carries strong connotational meanings. A simple example is found in the *Becoming Human* Web site (which will be described in section 1.3 below), where (supposedly) Ethiopian music lends connotations of "africanness" to a description of an excavation in Ethiopia. These secondary meanings apply to any interpretant to any of the sign systems involved, and cannot be captured by the above table either.

1.2.4.2 Mode of Signification and Rhetorical Convergence

We have repeatedly stated that a Web form is perceived to be the result of a rhetorical convergence if it blends rhetorical properties of two or more genres and media. The mode of signification is one set of rhetorical properties: it describes the particular combination of sign systems used in a text.

> *A Web form with a combination of sign systems similar to one genre in one medium, but in other respects similar to a genre in another medium is perceived as a result of rhetorical convergence of the two.*

Changing the sign system while keeping the rest of the rhetoric is the most obvious rhetorical convergence. All the examples used to illustrate the other three axes used the mode of signification as one dimension in the comparison. Many of them also in addition contain in the combinations of sign systems. The slide-motion film "Sights and Sounds of the Way west" described under 1.2.1.1 above combines a dynamic sound track with

still writing and imagery, but both photographs and written words are made dynamic by moving the frame and fading words in and out. *Yahoo! FinanceVision,* the example from 1.2.1.2, put paragraphs of written text next to a small video pane. The map from "Congo Trek" discussed under 1.2.2.1 uses the power of diagrams and maps to provide overview of a large number of written parts.

When two different modes of significations with different characteristics in all the ways listed above have to be aligned, two principles govern the combinations: the limits of the senses and perception, and what I call *containment.*

Our vision cannot read writing and images simultaneously (an observation elaborated by many scholars, for example by Michel Foucault in his analysis of Magritte's aesthetics in *This is Not a Pipe*), so we will have to move back and forth between the two. Thus, if a lengthy text is projected on top of a video segment for a short while only, it will be very difficult not to miss either some of the text or some of the images. Eye and ear may cooperate nicely, however, as when a voice-over explains images in a documentary film. It does also seem to me that our ability of language processing is such that we not only are unable to simultaneously comprehend two people speaking at the same time, but also that reading and comprehending several paragraphs of writing while listening to a speech is equally impossible.

Containment is a word I use to describe the fact that sign systems do not appear next to each other on the Web, but are convoluted. In Web pages, either a video pane is inserted into a text page, or text is inserted into a video window. Digital video is always rectangular, and text will normally either be within or around the rectangle. The fact that a video clip has to be a separate file from the HTML page makes it even harder to penetrate the edges of the video rectangle, if the author should desire to do so. Apart from these technical reasons, photographic video has always been separated by a frame, a basis for theories on film by Jean Mitry, Lev Manovich, and others. To escape the frame, or for text to penetrate

it, the video would have to loose all depth, all sense of foreground and background. It is imaginable to shoot a video against a monochrome background, matte it out, and script text to blend in with parts of it, but I have never seen it done in an actual Web page. What emerges is a master–servant or parent–child relationship.•

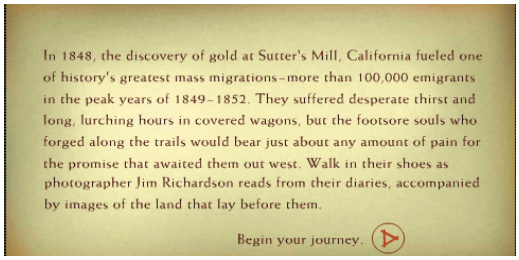
• Examples of different kinds of containment are given in the essay "Combining Video and Writing," the present chapter 8.

The parent-child-relationship between the containing semiotic system and the contained can also be treated as a time relation; one reads the parent before the child. This makes it possible to treat, for example, video inserted in a text page similar to video that opens in a separate window from a link in a text page. It is an advantage to do so, as the two often are used with comparable effect. To classify a page in one of the two categories, one would query which signs that reaches readers first; those of the text or those of video. Again, it is possible to imagine a page designed in a manner that makes this distinction impossible, that it is random what one reads first (which would require that the video loaded as quickly as the HTML). I have not seen it in reality, but were it to be found, it would be a third category requiring its own analysis.

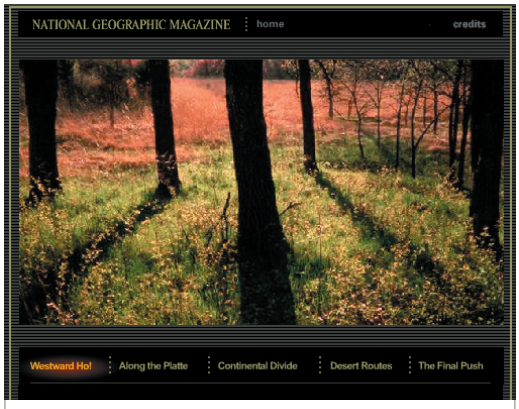
We will now go through the different distinctions between modes of signification. We will identify different examples of rhetorical convergence, and note the new intermediate forms that appear.

Eye and Ear. In the opening screen from "The Way West" (next page, top), music (harmonica, acoustic guitar, and double bass) plays as a background accompaniment, contributing to the Western pioneering mood. The two sign systems complement each other, each bringing one part of a combined message. Using music to evoke connotations in this way is common in film, but here it is coupled with an imitation of a printed page.

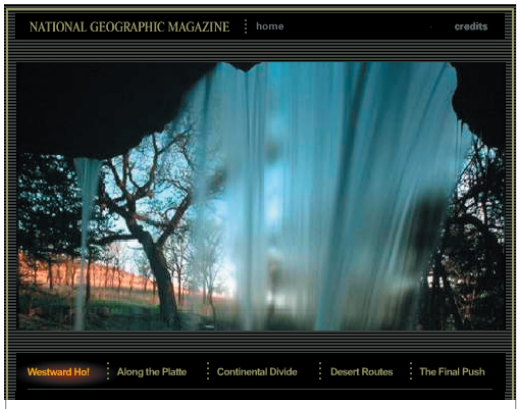
Later on in the same feature, still photography is coupled with radio – a different kind of eye/ear combination. The landscape photographs in the middle of next page are shown while a narrator reads the story of the pioneers, and the sounds of birds and a waterfall are also heard.



http://www.nationalgeographic.com/ngm/0009/feature2/media2.html



(Bird song) Voice-over: *Out onto the plains of Kansas.* New voice: *"Now we were out of civilisation and the influences of civilised society entirely, and cut out from the rest of the world to take care of ourselves for a while." Alicia D. Perkins, April 1849.*

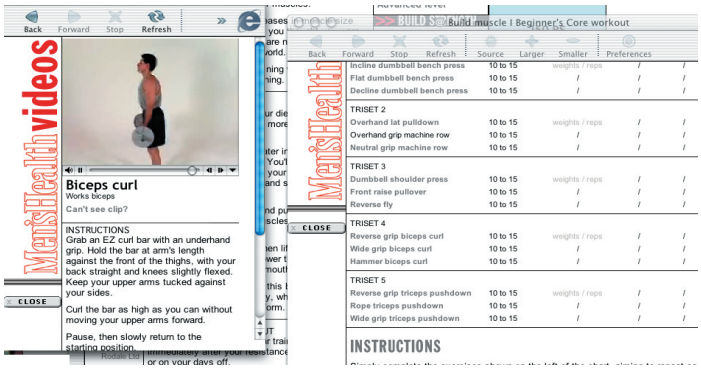


(Sound of waterfall) Voice-over: *When they came to Alcove Springs in Kansas, they described it as the most beautiful site on the whole trail, even though their whole trail experience had just been a few days.*

Rather than bringing different messages, sound and visuals here have the same content; the message is doubled. The sense of reality and immersion is heightened when they align. It is a little more like standing there, experiencing the landscape than images or sound alone would be.

