

PART 6

Moving Beyond the Visual





Multimodality and Multimodal Research

Theo van Leeuwen

INTRODUCTION

The term 'multimodality' dates from the 1920s. It was a technical term in the then relatively new field of the psychology of perception, denoting the effect that different sensory perceptions may have on each other. An example of this is the so-called McGurk effect: if people are shown a video of someone articulating a particular syllable, for example /ga/, while hearing another syllable, for example /ba/, they perceive neither /ga/, nor /ba/, but /da/ (Stork, 1997: 239). In other words, perception is multimodal. It integrates information received by different senses.

More recently the term has broadened to denote the integrated use of different communicative resources such as language, image, sound and music in multimodal texts and communicative events. The impetus for this came from linguistics and discourse analysis rather than from psychology. As soon as linguists began to study texts and communicative events rather than isolated

sentences, they realized what they should have known all along: that communication is multimodal; that spoken language cannot be adequately understood without taking non-verbal communication into account; and that many forms of contemporary written language cannot be adequately understood unless both text and image are taken into consideration. In time this discovery led to the development of multimodality as a field of study investigating the common properties of the different modes in the multimodal mix and the way they integrate in multimodal texts and communicative events.

It is not difficult to see why such a field of study should have developed. From the 1920s onwards, public communication had become increasingly multimodal. Film had changed acting, for instance, enlarging subtle aspects of non-verbal communication, and so influencing how people talk and move and smile the world over. Later, television had made non-verbal communication a decisive factor in politics, most famously in the televised

election debate between Nixon and Kennedy. Writing, too, had become multimodal, as illustrations and page furniture broke up and reshaped the pages of books and magazines. Like scholars in other fields of study, linguists took notice. In the course of the twentieth century four schools of linguistics would begin to engage with communicative modes other than language. The first was the Prague School, which, in the 1930s and 1940s, extended linguistics into the visual arts and the non-verbal aspects of the theater, and which included studies of folklore and collaborations with avant-garde artists (cf. for example Matejka and Titunik, 1976). In the 1960s Paris School structuralist semiotics also used concepts and methods from linguistics to understand communicative modes other than language. Largely inspired by the work of Roland Barthes, the Paris School mostly focused on the analysis of popular culture and the mass media rather than on folklore or avant-garde art (for example, Barthes, 1967, 1977, 1983). In roughly the same period American linguists took an interest in the multimodal analysis of spoken language and non-verbal communication. Ray Birdwhistell (for example, 1973) developed an intricate set of tools for analyzing body motion, and Pittenger et al. (1960) published a highly detailed and groundbreaking multimodal analysis of the first 5 minutes of a psychiatric interview. The fourth school emerged in the 1990s. Inspired by the linguistics of M. A. K. Halliday (1979, 1985), it was this school which adopted and broadened the term 'multimodality', and it is this school with which I will be predominantly concerned in this chapter. By the 1990s, of course, the production of multimodal texts was no longer the domain of specialists, of artists, designers, media producers, and so on, but accessible to all computer users. The study of multimodality therefore had become more urgent, and the Hallidayan-inspired approach was often applied to the development of multimodal literacy in young children and teenagers (for example, Kress, 1997; Kress et al., 2005).

Earlier work in multimodality had focused on individual communicative modes, using linguistic models to describe, for instance, visual communication as a 'language of images', musical communication as a 'language of music', and so on. As always, Roland Barthes was an exception here as he not only paid attention to the general principles underlying multimodal communication, devising a conceptual framework that could be applied to all communicative modes (1967), but also developed a theory of the ways text and image can integrate in multimodal texts (1977). The Hallidayan school also began its inquiries into multimodality with 'grammars' of individual communicative modes such as images (O'Toole, 1994; Kress and Van Leeuwen, 1996). But Kress and Van Leeuwen's theory of composition (1996: 181–230) already brought out that composition can be used to integrate, not just text and image, but also, for instance, objects in space and the elements of the built environment, and Van Leeuwen's account of sound (1999), too, was multimodal, in that it paid close attention to the principles common to speech, music and other sounds as well as to the way speech, music and other sounds integrate in multimodal texts and communicative events. In a later book, Kress and Van Leeuwen (2001) sketched the outlines of a more general theory of multimodality, and in the final part of *Introducing Social Semiotics* (2005), Van Leeuwen brought together his work on the integration of time- and space-based communicative modes. At roughly the same time, Baldry and Thibault (2006) published an influential and widely used method for the transcription and analysis of multimodal texts. A related approach to multimodality (for example, Norris, 2004) took up the American tradition of the micro-analysis of contextually situated communicative interactions which had begun with the already mentioned study *The First Five Minutes* (Pittenger et al., 1960). The growth of the field of multimodality studies can be tracked in the pages of the journal *Visual Communication*, and in a range of edited collections including

O'Halloran (2004), Ventola et al. (2004), Norris and Jones (2005) and Jewitt (2009).

