out one after the other as found in comics and similar picture sequences (cf. McCloud 1994). The temporal sequence is used to carry a range of additional meanings over and above those in the contributing images. The relations holding between successive elements in such image-flows have been analysed from several perspectives; they also overlap with the kinds of relations that can be observed in film (for extensive references and discussion, see: Bateman 2007). We will not discuss this mode any further here, however; we simply note that such elements also commonly appear as contributions to multimodal documents alongside, and in combination with, text-flow elements.

When a document starts to utilise the full two-dimensional spatial extent of the page for expressing rhetorical and other functional organisations, we move into a different semiotic mode: one which we term *page-flow*. Page-flow can combine elements in any of the semiotic modes appearing on a page, including text-flow, diagrams, graphs and so on. It adds to the individual contributions of these elements the possibility of a *rhetorical* unity supporting the communicative intentions of the document. And so, to documents in this mode, we use the multimodal extension to rhetorical structure that we have introduced in this chapter. Without this level of description, we are not in a position to explicate many of the spatial distribution decisions taken in page-based documents—although this distribution must also be considered in terms of any canvas constraints that are applicable, such as the grids and other properties of the page.

It is also only when we have a detailed view of multimodal rhetorical structure that we can start setting out the *alternatives* for expression that are open to designers. This is one of the main reasons that we require an account of rhetoric in our overall account. In addition, we have seen how intended rhetorical organisation and the relationships communicated spatially can diverge—usually resulting in incorrect or ambiguous interpretations of the document in question. This divergence shows the necessity of maintaining rhetorical structure as an *independent* contribution to a complete document analysis.

At the present time it is an open question as to just how much of the detail of rhetorical organisation is expressible visually. If we cannot make the fine distinctions that are commonly drawn linguistically in the visual mode then we will not require the full apparatus of RST. The resources which are actually employed in any document and the ways in which they are distributed around the semiotic modes mobilised will also depend to a large extent on the type of document considered. For this we need to consider the influence of document *genres*, which we now turn to in the following chapter.

5

Multimodal Documents and Genre

In this chapter, we address in detail the joint issues of *comparison* and *constraint*. We have raised the issue of comparison at various points in the discussion so far. Whenever we analyse a multimodal document, we need to consider the sets of documents that it resembles and the sets of documents with which it stands in contrast. This kind of comparison is of much more than academic interest: in fact, it is only *because* of this comparison that particular interpretations can be made rather than others. The overwhelming majority of constructs that we have seen in this book are elements and configurations that have, through use, *developed* the potential to carry meaning in multimodal documents: this potential is only made real in the context of actual documents standing in relations of similarity and contrast with other documents.

In short, when we see a text fragment shown slightly spatially offset in the text-flow, slightly larger and bolder than the text surrounding it, perhaps with a number before the text, this is *not* a section heading because of these features—it is (possibly) a section heading because there is an established body of documents in which this particular collection of typographical, visual and spatial properties is regularly deployed with the intention of signalling a textual division and a point of access into the navigation structure. Only then can we label this functionally as a 'section heading'. All such interpretations are dependent on the class of documents to which the analysed instance is allocated—a point recently made at length in terms of *conventions* by Kostelnick and Hassett (2003).

Allocation of this kind is part of the work of interpretation that any user of a document automatically performs. On this basis the user/reader brings appropriate interpretive schemes to bear so as to unravel what may be being

meant. Well designed documents necessarily take on the task of showing the user/reader just what class of documents is relevant for their interpretation. From this perspective, the class of documents intended establishes *constraints* on document design. These constraints inform design in as far as they clarify just what signals and conventions are going to be relevant to help some user/reader select an appropriate scheme of interpretation.

We shall argue in this chapter that a crucial theoretical construct for exploring this aspect of the meaning-making involved in multimodal documents is *genre*. The complex interplay of mutual relationships playing a role here is succinctly expressed by Lemke as follows:

> "We construct genres by construing certain sorts of semantic patterning in what we consider to be distinct texts, and we say that such texts belong to the same genre. Co-generic texts are privileged intertexts for each other's interpretation." (Lemke 1999)

And so, when readers allocate documents to particular classes of document, those classes bring with them certain interpretive frames and expectations. These frames guide readers to make sense of what they are seeing. Moreover, seen from the perspective of document *production*, we can take the intention of creating a document that 'belongs' to one class rather than another as a mobilisation of precisely those constraints that signal that some interpretative frames are to be applied rather than others. The decisions taken during production then rely more or less explicitly on the conventions and practices established for the class of documents to which the document is to belong.

The first explicit use of genre in a detailed model of the document design process was probably that of Waller (1987*a*). Genre within this framework is very much as it is seen in linguistic work: i.e., more or less stable categories formed by restricting theoretically available design choices to particular jointly selected combinations of choices. The position that genre takes in Waller's model is shown in Figure 5.1. On the left of the figure, we see the abstract paths available for moving from design task to design solution; on the right, we see the actual paths often followed in concrete design situations. As Waller states, solutions that only consider the norms of the situation, or the intended document, may produce unusable results because attention to those documents' functionality has not been paid; conversely, solutions that only consider functionality without considering the expectations that hold for documents of particular types may equally suffer from unusability simply because the users of the document do not know which interpretive frames to apply. Appropriate solutions require a process

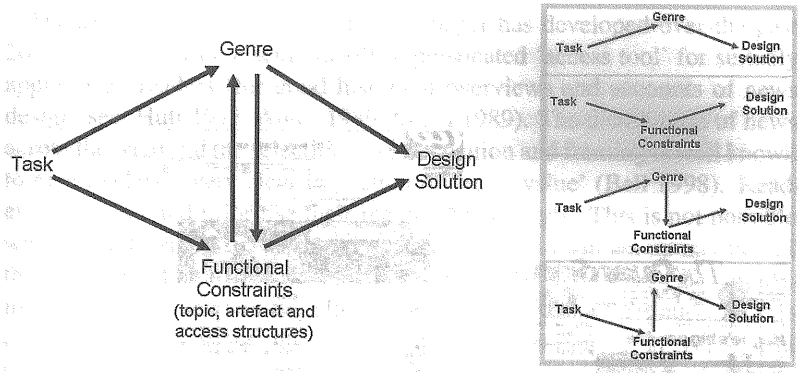
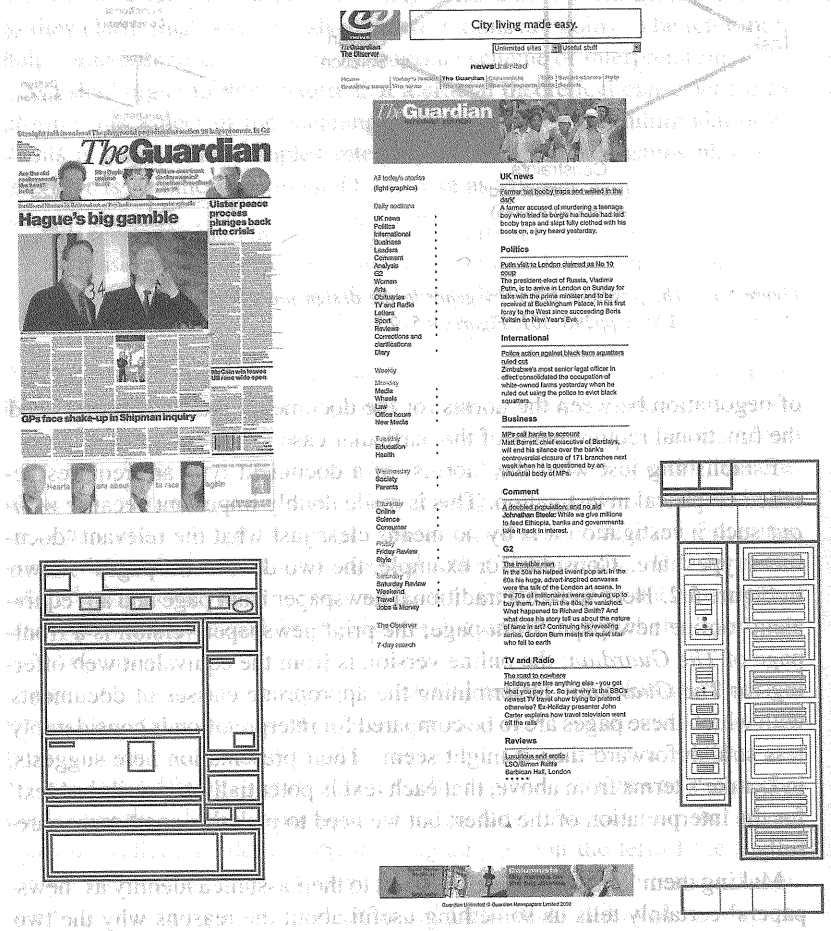


Figure 5.1 The positions taken by genre in the design process according to Waller (1987a, pp298–301; Figures 9.5–9.9)

of negotiation between the norms for the document type—the genre—and the functional requirements of the particular case.