S*tatic and Dynamic.* In these two examples from "The Way West," sound and visuals also bridge another division: that between static and dynamic sign systems. As sound also exists over time, it adds a temporal dimension and dynamic to the scene. In the still image of the waterfall, one can almost see the water moving when the sound is added.

Another effective combination of static and dynamic sign systems is found an exercise page from the British *Men's Health* site, which combines written words and print-like layout with video. The exercise program is explained in writing, and can be consulted over and over, while the exercises are demonstrated in video, thus showing the actual movements.



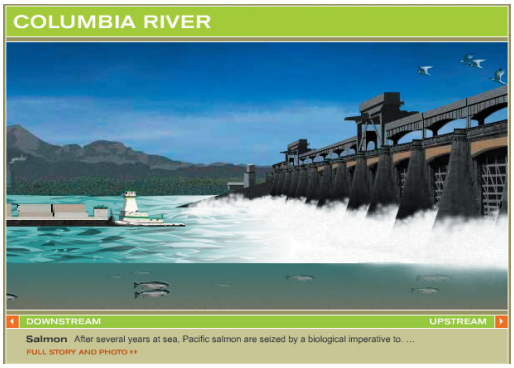
http://www.menshealth.co.uk/fitness/owen/

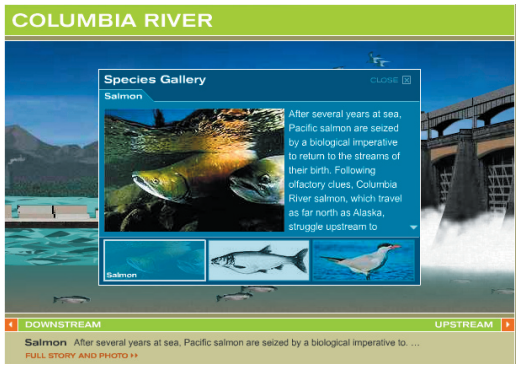
Differentiated and Dense. Not just different in being static and dynamic, writing and video are also differentiated and dense respectively. This adds to the effect of rhetorical convergence in the above example from *Men's Health*; as the video images are dense, all details are recorded, and may be studied by the aspiring weight lifter, including the finer points not mentioned in the written instructions.

Another combination of differentiated and dense sign systems is of course images and writing, a combination so common, it hardly deserves the name *rhetorical convergence* (was there ever a time when people never drew images on the same surface they were writing on?). A computer version of the combination, however is to let the user make the writing visible at will. *Nationalgeographic.com*'s "Columbia River" is a moving panoramic image of the river. When the reader positions

the mouse over an element in the image, a written label appears, and when the reader clicks, a smaller pane with more writing opens. This particular combination of image and writing makes it possible to combine a large-scale, detailed image with explanatory labels without cluttering the image with letters.

In the illustration below, a fish in the water was clicked, bringing up the pane with writing and images.





Attenuated and Replete. The image in "Columbia River" is a stylised drawing, which allows the artist to draw attention to the details he conciders to be important. In Goodman's vocabulary, it is relatively attenuated. When the fish is clicked, however, a photograph opens, showing what the fish "really" looks like in all its details. The photograph is replete. In 1.2.4.1, we noticed how effectively "Oleada" combined photography and drawing to reap the benefits of both the repleteness of photography and the attenuation of drawings, providing both overview and detail.

http://www.elmundo.es/elmundo/2001/graficos/septiembre/semana2/atentados/atentados2.html

Iconic, Indexical, and *Symbolic.* "Oleada" also demonstrates the combination of iconic, indexical, and symbolic signs. The photographs are indexical and iconic, the physical traces of the catastrophe, while the drawing is iconic but not indexical.

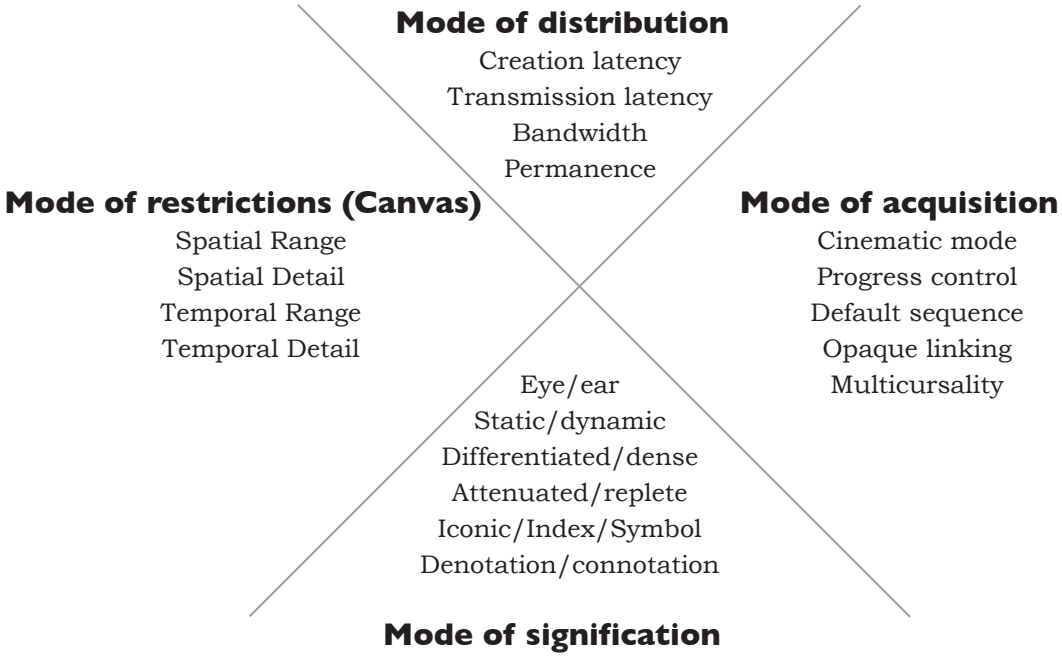
Similarly, *BBC News*• regularly combines the indexical with the symbolic, by linking parts of radio programs from newspaper-like written news stories. A quote in writing may thus be backed by the recording of how the statement fell, and the recording will also reveal the tone of voice of the speaker.