Multimodality, then, is first of all a term for a phenomenon rather than a theory or method. It denotes a key aspect of communication: namely, that it melds a variety of communicative modes into a coherent and unified whole. As such, the term is now gaining currency beyond the narrower domain of linguistically inspired multimodality studies. But multimodality is also a term for a particular set of theoretical concepts and analytical practices, aimed at understanding how a variety of communicative resources is used in multimodal communication. Although originally inspired by linguistics, in recent years it has become more interdisciplinary, with, for instance, work on visual communication drawing both on linguistics and art theory, and work on sound drawing both on linguistics and musicology, and so on. It is with some of these concepts and methods that I will be concerned in this chapter.

TEXT AND IMAGE

Barthes' early work on the relation between text and image (1977) remains a key reference for two reasons: (1) it introduced a basic conceptual framework for understanding the integration of text and image in multimodal texts and (2) it showed that text–image relations are culturally and historically specific. This is an important point, because while there is certainly a perceptual basis for multimodality, the roles given to different communicative modes and the values attached to them vary and are motivated by the interests and preoccupations of specific socio-cultural contexts and historical periods.

Barthes introduced three distinct semantic, content-based, relations between text and image. In the first two, 'illustration' and 'anchorage', image and text convey essentially the same content (though, of course, in different ways), but the order in which they are read and understood creates a subtle difference.

In the case of 'illustration', the text is primary, and the image interprets it in a particular context and for a particular audience. In the Middle Ages, 'illustration' was the dominant text–image relation. The most highly valued images illustrated the key stories and key concepts of the time – stories from the Bible and Greek mythology and theological concepts, predominantly. While these stories and concepts were relatively few and relatively stable, images provided ever new illustrations of the same few stories and the same few concepts, making them attractive and accessible to audiences who could not understand Latin, and placing them in familiar, recognizable environments.

In the case of 'anchorage', images are not understood with reference to a text but seen as naturalistic representations of the world. Therefore they do not come already impregnated with cultural meaning, but are potentially open to a variety of readings, so that they need linguistic interpretation, linguistic closure, both in terms of their denotation, of who or what they represent, and in terms of their connotation, of their more abstract, conceptual significance. 'Anchorage' began to take over as the dominant form of image–text relation in the Renaissance, when science and exploration encouraged images that could document the world, so making it amenable to scientific labeling, classification and interpretation.

In the case of 'relay', Barthes' third category, there is no redundancy between text and image. Text and image do not 'say the same thing' but convey different, complementary content. In the case of 'dialogue scenes' in films and comic strips, for instance, the image shows the speakers while what they say is conveyed linguistically. Text and image therefore depend on each other to convey the whole of the content. In contemporary communication, 'relay' is extending into many other fields. In signaling it, Barthes was ahead of his time, even though it was too early to see the future significance and ramifications of the move from 'anchorage' to 'relay'. The essence of this move is that while in the case

of ‘anchorage’ and ‘illustration’ text and image each convey a complete, whole message, in the case of ‘relay’ neither the image nor the text can convey the message on its own. In Figure 28.1, for instance, the idea of ‘mortal danger’ is expressed both visually, by the skull and crossbones, and linguistically, by the words ‘*pericolo di morte*’. Perhaps the sign-makers found the skull and crossbones icon insufficiently explicit and thought it was in need of verbal anchorage to be understood. In Figure 28.2, however, the visual and the verbal must be put together (and together with the immediate environment and placement of the sign) to understand the full message, which ‘says’ something like ‘workmen-working (*image*) ends (*text*) here (*placement of sign*)’. To put it another way, picture and text together form the equivalent of a clause (‘workmen-working ends’), with the subject expressed pictorially and the predicate verbally.