Establishing just what the 'norms' of a document type are requires detailed empirical investigation. This is made doubly important because *without* such investigation it is by no means clear just what the relevant 'document types' are. Consider, for example, the two document 'pages' shown in Figure 5.2. Here we see a traditional newspaper front page and an 'equivalent' online newspaper front page; the print newspaper version is a front-page of *The Guardian*; the online version is from the equivalent web offering for *The Guardian*. Determining the appropriate classes of documents with which these pages are to be compared for interpretation is considerably less straightforward than it might seem. Their presentation here suggests, in Lemke's terms from above, that each text is potentially a 'privileged text' for the interpretation of the other; but we need to pick this apart more carefully.

Making them 'co-generic' with respect to their assumed identity as 'newspapers' certainly tells us something useful about the reasons why the two documents contain information of similar kinds: i.e., 'news'. But in other respects this view is less helpful. It does not, for example, give us a particularly accurate indication of the interpretative strategies the documents require of their readers. Interacting with a web-page is something very different from interacting with a printed page—and *not* only because of issues of hyperlinks and scrollability. As suggested by the graphical versions of the layout structures shown in the figure, we are also dealing here with two very different *spatial organisations*—and this has significant consequences.



(a) Print version (b) Online version

Figure 5.2 *Two versions of* The Guardian *newspaper—one print, the other online—each accompanied by a graphical rendition of its layout structure; both newspaper versions used by permission*

The virtual artefact of the print newspaper has developed over the past 200 years to provide a semiotically sophisticated 'access tool' for suitably apprenticed readers (for good historical overviews and accounts of news design, see: Hutt 1973, Ames 1989, Moen 1989). The distribution of news across the page and the selection of size, position and framing is well known to give readers a very clear indication of 'news value' (Bell 1998). Readers can select and prioritise their reading accordingly. This is not possible with the web-page because a very different canvas is drawn upon—that of the virtual artefact of HTML tables. Although HTML tables can 'approximate' columns, they do not allow text to flow on from the bottom of one column to the top of the 'next'; the traditional text-flow column grid of print (cf. Section 2.5.1) is therefore surprisingly difficult to achieve. The apparent grid-like structure of the web-page layout is in fact made up of single columns set out next to one another. These are, apart from some relatively weak coordinating reason for the columns being on the same page at all, *essentially unrelated*. A reader has little choice but to start at the top of each column and proceed downwards: the truely two-dimensional deployment of space within the modern print newspaper is absent.

This situation commonly leads the spatial organisation to take on different roles. For example, a central column may provide the main news items, while off-centre columns serve navigational or 'other news' roles (cf. Knox 2007). The online page is then in many respects closer generically to online offerings such as web-browser result pages and simple lists, possibly with subheadings breaking those lists down into categories. The *multidimensional* access possibilities of the print newspaper are severely reduced.[1]

Although there are many issues of design and usability to consider here, what is most relevant for our present discussion is the way in which the properties of the online newspaper align with very different sets of co-generic texts than might have originally been thought on the basis of its informal classification as a 'newspaper'. This can be problematic in several ways. For example, a reader who unknowingly applies a newspaper interpretive scheme is likely to be frustrated—perhaps without even realising it—because the artefact is assumed to have affordances that it does not.

The divisions relevant here also do not respect simple technological divisions, such as that between print and web-based delivery. There are many print documents that are also structured very similarly to this online newspaper—which makes interpretative strategies developed for those classes of documents equally relevant here. This is essentially the style of

[1] As we have discussed elsewhere (cf. Bateman, Delin and Henschel 2007), this is one of the reasons for a resurgence of the traditional spatial organisation for online newspapers now being successfully promoted in so-called *digital editions*.

interaction found with telephone directories and other listings of information, including some styles of newspaper from the *19th* century rather than the 21st! The success or not of the web-offering is then to be seen (a) in terms of how well this latter set of co-generic texts also meets the current-day demands and habits of presenting and receiving news, and (b) in terms of new practices for news consumption that it may facilitate.

To take this further, we need far more information concerning just what documents are relevant for comparison, and methods for determining how those documents contribute to interpretation. Without this, we face the danger that decisions in document design will either resist functional motivation or suggest functional motivations that are not relevant or appropriate for the document at hand—simply because the *actual* reason for their selection was better sought in the generic constraints of a class of documents, including historical and technological constraints, than in the particulars of the individual solution under scrutiny. A stronger position for interpretation and explanation therefore requires that we establish genres for multimodal documents. Only then we will have a theoretical framework within which constraints on the interpretations and meanings deployed in documents can be effectively pursued.

Although there is a traditional belief that genre is not amenable to precise formalisation and aids only in *clarification* (cf. Fowler 1982, de Beaugrande 1980, p196, Miller 1984), we will push for a tighter notion. We need to develop a theoretical construct which allows us to make *predictions* concerning the form and content of multimodal documents. But working with genre, and particularly developing it further towards predictive and empirically motivated explanation, requires some careful discussion. The extension of traditional notions of genre to multimodal documents is not straightforward. Just what genres there are and how they are best to be defined multimodally remains far from clear.

The main contribution to our framework for multimodal genre will therefore be drawn from precise *linguistically-motivated* accounts of genre. We consider this view of genre as the theoretical construct of choice for unravelling issues of the kind facing us here. From this perspective, genre offers a method for relating any individual document encountered to its 'generic' context by means of explicitly identifiable design decisions. These design decisions, i.e., what selection of modal resources a document makes and how these resources are related to the meanings that they carry, must be isolated by empirical analysis.

To move us forward, the chapter will first set out the current state of the art in genre as it has been applied to text and discourse.[2] We then pick out some common problems faced when moving from textual genres to multimodal genres; we will see that most of the perspectives we introduce have already been applied in investigations of multimodality, although yet again there is little contact across the particular research communities involved. We also briefly discuss work that has focused particularly on finding general classes of *visual representation* (e.g. Twyman 1979, Burford, Briggs and Eakins 2003, Kostelnick and Hassett 2003, Twyman 2004) since this often includes treatments of document pages and so contributes directly to our goal of providing an overarching account of multimodal genre as such. We then present the concrete notion of genre adopted in our framework and conclude with some illustrations of using genre descriptions in document analysis.

5.1 Perspectives on genre

The term 'genre' is widely used to label broad classes of communicative verbal artefacts. In this largely pre-theoretical sense, it has been relatively straightforward and natural to extend traditional talk of genres such as 'sonnet', 'tragedy', 'comedy' and so on in literature, into newer, multimodal realms, including the 'Western' and 'Film noir' in film, 'advertisement', 'documentary' and 'sitcom' in television, as well as newer genres still, such as 'the online newspaper' mentioned above or 'the homepage' in discussions of the World-Wide Web. Most of this discussion has not aimed at making genre function predictively and criteria for assigning an artefact to one genre rather than another remain loose.

To provide a more empirically accountable view we orientate more towards the development of genre pursued in linguistics and discourse analysis since there it becomes possible to make far stronger statements about the properties that an artefact conforming to some genre must exhibit. Two broad approaches are particularly important for us here. One of these reconstructs genre within a specifically linguistic account; we term this the *genre as social semiotic* perspective (cf. Martin 1984, Hasan 1984, Ventola 1984, 1987). The other, initiated by the influential paper by Miller (1984), is directed towards understanding communicative events in terms of their

[2] Although there is much significant work on genre, our discussion here will be focused solely on achieving a notion of genre usable for multimodality; a recent review of genre theory more generally can be found in Muntigl and Gruber (2005), and further introductions and discussions of schools of genre and their interactions are given by, for example, Hyon (1996), Paltridge (1997), Lee (2001), Johns (2002) and Martin and Rose (2007).

socially situated 'rhetorical actions'; we accordingly term this the *genre as social action* perspective. We will see below precisely why we focus on these two orientations; we can note here at the outset, however, that they both draw a distinction between genre on the one hand, and style, or register, on the other. We will rely on this heavily below: there are approaches where genre and style, register, etc. are not distinguished and we will argue that this reduces descriptive adequacy, making the genre classification of documents, particularly multimodal documents, unworkable.

5.1.1 Genre as social semiotic

Considering genre as simultaneously, on the one hand, a semiotically constructed social entity and, on the other, a characterisation of a class of identifiable linguistic artefacts, derives more or less directly from the view of language proposed in Halliday's (1978) *Language as Social Semiotic* and systemic-functional linguistics (SFL). Most approaches taken within this perspective on genre are aligned with this theory, which, as we have seen in previous chapters, continues to contribute actively to almost all areas of multimodal theorising.