The ability to combine so many modes of significations within a Web site was the starting point for our investigation of rhetorical convergence of the Web, and its most visible and basic manifestation. But the obvious multitude of possible combinations of signifiers with different properties, and of different kinds

• <http://news.bbc.co.uk> This combination is discussed further in chapter 8, "Combining Video and Writing."

of semiosis makes the term *convergence* seem a less fitting description for the actual resulting text. Does not the discussion above indicate a *divergence* of rhetorical forms? We will return to that question towards the end of this essay, but first we should view the four axes of rhetorical convergence together.

The table below lists the four axes and the different terms we have discussed. (Neither their placement, nor the dividing lines have any significance, it is merely meant as a summary.)

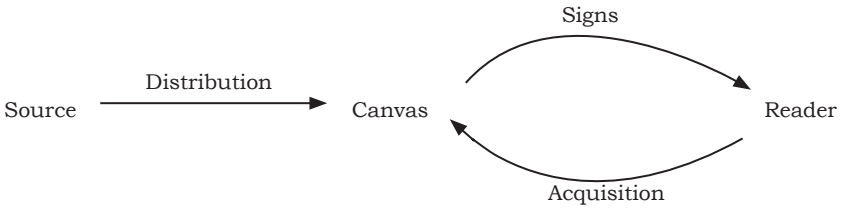


1.2.5 Limits of Technology

We have discussed the four axes of rhetorical convergence: *mode of distribution, mode of restrictions, mode of acquisition,* and *mode of signification.* Any rhetoric may be described by registering the variables listed for each of the axes, also established, well-known genre rhetorics. Each of the many Web texts discussed in this thesis will score so close to a genre rhetoric known from earlier media that we recognise them as fairly similar. In each of the cases, however, there are also some variables that are different, they will show similarities to a different rhetoric of a different established medium. It is this simultaneous resemblance to two or more established rhetorics we call *rhetorical convergence.*

What we have left to clarify is the relation between the four axes. It might be tempting to align the four axes with other familiar distinctions, for instance, to say that *mode of distribution* and *canvas* are aspects of technology or even medium, *mode of acquisition* is syntax, and *mode of signification* is semantics. Such a division would not stand up to inspection; technology is present as a factor in all four axes. As well as setting the limits for distribution speed and canvas, technology also governs which sign systems may be used, and the possibilities for user input and influence on the text. The mode of acquisition and the mode of distribution chosen by the authors will also influence on the semantic content of a text. We have already seen how a text may be read differently if it is live, or how bandwidth limits the use of video. Furthermore, multicursal aspects or merely playback controls open for a different reading and understanding of a text.

A perhaps more helpful way of relating the axes is to view them as a process of communication.

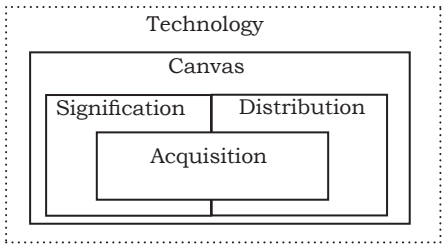


Then the mode of distribution would be how the message is brought to the reading surface, canvas is the properties of that surface, the mode of acquisition governs how the reader manipulates the surface to experience the signs, and the signs in turn are what the reader reads.

Such a highly abstracted view of a reading process may be helpful to memory, but as all models, this way of aligning the axes also obscures some relations as it highlights others. Although reading and comprehension take place in the meeting of text and reader, the rhetoric of the text is shaped by all earlier stages of the communication process. The axes of rhetorical convergence are not events that take place one after the other, but simultaneous dimensions that may describe any given rhetoric.

It is further important to realise that technology determines all four sets of aspects. Technology sets the premises for the possible modes of distribution and acquisition, technology dictates which sign systems may be used, and the canvas is just a subset of the possibilities of the technology.

As such, the four axes may also be seen as a set of limitations, of subsets of subsets of technology's possibilities, as in the following drawing.



The canvas is a set of chosen limits within the possibilities of the technology, a subset that sets the limits for which sign systems can be used and in which ways. The canvas also determines the distribution process, as a large and detailed canvas requires more bandwidth. The mode of signification and the mode of distribution are relatively independent, however, as the mode of distribution describes the communication as process, and not the semiosis itself. It is a way of expressing the temporal relation between signifier and signified, between story and discourse.

As modes of acquisition have to be signified, these depend in turn on the sign systems. The mode of distribution also determines the mode of acquisition, as all other acquisition modes than the cinematic require some permanence.

Such a theoretical model is not a description of a procedure or design workflow, it is probably rarely the case that screen size is decided on first, then writing and editing, and then linking. It remains a concern, however, that if one is to link from a video, for example, the presence of a link needs to be signified.

Also this perspective obscures one important relation: that the distribution process is a prerequisite for the communication to be possible at all. Each of these alternative views of the relations between the axes have its strengths and weaknesses, which is why it seems justified to view them as interrelated axes on which a multidimensional space is projected.

How does this multidimensional model relate to the rhetorical heritage? Traditional rhetorical figures reside inside what is signified, and are thus not visible in this model. This does not mean that they are not important, on the contrary: I have argued that the effects from choices along the four axes of rhetorical convergence have effect on the message. Still, this effect is probably less in terms of persuasion than the message conveyed by the signs.

After this long and detailed discussion of four dimensions of rhetorical convergence, we will attempt to bring the details together by analysing one concrete example, *Becoming Human,* produced by Arizona State University.

1.3 Rhetorical Convergence: *Becoming Human*

Becoming Human is a Web site made by Arizona State University's Institute of Human Origins, about paleoanthropology, that is, the origins of man. Its huge collection of well-written and engaging material on our predecessors and an extensive use of multimedia helped it win the *Webby* award for best science site 2001. Bringing together many different styles of writing, editing and "interactivity," it is an excellent example of rhetorical convergence.

The site consists of four main parts: A news section made solely in writing; a half-hour long documentary film (made of still images; a slide-motion film); forty-two "exhibits," that is, in-depth sub-sites on different aspects of man's origins, containing a wide range of sign systems and modes of acquisition; and a "learning center," also rich in sign systems and modes of acquisition. From the documentary film, links to selected exhibits appear, binding these two sections together.

Before moving on to the next section, I will invite readers to make themselves familiar with the site, which is found at <http://www.becominghuman.org>.

1.3.1 Distribution in *Becoming Human*

The mode of distribution in *Becoming Human* is halfway between a book and a magazine. Its contents are permanent, and can be accessed any time for the interested reader; nothing changes place or is removed (as it would in a news site such as *Cnn.com*). At the same time, new material is added to the site, although at a slow update rate. During 2001 and 2002, new items were added to the "news" section about every other month. In the same period, new "exhibits," "learning activities," and links to related Web sites were also added.

The most dynamic part of the site is the "News and Views" section, which is not a main feature. Its function is more of an afterthought, a source for updates for those readers who are already familiar with the main section, the "Documentary." Although it looks somewhat like a newspaper in its layout, its latency is much higher than in any news site. Most of the news items are links to other sites, and usually added several weeks after the articles were initially published in the remote Web sites, so the creation latency is quite long.