The concepts of ‘illustration’ and ‘anchorage’ are closely related to the concept of ‘elaboration’ in Halliday’s theory of conjunction (1985: 302–309), a theory which maps the possible semantic relations between clauses in a linguistic text where such relations may either be made explicit by conjunctions such as ‘and’, ‘but’, ‘because’, etc., or remain inexplicit. In the case of ‘elaboration’, one clause elaborates another, for example by rephrasing it, by making it more specific (or more general), or by providing an example of it. Nothing new is added to the



Figure 28.1 Danger sign



Figure 28.2 Work in progress sign

content of the first clause; it is just reiterated in some way, to clarify or strengthen it. When applied to text–image relations, this can add further detail to Barthes’s conceptual framework. ‘Anchoring’ texts such as captions can, for instance clarify what the image shows, or generalize it, and illustrations can add specific detail to what is mentioned in the text, or exemplify it. Van Leeuwen (2005: 219 ff) discusses the elaborative relations between the commentary and the images of non-fiction films and television programs in this way. The ‘relay’ relation, on the other hand, is not conjunctive, because while conjunction connects complete propositions (clauses and their visual equivalents), ‘relay’ relations connect incomplete propositions, parts of propositions (for instance, a ‘subject’ and a ‘predicate’), or a speaker and what he or she says.

Martinec and Salway (2005) have provided the most detailed Barthes-inspired account of

the semantics of text–image relations. The core of their approach is a distinction between the relative *status* of text and image and their *logico-semantic relationship*. When image and text are equal in status, they say, the whole of the image connects to the whole of the text. An image may for instance show a group of people walking to a courthouse while the text says ‘Janklow walks up to the courthouse with his legal team’. Image and text can be equal in both ‘independent’ text–image relations (‘anchorage’ and ‘illustration’ relations, as in Figure 28.1), and ‘complementary’ image–text relations (‘relay relations’, as in Figure 28.2). In unequal image–text relations, the image relates only to part of the text, or the text only to part of the image. An image may show the Nobel Prize winning author Naipaul seated in front of his book case. The image relates only to Naipaul, not to his winning of the Nobel Prize, which is mentioned only in the caption. The status of the image can therefore be likened to that of a subordinate clause in a sentence (something like, ‘Naipaul, who looks like this image, won this year’s Nobel Prize’).

Like Van Leeuwen (2005), Martinec and Salway follow Halliday in distinguishing different types of elaboration: ‘exposition’, where image and text are at the same level of generality; and ‘exemplification’, where either the text is more general than the image, or the image more general than the text, as in a skull and crossbones icon accompanied by the words ‘high voltage’. ‘Extension’ is then the addition of new, related information, as in the captions of paintings in art books, which may add details that cannot be seen, such as the year in which the painting was created, and the museum that owns it. They also introduce the possibility of ‘enhancement’, where the text may add a ‘circumstantial’ element to the image, or vice versa – for instance, the location or timing of an event, or its reason or result. In Figure 28.3 the text adds such a circumstance, indicating the extent of the stretch of road where reindeer may cross. And, finally, they include the relation of ‘projection’ – the relation between the



Figure 28.3 Crossing reindeer sign

image of a speaker and his or her words (this may also include the relation between a ‘thinker’ and his or her thoughts, as indicated by thought bubbles in comic strips).

Following Martinec and Salway, each image should be analyzed both for the relative status of text and image and for the type of semantic relation between them. In the case of the image of Naipaul in front of his books, the status is unequal, with the image subordinate to the text, while the relation is one of elaboration (exemplification), because the relevant text (‘Naipaul’) is more general than the image, which shows details of Naipaul’s appearance and relevant attributes (books). Figure 28.4 summarizes Martinec and Salway’s account in the form of a system network. The curly bracket (brace) indicates ‘both...and’ (i.e. text and image must be analyzed both for their relative status and for the logico-semantic relation between them), while square brackets indicate ‘either...or’ (for example, text and image are either equal or unequal).