In the systemic-functional framework language is seen as inextricably related to society, ideology and culture. The language system and its contexts of use are described at several distinct levels of abstraction, termed *strata*. These strata serve the function of successively mediating between, at the most abstract and general level, characterisations of language use as sociological, cultural and ideological phenomena and, at the least abstract level, accounts of the concrete linguistic forms, sounds and grammatical constructions employed. All of these levels of description are seen as capturing 'meaningful' contributions to a complete language event. This means that the account, in contrast to many other kinds of linguistic theorising, resists singling out any particular level as *the* level where 'meaning' is to be found. Any of the levels of the model may take up the role of constructing and carrying meaningful patterns.

This openness to considering linguistic patterns at any level of description contributed to the early development of a linguistic treatment of *register* (cf. Gregory 1967, Gregory and Carrol 1978). Register is generally defined as a particular 'style' of language use; this was taken further by several characterisations that related the concept centrally to social situation (Hasan 1973, Biber 1988, Labov 2001, Martin 2001), turning register variation into a powerful tool for exploring social and situational variation. Describing a 'style' of language use sufficiently exactly to reveal variation of this kind requires a fine-grained linguistic apparatus sensitive enough to

distinguish linguistic expressions that might, more superficially, all be said to 'mean' the same thing. Accepting style as an equal partner in the process of making meanings therefore drives linguistic description to become ever more detailed.

The linguist's task in such explorations is then one of specifying as accurately as possible mutual expectancy relationships between features of the social situation and *identifiable* linguistic features of the language occurring in that situation. This has always been a central goal of systemic-functional linguistics and has played a significant role in shaping the theory overall. It will also play a central role in our move to multimodal genre. Since we will need to make the features by which we can define genres as 'identifiable' as possible, we will need a descriptive apparatus that is sufficiently sensitive to disentangle alternatives—just as is the case with grammar and other areas of linguistic analysis.

The final move to genre within the SFL account was taken as follows. Whereas register offers a rather 'homogeneous' view of text and the linguistic options that texts take up in context, it is clear that actual language use displays more structure: when we describe the linguistic details of texts in close detail, it is rare that an entire text exhibits precisely the *same* range of stylistic options. More often we can locate particular phases or segments of a text showing a relatively homogenous range of stylistic options and other segments of the same text that show different options being taken up. Therefore, a single linguistic text, or linguistic event, may appeal to several distinct registers while it is unfolding and yet still be seen as a coherent example of a single 'type' of text. It is this feature of texts that has presented, and continues to present, the greatest problems for approaches to register, text types and style. Since texts need not be homogeneous, simple 'labels' for registers or genres are rarely appropriate. We will see this below in current work on classifying web-based genres, where the problem is made considerably worse by multimodality.

The theoretical construct that was introduced in systemic-functional linguistics to deal with this 'higher' level structuring and organisational integrity across stylistically distinct textual contributions was genre. Within this view, texts and communicative events belonging to some genre are seen as staged sequences of purposive actions, where each stage can take on a distinctive register. This established a fundamental link between observable patterning of collections of linguistic features, or styles, and patterns in the social activity unfolding: distinct stages of a social activity (e.g., selecting an item in a market or shop, fixing the price, etc.) co-occur with and are signalled by particular 'styles' of language use.

The first explicit representation of genre in this sense within the systemic-functional model was Hasan's (1978) *Generic Structure Potential*, or GSP. In this account, a particular *contextual configuration* gives rise to a particular text structure, whose individual elements manifest specific linguistic properties. The text structure is defined by the particular GSP associated with the context. This specifies precisely which structural elements are necessary for a text to be considered to be an example of a particular genre, which elements are possible but not obligatory, the relative ordering of these elements (preceding, following, interspersed), the particular *semantic functions* to be achieved by each element and, last but by no means least, the corresponding *linguistic features* of semantics, grammar and lexis that may be used for those functions' achievement.

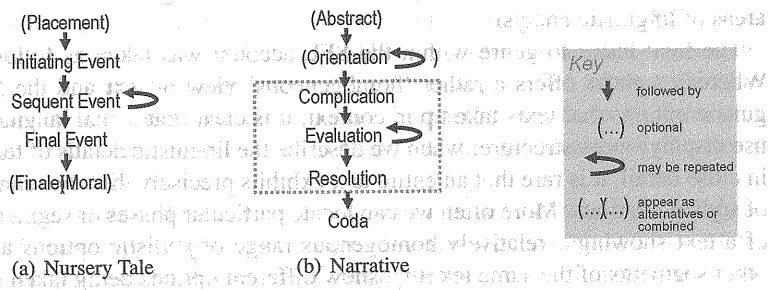


Figure 5.3 Graphical representation of the genre structures of Hasan (1984) for nursery tales and of Labov and Waletzky (1978) for narrative

One of the first examples of a characterisation of a genre developed by Hasan within this framework was that of the 'nursery tale' (Hasan 1984), the GSP of which is shown in Figure 5.3(a). Here we can see how this genre 'predicts' that recognisable elements will occur in a particular sequence, that one of these (the placement) may be omitted, that one (the sequent event) may be repeated and that the finale and the moral may be mixed or interspersed with one another.[3]

Hasan's discussion of the nursery tale made contact with several other approaches to textual organisation being pursued at that time, most significantly Labov and Waletzky's (1978) work on the structure of narratives. Figure 5.3(b) shows the narrative structure they propose in the same format as used for Hasan's GSP. In both approaches we see clearly the view of

[3] Analogies and forerunners of this account can accordingly be found in, for example, Propp (1968), Mitchell (1957) and others. We will not pursue the historical development further here; for more detail, see Martin and Rose (2007).

genres as "staged, goal-oriented social processes" that has subsequently become a cornerstone of approaches to genre in general (Swales 1990, Martin 1992, 1993).

Given the abstractness of the stages proposed within generic structures, it is understandable that much of Hasan's (1984) discussion actually centres on the *motivation* for distinguishing particular elements from others. That is, it is important—particularly from a linguistic perspective—to be able to find *evidence* that one particular structure holds rather than another; it is not sufficient to simply assert that some interpretation is 'plausible' or apparently the case. Hasan, Labov and Waletsky, and others approaching the issue of text description linguistically were therefore obligated to find linguistic evidence for the distinctions they were proposing. This then becomes the main criterion for establishing distinct categories. If, for example, a genre structure proposes that there is an orientation element in stories, then there should be linguistic features of grammar and lexis that reliably distinguish this element from, for example, the elements of the complication of those stories (cf. Figure 5.3).

It is this acceptance that distinctions may only be proposed if there is sufficient evidence *on the basis of linguistic patterning* for the presence of distinct elements that, on the one hand, places the investigation of genres and genre structures on a firm empirical foundation and, on the other, distinguishes the linguistically-motivated approach to genres from other, more interpretative approaches. It is also the step that is crucial for us here. When we move to multimodal documents, we also need to be able to find concretely identifiable empirical evidence to motivate particular structures and interpretations rather than others. Only then do we have a foundation sufficiently firm for further theory building.

Theory building also needs to proceed with respect to a sufficiently broad range of data. In linguistic genre analysis this has already proved its worth—just as it has in other areas of linguistics (and science at large). Martin and Plum (1997), for example, were able to develop a finer description of narrative genres by teasing apart distinct generic structures on the basis of more precise characterisations of the linguistic features found across genre stages. By close analysis, they discovered that the 'central' portion of narratives—shown within a dotted box in Figure 5.3(b)—manifests distinctive linguistic variation corresponding to several distinguishable *subgenres*. Not all of these are appropriately termed 'narratives' in the traditional sense, although they are nevertheless descriptions of chronological series of events. These more specific genres include *recounts*, which simply relate what occurred, *procedures*, which give somewhat anonymous instructions for activities to be carried out,

anecdotes, which tell a story to make a point, as well as the classical *narratives* themselves, which need not only to tell what happened but also to make it interesting—typically by setting up a complication which needs to be resolved. Each of these stages calls for *distinct collections of linguistic features* that allows it to be recognised and distinguished from the others.

Adopting an empirical foundation for genre investigation has therefore opened up the way for significant growth in our knowledge about genres and their structures. The resulting *genre families*, collections of more or less closely related genres, offer a clear illustration of the importance of pursuing fine-grained and empirically motivated analytic detail (cf. Christie and Martin 1997). But this is only possible when sufficiently accurate analytic tools are applied.

5.1.2 Genre as social action

Whereas the view of genre within the socio-semiotic perspective was anchored in the systemic-functional linguistic tradition, with its origins in anthropology and functional linguistics, the view found in the other approach to genre we discuss developed within a quite different intellectual climate. Within North American linguistics, the Chomskyan paradigm had already driven approaches to language as a social phenomena outside of the 'mainstream' to surface in parallel strands such as ethnomethodology (Garfinkel 1967) and work such as that of Hymes (Hymes 1974), both of which had strong roots in sociology and sociolinguistics. Interest in genre at that time was therefore not primarily a 'core' linguistic one.

Outside of linguistics, however, active traditions from literary theory and rhetorical studies contributed to the formation of the *New Rhetoric* (cf. Freedman and Medway 1994), which provided a setting within which explicit treatments of genre could grow. Here, language events were most readily seen as instantiating particular *social activity types* and these were then later assimilated to the notion of genre. Detailed linguistic description of the kind pursued for grammar and semantics was not a central part of this work and so remained very much on the edge of the account as genre studies in this tradition gained momentum.