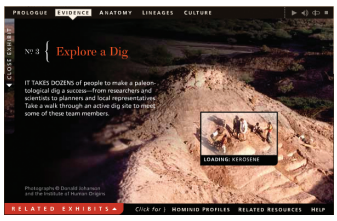
A long production latency is not a problem for a popular site about paleoanthropology. For a layperson with a limited interest for the finer scientific details, this is an area where new developments are slow. The site is so large – thirty minutes of slide-motion film plus forty-two engaging exhibits – that few Web readers will consume it all in one session. Clear navigation links to the exhibits and VCR-like controls and chapters in the slide-motion film makes it easy to return to the site to continue where one left off.

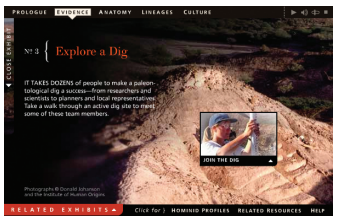
"Becoming Human – the Documentary" is made entirely in Flash. Its heavy use of moving images and sound makes a fast connection necessary to avoid long waits between sections. Yet, the choice never to include video, and instead relying on slide-motion-films seems to be an effort to save bandwidth.

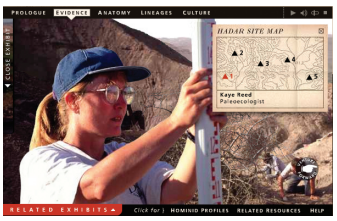
Even without video in the documentary, and using a speedy computer on a fast Internet connection, a reader often has to wait. Many rhetorical devices to cover the latency are found in the site. The most obvious

is the gradual appearance of the first screen of the documentary. More subtle is the choice to start each episode with a written title in white on black. Such a title screen loads fast, and can be shown while the rest of the segment still is in transit.

The largest exhibit, “Explore a Dig” (number 3 under “Evidence”), projects a sequence of images to cover download time: on top of the background image, a series of smaller images appear in the pane, with captions like “loading shovels.” When download is finished, and the reader enters the exhibit, the last of the small images is repeated as the background image.







Although looking quite a lot like television, *Becoming Human* is very far from TV in terms of distribution. It is slowly made; although there are continual updates both to the “News and Views” section and the other parts, there is quite a bit of latency, and live broadcast is never an option. There is no broadcasting schedule to pay attention to either. All parts of the site are kept permanent, and are consistently linked so they may be found again by a returning reader.

1.3.2 Canvas in *Becoming Human*

Becoming Human uses a small canvas. Everything but the main navigation links is kept within a pane of 600 x 380 pixels, set off as a window or a “hole” in the middle of the page. This is somewhat smaller than a television screen, which is 720 x 576 (in the European PAL standard). Flash movies have to have a fixed, rectangular size, so the “content” pane is made to fit the size of the main attraction of the site, the “Becoming Human” documentary.

Four Axes of Rhetorical Convergence

As Flash movies have a fixed size, it is common practice to ensure it will fit on a small VGA monitor of 640 by 480 pixels. The canvas size of 600 x 380 pixels is kind to low-end users, but even on a medium-sized monitor, it may look a little disappointing: large areas of brown and black, and a small area of interest in the middle.

Being fairly close to the number of pixels in a TV image, the canvas size is one of many similarities between the documentary film and television documentaries. The exhibits, on the other hand, are more akin to print. While having enough size and resolution for television-like images, the small canvas size heavily limits the amounts of text that can be visible at one time. Most of the exhibits contain more text and images than can be fitted in one page, so the reader needs to "turn pages" to read the whole, as in the example below. The total text of the exhibit is only two short paragraphs, but the second did not fit into the first screenful. Rather than to use the common computer practise of turning the text into a "scroll" by adding a scroll box control, the authors chose to split the text over two "pages," simulating the codex format of books and magazines.

Similarly, in exhibits that offer a "learn more" link to extra text pages, these texts have been split into two columns to fit the format of the canvas. Setting text in columns is also a print solution; it is good typographic practise to break wide pages into columns to avoid long lines that are difficult to read. In computer interfaces, "scrolling" one narrow column is again a more widespread practise. Writing displayed by a computer is not fixed like ink on paper, and does not need to be restricted by paper size. And while paper is a big cost in print, the "paper" Web pages are made of is free. The use of text in *Becoming Human*'s exhibits is thus a paradox: print-like in its use of typography and columns, but kept within the limits of a television-like canvas.

To save screen space, all navigation links inside the documentary are kept as narrow bands along the edges, which open up as drawers when clicked. Very unlike most Web sites, this is a device tailor-made for the small canvas in this site.•

The decision to keep the content inside a "hole" of consistent size also has consequences for the pages that are not made with Flash. *Becoming Human*'s home page (or "contents" page as it is called in the link texts in the site) can only fit one image and five short front-page leads. The "News and Views" section is similarly limited, its pages are viewed "through the hole," making frequent

• The sliding "drawers" would have been difficult to make without using Flash, but then again, it is not unlikely that the requirement of Flash movies to have a constant frame size is the reason for the small canvas. If this is the case, then Flash technology solves the problem it created.

scrolling a necessity. For the sake of illustration, I made my Web browser display the news in a full window, instead of restricting it to the small "hole" in the page. The difference is quite striking.

As noted in 1.2.2.1 above, the kind of "interactivity" that changes the display according to user input is often little more than a remedy for the low resolution of the computer screen. Similarly, much subject-activity in *Becoming Human* is caused by the small spatial range chosen for the canvas. We noted the need for "page turning" above; another example is the scrolling exhibit menu. Always is available from the lower edge, the "pull-up" menu covers a third of the height of the video pane (were it taller, it would obscure too much of the film when it is running). This is not enough to fit in all the exhibits, so the menu can be scrolled horizontally, revealing menu items that are outside the frame.

In *Becoming Human,* the rhetorical convergence of visual styles from television, print and computers within a television canvas has lead to several compromises, often requiring the use of canvas control to bridge the opposition between a small television canvas and the space-consuming practises of writing and print lay-out.

1.3.3 Acquisition in *Becoming Human*

To experience *Becoming Human* is to oscillate between many different modes of acquisition, as it brings together such different forms as television documentary, written essays and explanatory diagrams. The linear television rhetoric of the documentary film is coupled with hypertext linking known from computer texts. At the other end of the links are the exhibits, which are laid out with writing and images, much like popular magazines are. In these exhibitions, however, there is a wide variety of ways the user can use the mouse to alter the display. Most of these are features that can be recognised from popular computer applications and the CD-ROM publications of the 1990s.

It is difficult to single out a dominating reading mode in this site, it will vary during a normal reading session. We need instead to look at different levels.

At the site-wide level, what we called *default sequence* in 1.2.3.1 is probably a fair description. Although the site provides good navigation links throughout, it designed so the most effortless sequence is to begin with the documentary film, perhaps with a few detours into the

exhibits linked from the documentary, and later delve more deeply into the various exhibits, news items and other resources. Its construction is thus similar to the *delta model* proposed by Gunnar Liestøl in "Rhetorics of Hypermedia Design" (56–60). Liestøl recommends moving from video in the initial, top-level nodes of a hypertext to writing and interactivity in the lower levels, that, is from object-action to subject-action.