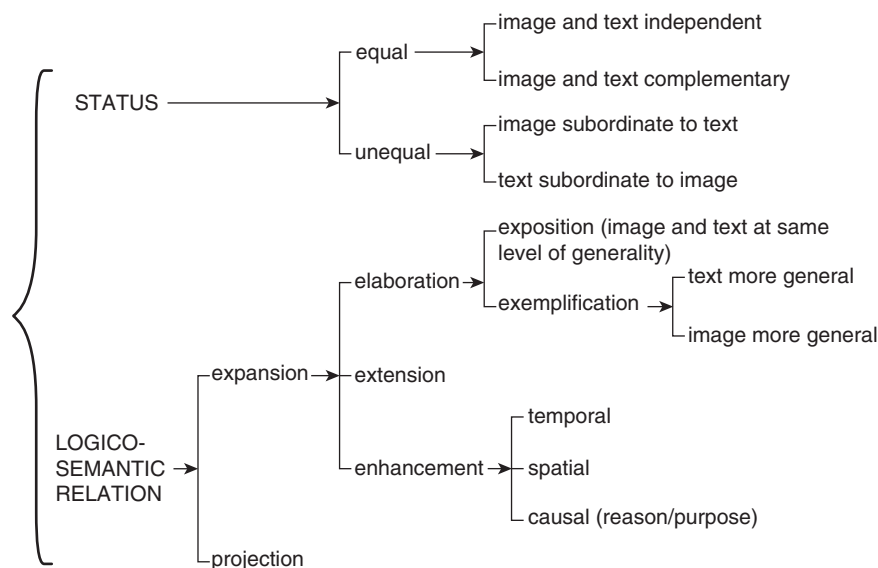


Figure 28.4 Network of text–image relations (after Martinec and Salway, 2005: 358)

The approaches I have discussed so far take their cues from the linguistics of conjunction and complex clause construction. But it is also possible to approach text–image relations from the other, visual, end, using theories of visual composition to explain the relation between text and image. These two approaches do not contradict each other. People are perfectly capable of understanding multimodal texts at two or more levels simultaneously.