Researchers in the North American and New Rhetoric traditions of genre generally cite Miller (1984) as the crystallisation point that brought notions of genre and social action together. Miller argued that those disciplines where the term 'genre' had traditionally been employed had not yet produced usable definitions. Many of the issues contributing to this problem continue in genre theory to the present day, and include concerns over the

nature of taxonomising and the value (or lack thereof) of relating texts to fixed and allegedly over-rigid notions of some ideal form. Miller's position on these issues, and on the value of genre as such, is that:

> "an understanding of genre can help account for the way we encounter, interpret, react to, and create particular texts. ... [G]enre study is valuable not because it might permit the creation of some kind of taxonomy, but because it emphasizes some social and historical aspects of rhetoric that other perspectives do not."
> (Miller 1984, p151)

To further these aims Miller takes the vital step that is incorporated in all those approaches that build on her work: i.e., that a sound definition of genre needs to rely not on the form of a discourse but on the social action that is being achieved through the use of that discourse. Recurring social situations give rise to recurring problems that can be approached through recurrent communicative solutions. These recurrent communicative solutions can then themselves be recognised as such by virtue of their recurrent form—but it is the social action that drives the account.

A description of genre in Miller's terms then picks out one particular dimension for discourse classification:

> "a classification based in rhetorical practice and consequently open rather than closed and organized around situated actions (that is, pragmatic, rather than syntactic or semantic)."
> (Miller 1984, p155)

Moreover, as a signpost for which practices will be found, she takes the 'de facto' genres that are already named and recognised in our societies as a good indication of just what genres will need to be described. A position also taken up by Swales (1990) in his standard introduction to the field.

This view of genre is one way of taking the actual practices and classifications of those using genres seriously. Genres are related to situations but those situations are themselves abstractions; they are social constructs that are maintained and created through action, including the genres that accompany them. A discourse of a particular genre is not then simply a reflection of an objectively defined situation, but is instead one of the ways in which an abstract social type of situation is signalled and maintained through communicative acts. Here Miller draws on other approaches that emphasise the constitutive role of language use, including ethnomethodology (Garfinkel 1967) and Halliday (e.g., Halliday 1978), allowing genre to take its place within a hierarchy of kinds of meaning, ranging from behaviour at one end through to culture and ideology at the other (cf. Miller 1984, p162).

Defining genre in this way allows Miller to set out some criteria by which classes of discourse can be accepted as genres or not. They will clearly

tend to show similarity of form, but most importantly they also have to function as bearers of meanings for the culture they are embedded within. This is similar to the relation seen by Martin (1992) between genre and ideology in the socio-semiotic tradition: different cultures will activate or use repertoires of different genres. Moreover, Miller holds that such entities do not support a treatment in terms of taxonomies, because genres "change, evolve, and decay" (p163).

A further move in this tradition relevant for the view of genre under development here is that of Bazerman (1994). Bazerman draws on the connection between social organisation and distinctive styles of discourse in order to argue for entire collections of genres that stand in systematic relationships. He characterises genres as actions, adopting further aspects of speech act theory, and argues that certain kinds of texts, particularly those that are intended to fulfill particular functions (such as filing a patent, which he discusses as one detailed example of genre change over time) can be considered as complex speech acts, quite analogous to speech acts formulated within single sentences.

Then, just as utterances rarely, if ever, occur in a vacuum and are generally components of larger scale 'conversations', Bazerman sees certain genres similarly. A text in a particular genre may require texts to have preceded it in other genres and certain other texts to follow: for example, court rulings, reports and legal decisions may each belong to their own genres but it is the entire ensemble of interlinked generic acts that make the social subsystem involved work. Seen more generally, Bazerman therefore invites an interpretation of society as a whole as a *system of genres*. The genres available at particular points in the network define the options for behaviour that a society provides for its members. Such larger groupings of genres are now receiving study within a variety of frameworks; they are termed *macro-genres* by Martin (e.g., Martin 1994, 2002), and *genre chains* and *genre networks* by Swales (2004, pp18–25).

Finally, the development arising out of the North American genre-as-action context that is probably the most significant for multimodal theorising at present is the research programme initiated by Yates and Orlikowski (1992). Yates and Orlikowski propose genre as a powerful means of investigating organisational practice and, in particular, *changes* in organisational practice over time. This approach has proved particularly influential for work on the genres of the new media, web-based genres and so on. Yates and Orlikowski begin by emphasising the essential role that communication plays in all organisations. Communicative genres, and their structuring and inter-relationships, then offer an incisive method for characterising organisations and their structure.

This framework represents a fruitful combination of several strands of research including, on the one hand, approaches to genre from rhetoric (cf. Simon 1978, Harrell and Linkugel 1978 and Chapter 4) and, on the other, the sociologically sophisticated model of *structuration* developed by Giddens (1984). Miller, as we have seen, had already defined genres as "typified actions based in recurrent situations" (Miller 1984, p159) and had moved 'situation' to be an essentially socially constructed, and hence semiotic, entity. This, taken together with structuration, provides a particularly important role for genre to play. Under this account, genre is understood in terms of a dual structure where genre both structures the stream of social practices and is shaped by that stream.

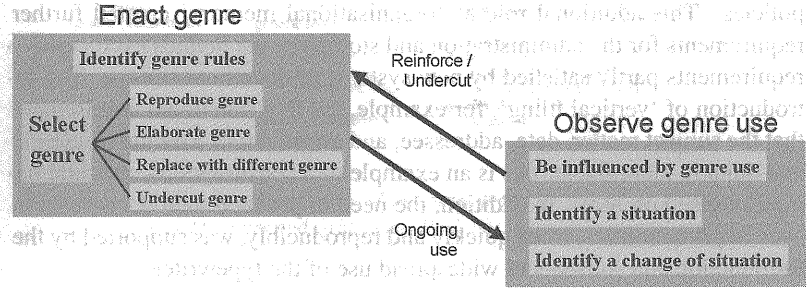


Figure 5.4 Characterisation of the cyclic process of genre change according to Yoshioka and Herman (2000); used by permission

The development of organisations over time is then addressed by Yates and Orlikowski by charting how the *genres* of those organisations change. Social practices in an organisation are given structure by the genre repertoires available to the organisation (Orlikowski and Yates 1994), and so changes in those genres also correspond to changes in those organisations' practices and structure. In other words, using a genre to structure a communication serves both to *reproduce* the genre and, whenever changes are introduced into the communication, to *modify* the genre.

This has led both to detailed analyses of changing document genres and to theories of just how such genre change occurs. One general model of the process is set out in Yoshioka, Herman, Yates and Orlikowski (2001), summarised graphically in Figure 5.4. Here we see how genres are selected and used, which can lead to a situation being identified and, conversely, how changes in situations can lead back to motivate elaborations, changes or even rejections of the originally associated genres. The model also corresponds well to the view of change in the socio-semiotic perspective: the

system provides the possibilities for instantiation that may be taken up and each instance of language produced in turn feeds back to either maintain or alter the system (Lemke 1988*a*, 1993, Kress 1993).

Yates and Orlikowski (1992) provide a detailed illustration of their approach to genre change over time drawing on the *business memo*. This distinctive form appears to have emerged as a progressive response to the need to manage intra-organisational communication and memory as companies grew beyond small groups with small numbers of management hierarchies. At first, the already existing genre of business letters was employed for this situation; this was then adapted by a reduction in formal language and other linguistic markers of a greater shared context of interpretation. Memos also began serving more as records of company-internal decision making and policies. This additional role as 'organisational memory' created further requirements for the administration and storage of the memos exchanged—requirements partly satisfied by new systems of document storage. The introduction of 'vertical filing', for example, led to a functional requirement that the subject matter, date, addressee, and originator of a memo be clearly and quickly identifiable; this is an example of a consumption constraint being exercised on form. In addition, the need to produce such documents in ever larger quantities, more quickly and reproducibly, was supported by the introduction and subsequent widespread use of the typewriter.

This technological change established production constraints that also served to shape the forms of the emerging memo genre. Underlining and capitals supported the use of subheadings for structural layout in memos and, subsequently, the introduction of tab stops further supported and encouraged the use of tabular information displays (Yates and Orlikowski 1992, p314). By 1920 this form was then established as a virtual canvas in its own right, one which continues in use to this day—perhaps in vestigial form even in the 'one-line table' discussed for our original Gannet example in Chapter 2.