We may then consider the documentary film and the exhibits separately. The film is linear, but the reader has what we called progress control in 1.2.3.1. The documentary has to be actively started at the beginning of each of the five sections. The reader also has access to links to the beginning of the five sections, a pause/play button, and full controls to any point inside each section. All these possibilities are available, but the film does not invite the reader to use them. Like a television or film production, it is made in one fluent line, so to use the controls is to interrupt, to intervene and stop the object-action mode. An engaged user is probably more likely to use these controls to return to something she or he saw earlier, than to actively alter the sequence of the film's parts.

Compared to earlier media, *Becoming Human*'s documentary film seems most similar to a television program recorded on a VCR. It is a compromise between two strong traditions: the linear pull of television rhetoric, and the "interactive" tradition of computer media, where the user is accustomed to have freedom of choice.

As for the exhibits, they are so varied that the best description of the mode of acquisition would be exploration. Not only exploring what is said, but the different ways of reading.

This becomes clearer if we consider the many different modes of acquisitions found in the forty-two exhibits. We will sort them into three groups: *page turning, menu selection,* and *display change,* and try to classify these groups according to the typology of modes of acquisition suggested under 1.2.3.1, but to do so will immediately make it clear that the five modes proposed

are ideal types; many exhibits will seem to belong in two categories.

Page Turning. All the exhibits open with a combination of writing and images similar to a magazine spread or a poster. Entering an exhibit from the film is thus a transition from object-activity to subject-activity. Inside each exhibit, the user can change the display with the mouse; either by clicking or simply moving the mouse so the cursor is over certain "hot spots." The simplest of these is just to move between – to "turn" – written pages. Twenty-two of the forty-two exhibits (as well as ten "learn more" sets of sub-pages) are in progress control mode, as they contain two to six "pages" of writing that are "turned" by clicking with the mouse. To read these exhibits is very much like reading a book or a magazine.

In two of the exhibits, the user has even less influence of the reading sequence. After the initial screen, the user clicks to start a sequence in cinematic mode. Two examples are shown below.

Exhibit number 5 under "Anatomy," titled "Bringing the Past to Life" contains a new slide-motion film. Set in a small pane inside a written introduction, photographs of the artist John Gurche at work are projected, while we hear a recording of Gurche explaining how he makes his models of prehistoric hominids. The film runs for a minute and eighteen seconds, and can only be interrupted by leaving the exhibit.

Similarly, "Consciousness" (exhibit 14 under "Culture") contains a one minute and forty seconds sound clip of psychologist Steven Pinker.

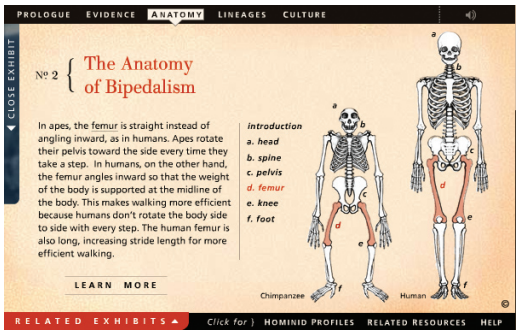
It may seem drastic to call so short clips "cinematic," but when reading *Becoming Human,* these exhibits make the reader markedly more passive, as one has to wait for the clips to finish.

Menu Selection. In each of the twenty-four exhibits I put in the group "page turning," there is one sequence the reader has to follow. The remaining eighteen exhibits are multicursal in one way or another: in seven, the text is broken up into "chunks" or "nodes," and the reader chooses the sequence he or she reads them in. I call the mode of acquisition in these seven "menu selection," as they all have a central list of the nodes. Reading thus becomes a zigzag motion: reading the menu, selecting, reading, going back to select another, and so on. The menus take different forms, suggesting ties to different media.

"Hominid Profiles" has a pull-down menu exactly like the menus of computer applications with a graphic user interface.

In "The Creative Explosion" ("Culture," number 13), seven archaeological sites are marked on a world map. Selecting one of them will display a collection of images from that site.

The collection is presented in a format known from photo collections on the Web and certain other computer interfaces: from a group of smaller "thumbnail images" the user selects which one is displayed in large format in the larger pane. Both these kinds of menus borrowed from computer interfaces are multicursal. Other exhibits have menus more similar to tables of contents from print, and these often suggest a reading sequence.



In "The Anatomy of Bipedalism" (number 2 in "Anatomy"), a menu selection in the middle column will alter the text in the left column, and highlight a different body part in the illustration. The sequence of reading is up to the reader. In the "page turning" exhibits, the text is broken arbitrarily, and mixing the sequence of the pages will make the text jumbled. This is not the case in this exhibit, the paragraphs about the different body parts may be read in any sequence. The positioning of the links from top to bottom and the alphabetic sequence of the initial letters strongly suggest a sequence, however. In 1.2.3.1 above, I called this mode *default sequence*.

Four Axes of Rhetorical Convergence

A default sequence does not necessarily mean less interesting, it is merely a description of a mode of acquisition. The elaborate exhibit titled “Explore a Dig” is what I elsewhere• have called a “virtual museum.” It is a depicted space in which the reader can move freely, a form most often seen in architecture models, flight simulators, and virtual reality installations. But unlike these forms, where either the space or the movement is the main purpose, a world of contents uses this space as a spatial table of contents. “Explore a Dig” is a wide, panoramic image of an archaeological dig, which the reader can move horizontally, not unlike walking alongside the actual site. When a person in the picture comes closer, the sound of him or her describing his or her job fades in. The reader may stop moving and listen, or move on, in which case the voice cross fades with the next person’s voice. A little map is inserted, and may be clicked on to speedily move to one of the five speakers.

• In the essay “Quests and Worlds,” chapter 9 in this thesis.

(Sound of thunder, scraping and digging)
Female voice: *As a paleoechologist my main role in the field is to find the fossils, and the second thing is to identify the fossils. And we find all kinds of exciting animals: antelopes, sabre tooth cats, hyenas, pigs, hippos, crocodiles, turtles, elephants...*
(Reader scrolls right. thunder and digging continues, female voice cross fades with...)

male voice: *...samples for purposes of dating. The samples are taken back to the laboratory where minerals are isolated from them, and we perform radiometric dating on these using sophisticated instrumentation.* (sound of pick) *Ahh. These are nice samples...*

Again, a default sequence is inscribed; the most natural reading mode is just to "walk along" from left to right. We may move any way we want, but this is the most likely, as we always start at the extreme left end of the scene, so we at the outset only can move to the right. This does not make the scene any less engaging, however.

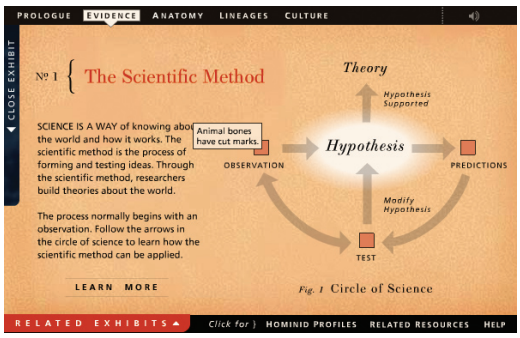
Above, I described the "Consciousness" exhibit as cinematic, as the reader has to sit passively and listen. It may then seem paradoxical that I will call a very similar exhibit multicursal.