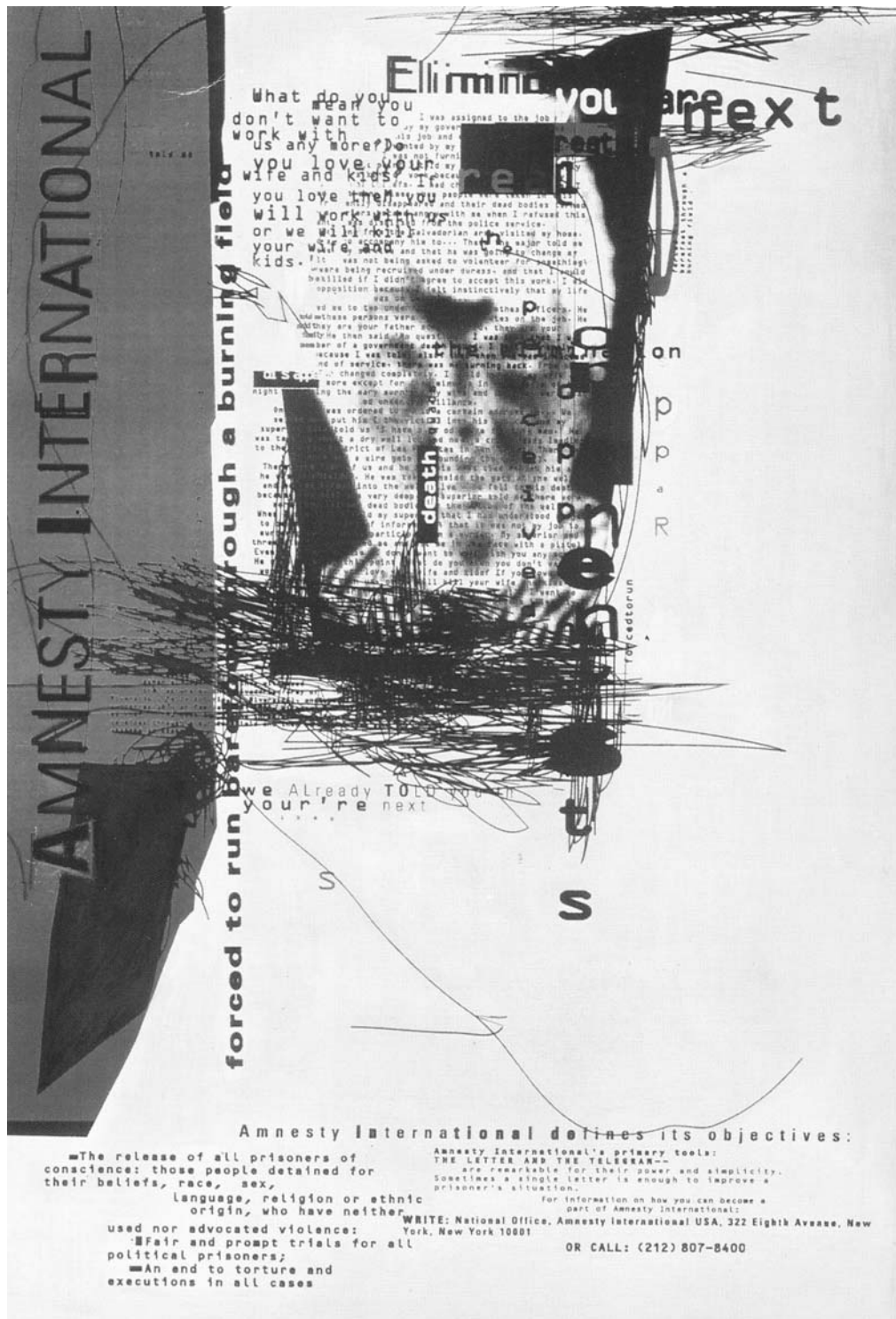
COMPOSITION

Today, visual analysis must go beyond the analysis of images. Texts like Figure 28.5, a promotional poster for Amnesty International (Joan Dobkin, 1991), or Figure 28.6, a web page, hardly contain any images. Yet they are profoundly visual, full of color, line and texture. Multimodality takes a step forwards here. Language itself becomes visual, and composition, color and typography play a key role in creating structure and meaning.

Until the twentieth century, a page of printed text was a single field of dense text surrounded by margins, although there had

been exceptions such as the two-column *Gutenberg Bible* or scholarly works which had the text in the center and surrounded it with scholarly commentary. Newspapers changed this, turning the page into a modular grid of equal-sized rectangles that can be filled with either pictorial or verbal content, and in which a given text or picture might fill any number of modules, making many different page layouts possible (see Lupton, 2004). Today's websites use the same principle, with designers making creative use of the table facility provided by the designers of HTML (Hypertext Mark-up Language) in ways that have little to do with the original function of tables (see, for example, Figure 28.6).

If linear text gets its shape and structure from the progressive development of an argument or a storyline, how do such non-linear, fragmented texts cohere? Kress and Van Leeuwen (1996) have attempted to answer this question by proposing that the spatial 'zones' of pictures, pages and screens (left and right; top and bottom; center and margins) interrelate textual elements by providing them with specific 'information values'. To start with left and right, if there is *Polarization* (some kind of difference or contrast) between an element placed on the left and



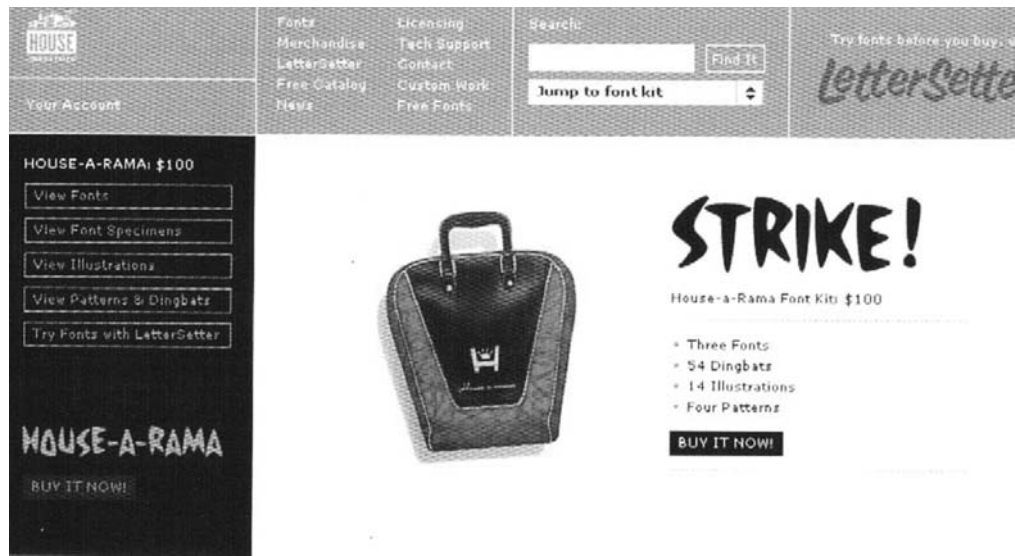


Figure 28.6 Housind.com website. (Designed by Andy Cruz, Tal Leming, Ken Barber, Rich Roat and Bondé Prang, 2004.) © 2004 Princeton Architectural Press, Inc. Reproduced with permission

an element placed on the right, then the left element will be understood as the *Given*, as a departure point for the message that is, or should be, already familiar to the reader or viewer, while the right element will be understood as the *New*, the element that contains the information the message is trying to get across. This left–right information flow clearly corresponds to the left–right mode of writing and reading in Western culture, and is indeed reversed in cultures that write from right to left.