Yates and Orlikowski go on to show how the memo form has contributed significant aspects of the 'email' genre as well. In the context of electronic mail systems:

> "System designers embedded the structural features of the memo heading into the new medium. In this case, computers rather than people routed the messages, so the fields of the memo heading were designed to be readable by computers (as well as humans). A typical memo layout for the fields was not required by computers, so its widespread adoption shows that designers (whether implicitly or explicitly) retained elements of an existing and familiar genre in moving to a new medium." (Yates and Orlikowski 1992, p316)

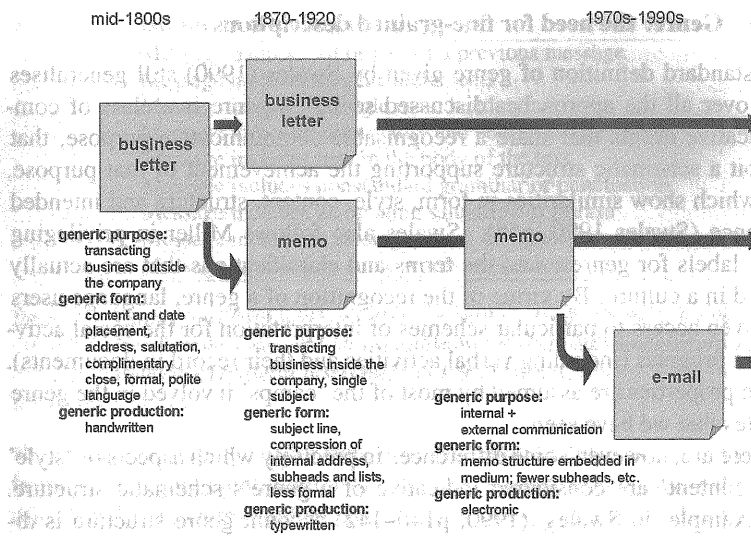


Figure 5.5 Graphical representation of the development of the memo genre as discussed by Yates and Orlikowski (1992, p315)

This is another good example of the power exercised by the virtual canvas.

We summarise Yates and Orlikowski's discussion graphically in Figure 5.5. This suggests more readily the way that genres split and grow and also includes reference to the production modes employed in the genres. All of the genres involved here continue to develop; indeed, in certain respects the email genre is already dividing in numerous ways, ranging from the (almost) recognisable genres of spam messages to variants that are encroaching on the traditional preserves of the non-digital business letter.[4]

As we will suggest in detail with respect to our own examples below, changes such as those reported here for the business memo can be tracked even more effectively when we have a sufficiently detailed characterisation of the features that will change. We see the distinct layers of description given by the GeM model and the correlations that we can set up between these layers for particular genres as providing a particularly good foundation for such explorations.

[4] Strengthened in some places by legislation: for example, there is a European Union regulation that all business communication, including emails, should carry appropriate legal information concerning the contact details of the originating business, who is legally 'in charge', and so on. This regulation is gradually finding its way into corresponding national legislation.

5.1.3 Genre: the need for fine-grained descriptions

The standard definition of genre given by Swales (1990) still generalises well over all the approaches discussed so far: a genre is a class of communicative events that share a recognisable communicative purpose, that exhibit a schematic structure supporting the achievement of that purpose, and which show similarities in form, style, content, structure and intended audience (Swales 1990, p58). Swales also follows Miller in privileging 'folk' labels for genres: i.e., the terms and classifications that are actually named in a culture. By virtue of the recognition of a genre, language users are given access to particular schemes of interpretation for the social activities in progress (including verbal activities and their record in documents). These properties are assumed by most of the 'camps' involved in the genre debates that we have seen.

There are, however, some differences in precisely which aspects of 'style' and 'content' are considered indicative of a genre's schematic structure. For example, in Swales' (1990, p140–142) account genre structure is divided into 'moves' and 'steps' and is closely associated with broad 'rhetorical strategies', analogous to those found in traditional rhetoric. A similar relation between genre moves, or stages, and rhetorical organisation is proposed by both Lemke (1988*b*) and Thibault (1990, pp106–108). In other approaches, for example in Hasan's model of generic structure and approaches derived from it, there is a particularly strong expectation that generic stages will correlate with fine-grained lexical, grammatical and semantic differences. For both rhetorical organisation and notions of 'style' and 'content' in terms of linguistic features, generic phases are recognised by finding segments of text over which the variation of features selected is significantly less than that theoretically possible. Constraint on variation is then considered as a significant symptom for an associated genre phase.

The kind of information appealed to in approaches of the North American school appears, in contrast to this, to be more linguistically *ad hoc* or opportunistically selected. This can be revealing of distinct genres but is not necessarily so. As an example, the 'form criteria' established in order to find genre distinctions and developments in one investigation of the use of email within a 'virtual community' reported by Orlikowski and Yates (1994) show rather little systematic organisation; a selection of these features is given in Table 5.1. Although, on the positive side, we can already see features relevant for a multimodal view of the documents analysed, rather more problematic is the loose, pre-theoretical nature of these features. In particular, the characterisation of *linguistic* detail that is assumed in linguistic-based analyses of genre is not at all present. As mentioned

Definition of coding category
Message includes all or part of a previous message
Message includes graphical elements
Message includes a single main heading
Message indicates informality and colloquialism
Message includes lists in the body of the text
Message includes nonstandard grammar or punctuation
Message includes an opening salutation or phrase
Message includes a closing remark or signature
Message emphasises some words or phrases

Table 5.1 Extract from Orlikowski and Yates's definition of coding criteria for their investigation of genres deployed within an extended email-mediated organisational effort (Orlikowski and Yates 1994, p552).

above, incorporating broader ranges of linguistic features has already been found to aid genre identification significantly—even largely automatic identification (cf. Kessler, Nunberg and Schütze 1997)—and so its absence here can only weaken the accounts possible.

We see this as a major contributor to some of the problems that Yates and Orlikowski raise with respect to the degree of detail desirable in genre descriptions:

> "undue proliferation of genres may also weaken the usefulness of the concept. ...If ...such cataloguing were extended to every industry, it would result in endless lists of genre comparable to the exhaustive (and exhausting) model letter books common in the 19th century. [] Thus there is a tension between too broad and too narrow a definition of genre." (Yates and Orlikowski 1992, p305)

We do not accept this tension at all; the situation is actually the same for any linguistic analysis. The question is simply one of how fine an analysis needs to be for a specific analytic purpose. This decision should rest with the analyst and not be imposed *a priori* by the *framework*: the framework employed should be able to support analytical resolution of any granularity that is appropriate for the kinds of research questions raised. We should not, therefore, build into the definition of genre itself a preference for, or an avoidance of, particular resolutions.

Crucial for an effective answer to the problem is the capability to define whether and in what ways one genre might be more 'specific' than another. The experiences of linguistically based investigations of genre lead us to expect that the more detail that is available to us, the more refined our views of individual genres and their interrelationships will become. The discovery of genre families mentioned above is a clear example of how this can work to good effect. Drawing on such detail is one way by which we can avoid

our study of genre collapsing into the simple catalogues of genres that Yates and Orlikowski warn against.

The serious attention given to the value of linguistic form found in strongly linguistic-based approaches is then not equally available to the accounts deriving from the North American School of genre. But we consider this to be crucial for turning genre into a usable construct. Fine linguistic detail is a prerequisite for fine-grained genre classification since only then do we achieve sufficient detail (i) to allow predictions to be made and (ii) to reveal more genres than superficially available by inspection or folk-labelling within a given discourse community. When we turn to the even less well understood areas involved in multimodal genre, a fine-grained specification employing a greater degree of linguistic sophistication and systematicity in the *kinds of forms* that can be used for evidence for or against the recognition of a genre category is even more important.

5.2 The move to multimodal genre

The previous section has presented how genre has been developed as a theoretical notion for analysing text. We argued that the most important theoretical property that makes an account of genre usable for empirical study is its potential to function *predictively*. Knowing the genre to which a text belongs leads to particular predictions concerning the form, function and contents of that text. This link between form and genre is a crucial feature that distinguishes linguistically-motivated approaches to genre from non-linguistic approaches. In this section, we discuss moves that have been made to extend existing accounts to include *multimodal* configurations in addition to texts alone. In taking this move, maintaining and extending the ability to function predictively is a, perhaps the, major challenge.

We will also address approaches to multimodal genre that have not been concerned primarily with text. This includes notions of document types from design, of particular classes of document based on visual organisational features, as well as results that have been obtained in automating document class recognition in the document analysis community. We will argue that these diverse orientations to genre usefully converge in, and can be improved by, the theoretical account of genre we are pursuing.

The section concludes with a brief discussion of one current area of concern in multimodal genre where a lack of an appropriate theoretical foundation is leading systematically to problems of application: this is the area of the so-called *cybergenre* (cf. Shepherd and Watters 1998), which is currently receiving considerable, and increasing, attention. We will argue that

current efforts to characterise the kinds of documents found on the World-Wide Web are seriously handicapped by a relatively simple notion of genre that has only been extended minimally from traditional, non-multimodal conceptions of register and style.