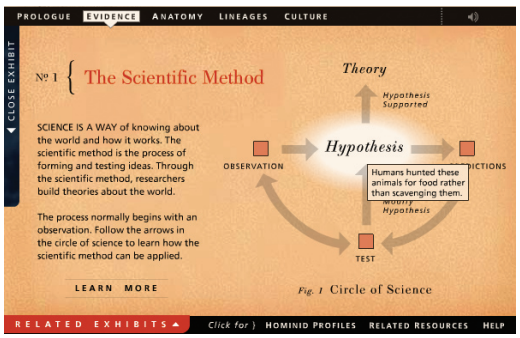
In "Humans' Place in Nature" (exhibit 15 under "Culture") there are not one, but four audio clips with people speaking their opinion. It is the fact that there are several clips to choose among that makes the difference. They are all answers to the same question ("What is our place in nature?"), and do as such not have any inherent sequence. We understand their relation, and are free to move between them: if one recording, is started and then immediately another is clicked, the first stops when the second begins.

Why do I call this multicursal why I called "The Anatomy of Bipedalism" default sequence? Because "Human's Place in Nature" have small written excerpts from each recording. They giving the reader a little more background to base the choice of sequence on, and they reveal that the recordings are both independent and containing similar material, the very requirement of multicursality (see 1.2.3.1 above, and my essay "Linearity and Multicursality," chapter 5 in this volume).

Display Variation. In most of the third and last group of exhibits, the user does not need to click on anything to change the display, it is sufficient to roll the mouse over certain spots.

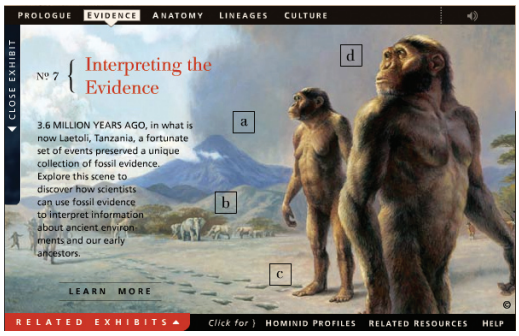
In "The Scientific Method" ("Evidence," number 1), the illustration shows a general "circle of science." By rolling the mouse over the stages in the circle, an example is shown in writing to explain the model.

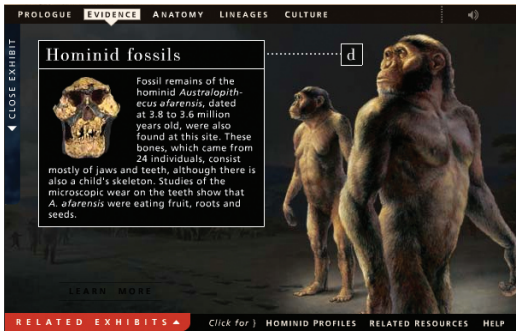




Similarly, in "Interpreting the Evidence" ("Evidence," number 7), positioning the cursor over parts of the painting of a scene will display an explanatory paragraph and even an supplementary explanation of that detail. At the same time, the rest of the image is dimmed.

Rather than selecting different parts from a menu, I view these "rollover" labels as manipulations of one figure, or model, inspecting its parts in detail. This kind of "interaction" is also well known for computer users. Explanatory labels are used in the interfaces of operating systems such as Mac OS or Microsoft Windows, and was also a popular feature of the mid-1990s CD-ROMs made with Macromedia Director.





In a few exhibits in *Becoming Human*, both clicking and "rollover" is combined, as in the complex exhibit "A World of Diversity" ("Lineages" number 10). Pictures of human faces are placed on a world map. Rolling over a face will display a label with the name of the person's race.

In the table in the left half, the reader can select "Skin Colour," "Lactose Tolerance," or "Blood Type." When one of the three genetic traits is selected, the world map will change, highlighting the races that share one of the varieties of this genetic heritage, for example, blood type A. On a new menu that has appeared, the other blood type (B) may also be displayed.

The exhibits I have grouped under the heading "display variation" are all multicursal, they invite the reader to comparison and exploration without any set sequence, and with few or no surprises as to what will happen when a link is activated by clicking or rolling over.

Another way of describing the modes of acquisition in *Becoming Human* is to classify the links as navigation links, presentation links and relation links. All the "page turning" and "display variation" exhibits are examples of what I have called presentation links, while the "tables of contents" are lists of navigation links inside each exhibit. Then there are also the site-wide navigation links that always are visible at the top of the pane. *Becoming Human* does also contain relation links, but only a few. There are twelve links from the documentary film to related exhibits, and then there are links from one exhibit to articles in the "hominid profiles" section. Obviously, there could have been many more relation links in *Becoming Human.* There is no part of the documentary film that is not related to one or more of the exhibits, so there could have been "learn more" links displayed constantly. There are furthermore no links from the exhibits to the film, except that if one followed a link from the film to an exhibit, then the "close exhibit" link would bring the reader back to the same place in the film. Finally, many of the exhibits are also clearly related, but with no hyperlinks between them.

The result of this restricted use of relation links is an orderly and hierarchical hypertext. It is closer to mainstream Web sites than it is to literary hypertexts of the "Eastgate school" or George Landow's large hypertext collection *Victorian Web.* Rather than being a network of crossing relations, it is a combination of a strict hierarchy and what Ted Nelson called a stretchtext ("Stretchtext"). A stretchtext is a text that follows a set sequence, but where the user may stretch or contract this sequence by regulating the level of detail. In *Becoming Human*, there are two basic ways of reading: either to watch the documentary film, that follows one sequence which may be expanded with "detours" into selected exhibits, but

the basic sequence stays the same. The other way is to use the menus of available exhibits to navigate in the hierarchy.

Becoming Human brings together several modes of acquisition: a cinematic mode, a VCR-mode, a book mode, and many different forms of computer reading. Effectively the text bridges two popular modes of acquisition and combines the linear sequence with "browsing." Reading *Becoming Human* also made one effect of the rhetorical convergence evident to me: as a reader, I start to expect ways to alter the display with my mouse. I became accustomed to the variety of acquisition modes, and wanted moving parts. I often felt a little disappointed when I came upon exhibits that just had written pages and still images.

1.3.4 Signification in *Becoming Human*

From the presentation above, it is evident that *Becoming Human* uses a wider variety of sign systems than is usual in any established media before the Web. Moving images, sound effects, music, speech, writing, photography, drawings, diagrams and maps are brought together, making it resemble many known forms in places, but as a whole, it is something new. This is perhaps the most obvious form of rhetorical convergence, and most of the many combinations of sign systems in *Becoming Human* have been described in the above sections. We will only indicate a few further details here.

In contrast to most of the Web, *Becoming Human* attract both eyes and ears. When the site won the Webby award for best science site 2001, one of its makers, Leonora Johanson, described the site in this way: "I harken this to a combination of a Ken Burns documentary and an illustrated NPR radio show,"• thus clearly indicating the rhetorical convergence found in the site. One might add that the length and arrangement of the storyline (the dispositio) is very familiar for TV viewers who have watched the Discovery channel.