5.2.1 Multimodal moves within linguistic and rhetorical approaches to genre

To the extent that the approaches to genre that we have discussed so far have addressed phenomena of multimodality, we will see in this subsection the problems that such extensions cause. This will allow us to set out more clearly just where a more developed treatment of multimodal genre has to start. As above, we begin with the systemic-functional approach and then move on to the social action perspective.

We have already discussed several applications of the systemic-functional linguistic approach to multimodal documents in previous chapters and so it will come as no surprise to see that there have been similar attempts to extend the systemic-functional notion of genre to deal with multimodality also. This development is one specialisation of the general assumption that is always made when extending SFL multimodally—i.e., that some artefact is subject to similar functional pressures to those that affect language and so will exhibit similar organisational properties (cf. Kress et al. 2000, p44). This applies equally to genre and its use multimodally.

We have seen in the previous sections, for example, that SFL attributes genres with a very particular social significance. As Kress emphasises in his account, the term *genre*

> "is best used to describe one aspect of textual organisation, namely that which realises and allows us to understand the social relations of the participants in the making, the reception and the reading/interpretation of the text."
>
> (Kress 2003, p96)

Kress and van Leeuwen then claim that since both verbal and non-verbal artefacts—and, indeed, social practices—can be indicative of social relations, then genre must also be a useful category to apply regardless of mode.

SFL approaches to multimodal genre have also focused on issues of change over time. One detailed example of the combination of issues of change and issues of multimodal genre is Baldry's (2000) analysis of two 'genres' of scientific text over the past 100 years: texts in plant biology and in economics. Baldry suggests that combining the contribution of various modalities can be seen as a *social achievement* of the relatively

understand users' interpretations of the documents they encounter. The importance of such *generic expectations* brought about by convention when dealing with multimodal artefacts cannot be over-estimated.

Proposals from automatic analysis

Arising out of the school of thought mentioned above that genre is not amenable to precise formalisation, it is sometimes suggested that it makes limited sense to attempt detailed and explicit characterisations of genres, particularly multimodal documents. But, in fact, there is good evidence that substantial detail can be extracted from such artefacts—and this can be used to drive considerably more refined genre characterisations than are currently available.

Within the automatic document analysis tradition, for example, the task of classifying documents according to their visual appearance in order to assign them to particular categories has received close attention for many years. This research is relevant for us here because significant progress has been made in 'clustering' visually related documents together into families of similar documents. Such families consist of documents exhibiting similarities of 'style', both linguistically and visually.

With respect to linguistic similarity, statistical techniques of various kinds have now been applied with considerable success. This is closely related to research into linguistic register as introduced above (cf. Biber 1988), extending this to apply to linguistic genre classification (cf. Kessler et al. 1997). There is, however, little differentiation here of 'style', 'register' and 'genre'—the main focus is on providing labels for families of similar documents. The term 'genre' is used when those labels correspond to existing categories such as 'legal text', 'fictional', 'editorial' and so on.

Similar techniques have now also been applied to visual similarity (e.g., Hu, Kashi and Wilfong 1999, Cesarini et al. 2001, Diligenti, Frasconi and Gori 2003, Mao et al. 2005). Approaches here take data consisting of a collection of layout structures, often based on some variant of the XY-trees that we saw in Section 2.4, and organise that collection into specified classes according to the closeness of the structural relations exhibited. Members of the same class are considered to be visually similar; members of distinct classes are visually dissimilar. The fact that some good quality (i.e., reliable) clustering results are being achieved indicates that there is indeed a considerable degree of visual generalization to be drawn on. Although a visual difference *per se* may or may not correspond to a genre difference, the ability to find visual similarities is already a good first step. Considerably more work will need to be carried out to ascertain just how visual similarity and judgements of genre membership relate. This will no doubt feed in future into more developed accounts of multimodal genre.

We see very similar results and methods being applied in the context of automatic recognition and classification of documents found on the World-Wide Web. Doermann et al. (1997) argue that classifying web documents according to their basic intended function will provide considerable benefits. They accordingly propose several broad categories of distinct web-based families of documents that we can consider as genre 'dimensions' or facets (see below). For example, one distinction involves a separation of documents according to 'use' and to 'type'. Use-related categories include documents for reading, documents for browsing, and documents for searching. Each of these can be correlated with particular indicative sets of document features; for example:

> "In a searching document, no more than 25% of the text blocks should have more than five lines. There should be no image components, and few or no graphic components."
>
> (Doermann et al. 1997, p1080)

Collections of descriptive criteria of this kind show a broad similarity with notions of genre, serving to associate realisational features of documents with generic purpose. We will see further applications of this kind of approach for web-page classification in the section following.

Doermann et al. (1997, p1080) go on to observe, however, that criteria of this kind cannot be expected to perform well on documents with complex structures. This is again strikingly similar to the state of affairs in linguistic approaches to classifying text types. Attempts to find 'dominant' text functions to generate predictions about the features of a text have not in general been successful because texts are diverse and can be satisfying multiple functions. This is one of the primary motivations for applying a more differentiating view of genre and its realisation—i.e., accepting that texts, and particularly multimodal texts, may exhibit a range of distinct styles across the structural elements that they involve.

There are accordingly proposals in this tradition for processing documents in order to capture their constitutive layout elements (cf. Chapter 2). This can then be taken further by assigning those layout elements distinguishing *functional* labels. Doermann et al., for example, define:

> "...a level of document organization, which can be regarded as intermediate between the geometric and semantic levels, that relates to the efficiency with which the document transfers its information to the reader. We refer to this level as the *functional* level."
>
> (Doermann et al. 1997, p1078)

This additional layer characterises the particular communicative roles played by the logical structure elements of a document's pages and so is already very similar to the *generic stages* defined by Hasan's Generic Structure Potentials and related approaches (cf. Figure 5.3 above).

Just as is the case with generic stages, this assignment of functional value also depends crucially on the type of document analysed. For example, we might be able to recognise some short block of text separated by whitespace from its surroundings, identifying many features of the block purely on the basis of its appearance. In the context of the page as a whole, however, that text block will take on further particular roles—such as the *caption* of a figure, or the *by-line* of a news article, or a *label* (or *callout*) in a diagram. Allocating these functional values requires knowledge not only of the geometric and physical properties of a block but also knowledge of the document type, i.e., the genre. Given such information a system may then predict what kinds of functional parts are expected.

Eglin and Bres (2004) formulate a model of functional document analysis for assigning functional labels of this kind automatically. In this account, documents are first decomposed into textual and non-textual elements using filters of various kinds as described in Chapter 2. A range of features are extracted from the visual properties of these elements, including visibility, salience, and position with respect to other elements. This information is taken as the basis for functional labelling. The labels themselves are drawn from a classification hierarchy of families of related functionalities. The main divisions of this classification define three families: titles, intermediate, and paragraphs. A subsequent process then attempts to provide hierarchical structure for the elements discovered, and this can in turn help to further constrain functional labelling and to correct misclassifications.

Eglin and Bres have applied the framework to a range of documents, including newspapers. We show in Figure 5.7, therefore, how our example newspaper page from above might appear if analysed according to this framework. Although the classification is not yet fine enough to reach the standard terms and categories used in newspaper design, the similarities that this functional analysis shows to Hasan and others' generic stages is already suggestive.

Characterising both generic structures and the particular constraints holding on elements of those structures is certain to become increasingly important for all kinds of genre recognition—both of traditional print documents and in new media. Among these, web-pages in particular form a challenging area where a strong and well theorised notion of genre is critical. In the subsection following, we suggest that many problems currently discussed with respect to web-based genres in fact arise from weaknesses in the theoretical foundation adopted for genre rather than in the complexity of the documents being analysed. Most problematic is the tendency not to draw on those properties of genre central to its definition—particularly in relation to structure—and to collapse 'genre', 'register', 'style' and similar notions together.

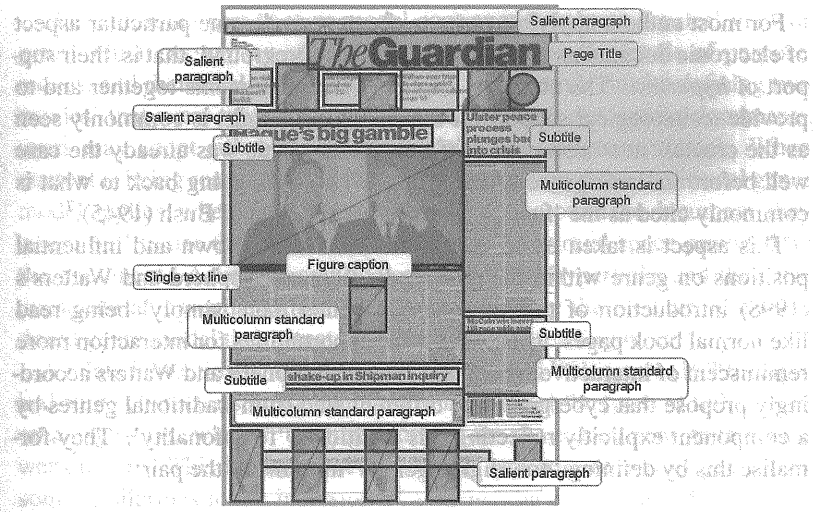


Figure 5.7 Constructed analysis of an example newspaper page following the style of description of Eglin and Bres (2004)

5.2.3 Cybergenres: a brief critique

Several research communities are currently applying notions of genre in the context of the World-Wide Web. In fact, the treatment of web documents is currently one of the fastest moving areas of application for notions of genre that we can find—presumably motivated by the increasingly dominant role that web-based information has taken on in modern life. For approaches to genre that draw on linguistics, this is a natural development: whether the artefacts analysed are on the web or printed is of little concern, one would expect the same analytic frameworks to find broadly similar application. For other approaches, the technological basis of the web has been a major driving force and, as we saw in the previous section, practical issues of web-page classification, clustering for information delivery and retrieval, user evaluation and automatic information extraction play more of a defining role.

For most authors applying genre to the new media one particular aspect of electronic documents is placed firmly in the foreground: that is, their support of *hypertext*. The ability to explicitly link documents together and to provide technological support for following those links is commonly seen as the crucial 'new' component of new media. This was already the case well before the emergence of the World-Wide Web, going back to what is commonly cited as the 'first' work on hypertext, that of Bush (1945).[7]

This aspect is taken up in one of the most well-known and influential positions on genre within the new media, that of Shepherd and Watters's (1998) introduction of the *cybergenre*. Rather than 'simply' being read like normal book pages, web-pages offer opportunities for interaction more reminiscent of interactive computer systems. Shepherd and Watters accordingly propose that cybergenres be differentiated from traditional genres by a component explicitly reflecting this additional 'functionality'. They formalise this by defining 'non-digital genres' in terms of the pair:

$$< \text{content, form} >$$

in contrast to cybergenres, which they define with the triple:

$$< \text{content, form, functionality} >$$

They thus build into their account a distinctive and defining role for the *medium* of the documents encountered, providing at the same time a place for characterisations of a user's *interaction* with the web document.

A number of developments of this kind have now been made. For example, a similar line of argument follows in Askehave and Nielsen's (2005) proposals for extending the notion of genre developed by Swales (1990) to new media by incorporation of a sensitivity to the medium of the web. They propose what they describe as a 'two-dimensional' genre model, in which the generic properties of a web-page are characterised both in terms of a traditional *text* perspective and in terms of the *medium* (including navigation). They motivate this by means of a discussion of a proposed 'homepage' web-genre thus:

> "This duality inherent in homepages may best be described by conceptualising the homepage as a front door with a door sign. The door sign indicates the name of the residents (i.e., the 'content' of the house) while the door itself is the gateway (the medium) which enables guests to enter the house and visit the residents inside." (Askehave and Nielsen 2005)

[7] We will not follow up this history of the medium in any detail here; for further information and many references, see Conklin (1987), Landow (1991, 1994, 1997).

The traditional part of their model continues to rely on Swales' view of genre which, as we saw above, analyses genres at the levels of purpose, moves, and rhetorical strategies. The new part extends this by defining two 'modes' that users take up in their interaction with new media documents: users may adopt either a 'reading mode' or a 'navigation mode'. When content is being read, readers 'zoom in' on content and the traditional genre model applies; when, however, users are moving around the document, they 'zoom out' to use the medium to navigate in the virtual space provided by the document. Askehave and Nielsen argue that hyperlinks and their use constitute an essential extension brought about by the medium; the traditional genre model they situate solely within the reading mode.

A related direction is followed in Crowston and Williams's (1999) empirical study characterising distinct genres in terms of the types of links found on a page—i.e., whether the links point to the same document, to the same website, etc. This inclusion of the linking structure of web-pages shows some similarities to the inclusion of a navigation and 'access' layer in our framework. We do not believe, however, that this is usefully restricted to be solely a feature of web-pages. Indeed, accepting a defining role for medium among the constituents of genre has traditionally been controversial (cf. Yates, Orlikowski and Okamura 1999, p100).

Medium is most commonly associated with aspects of form and Crowston and Williams (1997, 2000) argue, following Yates and Orlikowski, that it is the communicative function and intent that is critical for defining genres, not the 'physical form' of the supporting artefact. Thus, whether a document appears as a 'brochure, booklet or a flyer' should not necessarily be indicative of different genres.[8] But the usage of the term 'genre' nevertheless shows extreme variation. Some researchers even talk of photographs and diagrams as being different 'genres', which goes to an opposite extreme and makes medium criterial.

The precise relationship between physical form, the affordances of that form, and genre is therefore complex. There are many cases where the physical form provides or prevents particular deployments of resources, visual or verbal, from being effective: and in precisely these cases the physical form and a genre definition will need to interact. Moreover, although the tendency to avoid questions of medium has certainly been encouraged

[8] In the list of web genres found in Crowston and Williams' study, their description of the characteristics of those genres is accordingly given predominantly in terms of function and purpose (e.g., *meeting minutes*: "the record of the proceedings at a meeting of an assembly, corporate body, society, company, committee, or the like"); this stands somewhat in conflict to their observation that they could often assign genres for web-pages in languages that they did not themselves speak (§ 3.2): the visual side is still clearly under-represented therefore.

by monomodal considerations of text, this is far from the case in socio-semiotic constructions of genre such as those offered by Kress and van Leeuwen (2001) and Kress (2003), where artefactual properties are considered crucial. Significantly, this development has again been proposed for *all* documents, not just those of the 'new' media, and so it is beneficial to return to a consideration of this, rather broader, starting point when theorising web-based genres also.

Many studies, however, have already begun characterising the genres found in documents on the World-Wide Web, usually adopting informal labels for perceived document purposes. Crowston and Williams (1997) report a classification of 100 randomly selected web-pages in which each page was coded according to a hierarchically organised collection of genre labels. The acceptance of a hierarchy follows from Yates and Orlikowski's notions of genre 'specificity' mentioned above. The genre labels themselves were predominantly familiar from non-digital publications, with the addition of some characterised as novel, such as the *hotlist*, webserver statistics, and search engines. But of the pages sampled, 80 appeared to belong to familiar genres, leading Crowston and Williams to comment:

> "Perhaps our biggest surprise was just how mundane our sample was. The 100 pages in our pilot sample did not include anything particularly radical..."
> (Crowston and Williams 1997, § 6).

This reaffirms the proposals of Orlikowski and Yates (1994, p547) and others before them that, in novel situations, people adopt and adapt existing genres thereby reproducing them in the new situations of use. It is difficult, for several reasons, to move beyond the realms of established genres—even in the new media.

In a similar study, Shepherd and Watters (1999) characterised a random selection of 96 websites according to properties drawn from their content-form-functionality definition of genre. They used these properties to assign web-pages to a list of rather general genre labels, consisting of homepage (personal and corporate), brochure, resource, catalogue, search engine and game (cf. Table 5.2). After mapping Crowston and Williams' much finer set of genre categories to their own, they suggest that there has been a substantial change in the genres found on the web over the short time period spanned by the two studies: they remark on a significant increase in homepages and a corresponding decrease in 'resource' pages.

But the extent to which such claims are accurate or revealing is questionable. As we can see from Table 5.2, the actual attributes appealed to are very broad and may be subject to a range of variation without being strongly predictive of genre. When it is as straightforward to include a search facility on a web-page as to leave it out (as has almost become the case), we can expect such a functional element to appear whenever the author thinks of including it. Functionalities currently listed in the table as corresponding to specific cybergenres may then spread rapidly across genres, losing any predictive power they might have fleetingly enjoyed concerning genre membership.

Cybergenre	Content	Form	Functionality
Homepage	information about person or company	• introduction • hierarchical • images • animated images	• browsing • e-mail
Brochure	products and services	• shallow hierarchy • high-impact visual	• browsing • e-mail
Resource	subject-specific information	• hierarchical • images • video • audio	• browsing • e-mail • search • discussion • interaction
Catalogue	products and services	• hierarchical • images	• browse • e-mail ordering • search

Table 5.2 Selection of the results of Shepherd and Watters' analysis of 96 websites according to their content, form and functionality scheme

Moreover, for the majority of simple websites—such as those sampled both by Shepherd and Watters and by Crowston and Williams—the forms supported by the web are actually extremely limited. The constraints imposed by their use of straightforward web technology constitute a sufficiently strong straightjacket to keep attempts at innovation within familiar channels. The parallel with arguments made elsewhere concerning the consequences of introducing new technology is striking. With the emergence of print technology, for example, instead of immediately opening up a wealth of new possibilities as sometimes assumed, the immature technology actually served to *restrict* variability. Much previously existing use of illustration, colour and their integration with text then went missing until the technology had, in a sense, caught up again with the demands and challenges of design (cf. Twyman 1986 and Waller 1987*a*, pp248–251 for further discussion and references).