Johanson's description fits the documentary film: it is a documentary made by adding camera movements to pre-existing still images (as Ken Burns did in his

• "Institute of Human Origins Wins Webby for Best Science Site." *Becoming Human*, April 20, 2002. <http://www.becominghuman.org>.

documentaries about the American civil war). Composed as still images, the photographs belong to an aesthetic that strives for a balanced effect of lines, volumes and colours. This is not counter to television practise, but when shooting moving images, a balanced composition may be sacrificed in order to capture the movement better. In *Becoming Human*, speech is always superordinate to the illustrations, and little would be missed by a reader listening with her eyes closed. Whether describing this from as "illustrated radio," as the producer quoted above, or "still-image TV" is a matter of taste; both terms would point to the fact that "Becoming Human: the Documentary" occupies a middle ground between TV, radio and photography.

Little would be missed with closed eyes – except the links. For the reader to notice them, they need to be both noticeable and present for a while. This is accomplished by a noticeable orange colour, and the use of writing, which is a static sign system, different from the soundtrack and the moving photographs and drawings.

Viewing "Becoming Human: the Documentary" is a saturated experience, like that of watching television and film. We do not just watch, we listen simultaneously. We are also able to comprehend the spoken narration while sound effects and music play in the background, and view short written sentences while keeping an eye on other visuals.

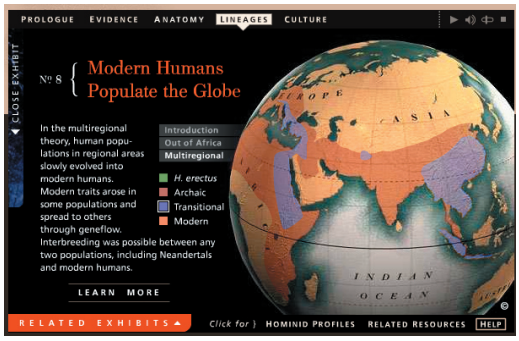
The exhibits are much more static and print-like. As noted above, they all open with a screenful of writing and images, laid out similar to a magazine or a poster. Some are saturated with sound too, however. Two exhibits combine several pages of writing with background sound effects; bird sounds in one, a roaring fire in another. Both add depth and reality to what is described in the written text, just as sound effects do in radio or film. Sound is dynamic, it exists over time, but in these cases that is almost not the case. The recordings play indefinitely in a loop, and as soon as we recognise the sounds as birds in a landscape or fire, we do not perceive any changes over time, and there is no problem in aligning the dynamic sound with the static writing.

When speech and writing is combined, however, they both require the language facilities of the reader, so then the sound recordings wait for the reader to start them, as in “The Neanderthal Question” (“Lineages, number 9).

Note also that two other sign systems are introduced to signify this organisation: a photograph of the speaker and a graphic symbol (a stylised loudspeaker with visible sound waves) indicating the presence of a sound recording.

The voices are recorded (indexical-iconic). This adds information that is could not have been captured in the same detail in writing or imagery: how the person actually spoke, such as tone of voice, emphasis and accent as well as hints of the speaker’s age.

Also the exclusively visual rhetoric in most of the exhibits contain converging uses of sign systems. Movement is impossible to show in print, but is used in a couple of diagrams to show change.



In “Modern Humans Populate the Globe,” two different theories on how the human species spread to new continents are illustrated with growing coloured fields on a map. Impossible in print, these animations could have been shown on TV, but then, the reader would not be able to replay and compare the two animations. In this exhibit, each stage in each animation can be played and replayed for study, just as still drawings are read by repeated comparisons between parts

Similarly, the 3-D models of hominid skulls found in the site are midway between static and dynamic, in that they are only dynamic when the reader wants them to be.

When two or more sign systems addressing the same sense are combined, they tend to be convoluted or contained, and also in the kinds of containment can we observe rhetorical convergence. In the documentary film, writing is contained within the frame of the image. In contrast, the short narrated slide-show "Bringing the Past to Life" is contained within a page of writing, similar to how an illustration is inserted in a book. Above, we noticed how audio recordings in some exhibits are available from links placed like illustrations in the running text, and the audio's effect is not unlike visual illustrations: a short distraction from the reading, explaining, or exemplifying something. In several cases, the audio recordings are quotes from interviews placed in the running text almost like written quotes.

Studying *Becoming Human*'s mode of signification, we are able to find many examples of rhetorical convergence, as we did with mode of distribution, mode of restrictions, and mode of acquisition. If we should summarise the findings, we could say that *Becoming Human* brings together a certain kind of television rhetoric with many devices known from print and computer applications. I do not want to summarise too much, however, but rather note the variety. By way of a summary, the table in the Appendix (page 366) renders some of the many relations we have noted.

Nowhere is *Becoming Human* a direct replica of rhetoric from another medium; we always recognise several rhetorical traits at once. As such, there is convergence on every page. Unlike some uses of convergence, however, where it is used to describe a development where several things come together into one, *Becoming Human* appears to be the result of a process of convergence where the result is more forms than before. *Rhetorical convergence* is at the same time a *rhetorical divergence.*

1.4 Rhetorical Divergence?

In an area in the Pacific, known as the intertropical convergence zone, the strong trade winds from the east collide head on with winds from the west. It is a phenomenon meteorologists know as *convergence.* The colliding winds bend upwards, and the high pressure is released as the winds spread out again in a *divergence* on a higher altitude.

This thesis argues that the perceived convergence of media may be viewed as combinations of earlier form, or forms sliding towards each other as variables change. Some combinations are time-tested and traditional; others are new and creative. Combinations adhere to certain restrictions and follow certain patterns, described in the section above using *mode of distribution, mode of restrictions, mode of acquisition,* and *mode of signification.* Together, the model of the four axes collect the findings in this thesis. It is an inductive construction based on the examples collected here, so there is no way of knowing if it will be sufficient to explain other texts. I have, for example, chosen not to include fictive texts or texts were users may contribute their own writing or images. I believe the four axes outlined here still would be useful for such texts, but it seems reasonable that dimensions such as questions of fact and fiction and patterns of sender and receiver would need to be added in order to map convergence in such genres.

As each of the four axes span many modes and dimensions, their possible combinations are numerous. Traditional media utilise only a few combinations of modes, many more are possible on the computer. What has become possible is a veritable *divergence* of rhetorics. I have neither hope nor ambition of describing all aspects of rhetorical convergence, but even the small number of perspectives I have discussed can combine in a vast number of different forms, as a little math would easily demonstrate. And the likely finding of just one other rhetorical dimension would drastically increase the total number of possibilities. Hence, although I have argued that *convergence* is a meaningful term to use,

this *convergence* can only result in a *divergence* of forms. We haven't seen most of them yet.

This thesis has given a long answer to a simple question: "Do the means of expressions on the Web converge, and if so, how?" The short answer to the first half of the question (before the "how" that serves to legitimise the length of this text) is: Yes and no. Yes, there is convergence of forms, but not just one or in just one direction. Rather, a new range of new Web media rhetorics are established alongside others.