This should be taken as a word of caution for the currently exploding 'study of websites' being undertaken in multimodal linguistics: there are useful and interesting aspects of websites to investigate, but they are found exceedingly rarely directly on the surface in terms of novel design and new

genres. To what extent the web-page has moved beyond technological dependence to support semiotically interesting meaning-making possibilities is an open question, a question that is not addressed adequately simply by assuming it to be the case.

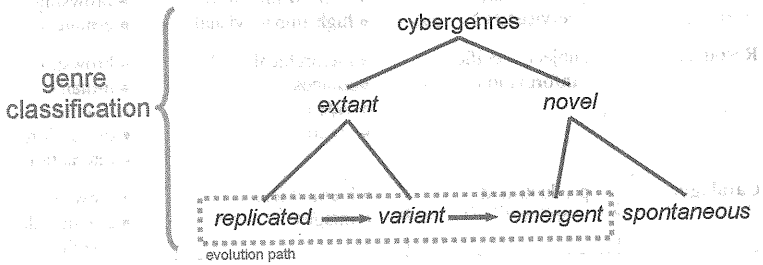


Figure 5.8 Cybergenres and their development according to Shepherd and Watters (1998)

Shepherd and Watters (1998) also propose an abstract characterisation of how genres on the World-Wide Web develop; this is shown in Figure 5.8. They define an 'evolution' path for cybergenres by which initially non-digital genres tend to be simply recreated in digital form: these are the *replicated* genres on the lower left of the diagram. These subsequently pick up additional functionalities supported by the web leading to *variants* on the original replication. This can then extend further to support completely new uses and new genres, described in the diagram as *emergent*. Alternatively, totally new genres can be produced spontaneously, supported by the new functionalities and very rapid distribution possibilities that the web opens up.

Although appealing, this characterisation does not yet do justice to the 'technological straightjacket' imposed by the new medium. It is far more difficult to simply 'replicate' existing genres than might have been thought. This fact contributes both to the dominance of 'familiar' (or *replicated*) genres on the World-Wide Web and to the claimed low degree of *variety* of genres on the web. The range of genres available is still being restricted because the technological limits channel documents into a narrower range of options than a technologically mature medium might support. This manner of genre 'coercion' needs to receive far closer attention. It is probably the most frequent style of 'adopting and adapting' genres that occurs during the emergence of web-based genres.

We can see this in the examples that Shepherd and Watters themselves offer. As illustrations of replicated documents they include early examples of 'electronic newspapers', which tended to be simply versions of print

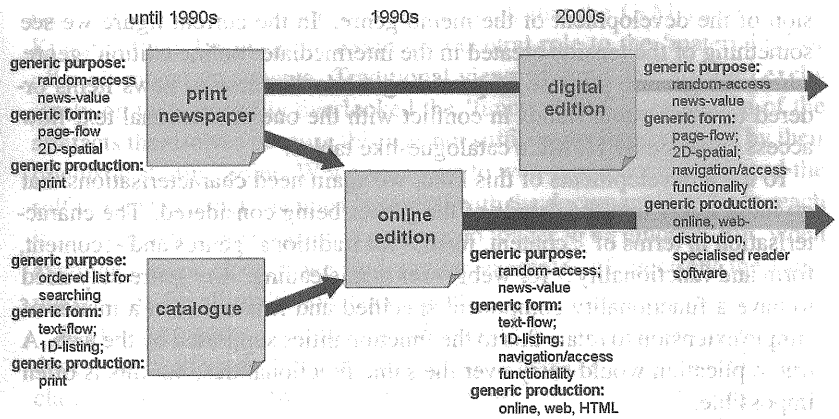


Figure 5.9 Development of parallel distinct online news genres

newspapers with extra facilities. However the documents that they refer to were actually experimental results from research projects in the early 1990s (cf. Ashton and Cruickshank 1993, Haake, Hüser and Reichenberger 1994), which were not at all the electronic newspapers that then arrived on the web *en masse* from newspaper companies. As we discussed above, these actual electronic newspapers were, and often remain, strongly influenced by the weak technological basis provided by the web and were not able to 'replicate' the original non-digital genres at all. It is only with the very much more recent *digital editions* alluded to above that replication has become remotely possible.

An arguably more accurate characterisation of the development process in this case, therefore, is one in which an originating document type is approximated applying a different technological and material basis. This difference may coerce the originating genre into another, more or less related, genre that preserves some aspects of the original while replacing others. The originating newspaper genre here was therefore coerced so that what was produced was not a newspaper but a 'catalogue' of news items with little use of spatial composition for organising meaning and access structure. Shepherd and Watters present the emergence of digital editions as a process of 'evolving back' to the print genre (§ 2.3); we would suggest instead that we have here the emergence of an *additional*, more accurate replication of the *originating* genre for which the supporting technology is only now becoming available. We show this graphically in Figure 5.9 in the style used in Figure 5.5 for our illustration of Yates and Orlikowski's (1992) discus-

sion of the development of the memo genre. In the current figure we see something of the tension created in the intermediate 'online edition' genre, where the generic purpose of providing random-access to news items ordered by news-value stands in conflict with the one-dimensional text-flow access enforced by HTML's catalogue-like tables.

To track developments of this kind, we again need characterisations that go more finely into the details of the genres being considered. The characterisation in terms of <content, form> for traditional genres and <content, form and functionality> for webgenres is misleading. *Any* genre may need to have a functionality component specified and it is not then a matter of simple extension to relate this to the functionalities supported by the web. A true replication would carry over the same functionalities, but this is often impossible.

Many kinds of 'faithful replication' of genres according to the schemes seen so far must therefore be seen as far from faithful. Simply preserving folk-labels across the genre coercion that has occurred can obscure important differences. We need instead to allocate genre on the basis of the documents themselves rather than on claims of their lineage. Only then can we see more clearly the contributions made when genres, and specific aspects of genres, combine. In terms of the metaphor used in the introduction to this book, a loose genre classification that privileges intended 'purpose' can easily swamp (for the analyst) the signals that the document itself might be sending. In Figure 5.10 we illustrate the difference in modelling approach that this entails between the cybergenre view and that presented here.

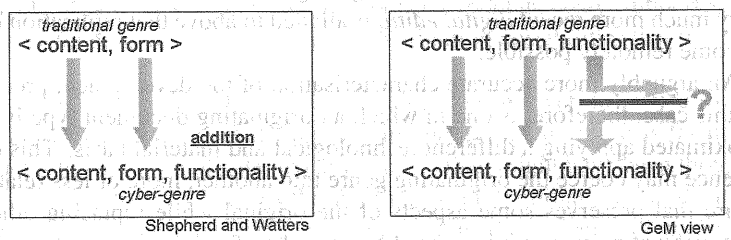


Figure 5.10 Contrasting Shepherd and Watters (1998) view of cybergenres with the approach proposed within the GeM model

There is very much more to be said concerning appropriate ways of approaching the issue of genres and the application of the construct to the digital context. Most important for our current purposes, however, is the consideration that a more appropriate definition of genre does *not* in fact open up a divide between 'digital' and 'non-digital' artefacts. We emphasised already in the introduction of the book that the GeM view follows Kress and van Leeuwen in assigning a central role to the 'materiality' of *all* multimodal documents. Traditional views of genre that focused exclusively on texts typically overlooked the 'functionality' contribution of the artefacts themselves because this was not sufficiently foregrounded by their linguistic starting point. When we move to web-based documents, and the self-evident fact that one must 'interact' with the document in order to reach it at all, then this component of genre is no longer to be ignored. But, when examined more closely, we find similar examples of precisely the same phenomena for all multimodal documents.

Anchoring genres in both societal context and time as we propose reduces the likelihood of characterising any cybergenre as an entirely 'new' class of genres, thereby providing additional ammunition for arguments that have been made against an alleged 'new media revolution' (cf., e.g., Manovich 2001, Hocks and Kendrick 2003). As Shepherd and Crowston write, the replicated, variant and emergent genres should preserve aspects of their original content-form mappings as a useful design feature. We go further and suggest that technological limitations and design practices both combine to restrict innovation in this, as well as in every other, medium. Much of what is currently seen as 'novel' or 'new' on the web is then actually caused by it not being possible to track originating configurations of genre components in sufficient detail. And it is to this, more general task that we now return.

5.3 Representing genre

We need now to set out how we can begin to represent genre in a way that can drive empirical inquiry and organise empirical results concerning potential genre descriptions and inter-relationships. This section sets out those approaches to representing genre that have been proposed previously in the literature. We briefly discuss the merits and limitations of these approaches, and propose some working methodological assumptions for taking us further. At the current state of the art, there is no adequate general solution to the problem of representing genres available; the best we can do, therefore, is to establish the kinds of properties that any such solutions are going to require and to provide a framework within which we can move towards this.

Traditionally, genres have been represented in isolation or in terms of loose collections of more or less 'similar' genres. To make an account of genre maximally predictive and useful for organising empirical results, we need to find better ways of bringing out both the individuality of individual