As such, it is an old story. People have reminded me throughout this project that it is an old (maybe the first) insight of media studies that new media first borrow forms of earlier media. Even polar opponents such as Williams and McLuhan agree on this point. It is also the case of the Web, as one could expect.

Still, I hope to have contributed a little to media studies. First of all, I believe to have introduced some new pieces of terminology that might serve other intending to describe not only Web sites, but also other texts that are "multimedia," "interactive," or "online." Also the job of fitting semiotic terms to Web communication is begun.

It is further my hope that this thesis has demonstrated that there is more work to be done in the question of sequence in texts, be they linear or multicursal. This is also just one example of the lack of theory of the Web as text. In this respect, this thesis has been a first in many areas. By trying to position Web texts in relation to texts in earlier media, I also believe that some rarely asked questions of media difference, of the relation between signifying systems, and of the relations of genres and media have received some attention.

None of these issues have been treated at sufficient length in the following pages; there is much left to do. Indeed, where to stop a study of this character is a pragmatic question. It could go on and grow infinitely, and would only be complete when every piece of human communication is charted, an infinite and impossible outer limit. Though this study is far from complete, I hope my readers agree that we have reached a sensible point of conclusion. The essays span convergences

of newspapers, magazines, photography, graphics, television, and computer games. The four axes outlined give an overview over the many kinds of convergence found, and show that many more are possible, revealing the dynamic of convergence and divergence.

The many possible further routes will be left for future works. It has never stopped amazing me that there is so little theoretical work on Web textuality. I hope my shortcomings and errors can provoke others to start filling this gap.

1.5 Further Perspectives

The process of rhetorical convergence and divergence opens a rich palette for creative authors, whether they call themselves Web designers, information architects, new media producers, or simply writers or people with a home page. Their diverse products provide much fertile land for the textual analyst. The Web has already yielded an enormous range of widely different forms, genres and even media. I have tentatively described some Web genres such as photo collections, interactive graphics, video-on-demand and annoted video, but these need more serious attention, both in charting their means of expression and in knowing their range and variation as genres.

Many new and powerful Web media have appeared while this thesis was written: Personal, collaborative and networked forms of writing have appeared in the form of user-driven Web magazines such as *Slashdot* or *Plastic*,• and in "blogs" (short for Web Logs) and Wikis (Web sites for collaborative writing using the Wiki technology). So-called "file sharing" or "peer-to-peer networking" reached the headlines of older media during the rise and fall of the file-sharing network *Napster*•• with its millions of illegally distributed music files. *Napster* was closed, but other networks such as Gnutella have replaced it. The Web has also become an outlet for independent film and television producers, who may get an audience for their creations on *Animation Express* or *Atomfilms*.••• Streaming audio technology has created thousands of

• <http://slashdot.org>; <http://www.plastic.com>

•• <http://napster.com>

••• <http://www.wired.com/animation/>; <http://www.atomfilms.com/>

"Web radio" channels and even television channels that are only broadcast on the Internet.

These examples differ from the Web sites I have studied. Where I have focused on Web media that clearly derive from earlier media, these other examples are media that have different power relations, being more networks than the traditional star-shaped pattern of one source and many receivers. These new patterns of communication deserve study. At the same time do all these media rely on old forms, and may thus be studied from the perspective of rhetorical convergence. As these new Web media and numerous others gain larger audiences, they should also get the attention of media scholars.

• Chapter five in this thesis.

In the essay "Rhetorical Convergence,"• I called for a catalogue of convergent forms on the Web. Although this is modestly started in this thesis, it should be clear that much more work in this direction is needed, as also the above thoughts on new Web media indicate. It would also be desirable to get a better overview of what the Web actually looks like, an overview beyond the growing numbers of servers and pages in the databases of the largest search engines. Also the actual reading patterns of real users, what they read and how, is something we know too little of. Even if Web audience measurement is being conducted for the purpose of market analysis, these are restricted to the largest sites of the largest nations. That is not satisfactory for the Web, which is characterised by the exact opposite pattern: a multitude of small sites, knowing no national borders in reach and readership. The challenges of getting to know this terrain better can only be met by large-scale, international cooperation.

I have argued throughout my essays that there is a tradition of computer rhetoric in the form of interface design, games, and applications that influence on computer texts, and needs to be studied in its own rights. Although much has been written on computer texts, most has been limited to contrasting computer media with print, television or cinema. Actual insights into the special nature of computer rhetoric are few

and scattered among several works that do not refer to each other. Just as we need an overview of the many forms of the Web does it seem overdue to begin to form an overview of the traditions and conventions of five decades of earlier computer media.

There were many computer media before the Web, and many more have been designed after its arrival. Although the model of rhetorical convergence, and the rhetorical divergence it implies as its consequence is the result of studies of Web media, do I find it reasonable to believe that it could be applied on other "new media" too. Many solutions for digital radio and television have been launched, and more will probably appear. Portable devices such as mobile phones and small computers with calendar and note-taking tools known as PDAs (two classes of commodities that are themselves converging) promise to deliver news and entertainment in writing and video, wirelessly delivered to the small screens. As I am writing this, Microsoft has just advertised new tablet computers and wrist watches with the same features.

It seems likely that any new media machines that appear in the future will have to rely largely on established rhetorical means. Thus, the four axes of rhetorical convergence may serve as a roadmap also when studying the texts of future computer media.

The massive investments in new technology also creates a demand for new "content:" something to use the new media machines for. In keeping with the rhetorical tradition, the four axes may also be seen as a list of topics for new media rhetorics. Rhetoric began as a generative art, that of speaking in public, and to use rhetoric merely for analytical purposes without a generative aim is a fairly recent development. This development is part of a general trend in the humanities in Norway and many other parts of the world, where scholars study the works of others rather than themselves practising and teaching authorship. Within the generative branch of rhetoric, however, lists of topics that could be applied to any subject were always and important part, and included in most of the classical treatises on rhetoric. Our four axes of rhetorical convergence could be used

in a similar manner as a list of topics for multimedia authors. They outline a range (or rather, four ranges) of choices available for any Web author. Going through the possible positions on the four axes may assist the author in choosing when constructing a Web text, just as much as it may assist the analyst when studying the resulting text later on. Viewed as a list of topics, the four axes may outline the full potential of rhetorical convergence as rhetorical divergence.

I argue throughout this thesis that existing media theory cannot fully account for Web media. My choice has been to combine approaches from different theories on older media with new insights based in empirical analysis. Concepts from different theories need to be aligned and fitted together, a process I have called *conceptual convergence*.• The four axes outlined above may serve as a list of topics or a roadmap also for this process. If we, following an analysis in the manner demonstrated above, see that a certain Web form is similar to rhetorics of both print magazines and radio, this is also an indication of which bodies and parts of media literature one may look into for insights and inspiration.

• This argument, and the tentative concept of conceptual convergence is presented in "Rhetorical Convergence," the present chapter 3.

All communication is based in convention such as the semiotic codes, communication registers, genres and media. Such conventions are slowly established; rather than being designed, they evolve. New media are always just a little new. A theory of new media should ideally be able to account in a unified way for both what is new and what is heritage.