It is important to note that the Given–New relation applies to pictorial compositions such as the painting shown in Figure 28.7, as well as to multimodal, text-and-image compositions such as Figures 28.5 and 28.6. In Figure 28.7, a seventeenth-century painting by Guercino called ‘Et in Arcadia ego’ (‘I am there, even in Paradise’; also known as ‘The shepherds of Arcadian’ and ‘The Arcadian Shepherds’), the left shows a figure who contemplates the New; the object the painting invites viewers to also meditate about is death. Examples like this illustrate that visual art has made use of Given and New composition for a long time. What is new is its introduction into the domain of modern, multimodal writing.

When there is vertical polarization – polarization between an element placed in the upper and an element placed in the lower section of a picture, page or screen – the top element is the *Ideal*, the idealized or generalized essence of the message, and the bottom element the *Real*, contrasting with the Ideal in presenting factual details, or documentary evidence, or practical consequences. In single-page magazine advertisements, for instance, the Ideal usually depicts the



Figure 28.7 ‘Et in Arcadia Ego’ (Guercino, 1621–1623)

'promise' of the product, the glamour or success it will bring to consumers, or of the sensual satisfaction it will give them, while the Real shows the product itself and provides factual information about it. In Figure 28.5, the name of the organization, 'Amnesty International', appears on the left, as the Given, while the New is the typewritten story of persecution, which is to persuade readers to support Amnesty. This story is also the Ideal, the essence of Amnesty's message, and the Real provides the official version of Amnesty's objectives and methods as well as a contact address. In Figure 28.6, the Ideal contains the branding, the Given the global

navigation, while the New and Real highlight a key product.

The *Center* is another key compositional zone. Instead of polarizing the elements of the composition, the Center unifies them, providing the *Margins* that surround it with a common meaning or purpose. In Figure 28.8, for instance, the Center represents a happy, indeed overjoyous, Rank Xerox employee, and the words that surround him suggest the various ways in which Xerox makes its employees happy by 'recognizing and rewarding' their efforts.

The compositional schemas discussed here can be combined in various ways.



Figure 28.8 'Recognizing a job well done' (*Document Matters*, July 1996). © 1996 Xero28. Reproduced with permission

Figure 28.6, for instance, contains three horizontal zones – the name of the organization, Amnesty International on the left; an empty yellow zone on the right, as the New, the zone of the viewer's action ('You are next'); and the typewritten and partially obscured story in the center as the 'Mediator' (Kress and Van Leeuwen, 1996: 208–211) that connects the two.

Such compositional schemas are multimodal for two reasons. They can apply to any kind of spatial configuration, whatever its mode – image, text, museum display, stage design, architectural façade – and they can integrate different kinds of element (for example, text and image) into a multimodal whole. But it is a different kind of integration from the one discussed in the previous section. The connections it establishes between elements are visual rather than verbal, informational rather than semantic, and geared towards hierarchies of importance and attention rather than to internal, logical coherence. The system network in Figure 28.9 provides a summary of the key terms.

DIAGRAMS

The pie chart in Figure 28.10 summarizes the result of a staff feedback survey. It tells us that the survey had three kinds of responses: responses with a positive attitude to change (the top segment, in yellow); responses with a negative attitude to change (the bottom left segment, in blue); and responses with a 'wait and see' attitude to change (the bottom right segment, in purple). The shading in the center indicates the amount of responses in each category. In many other contexts, a summary of this kind would have been entirely verbal, announcing the three types of response, defining and illustrating each, and then indicating the percentage of respondents in each category (which is what pie charts are normally used for). But here only the quotes are verbal. The categories are defined, not by words, but by a color (a sunny yellow for positive responses, cooler and dimmer blue and purple for the less positive ones), and the proportions are shown only visually. Composition also plays a role, centering the

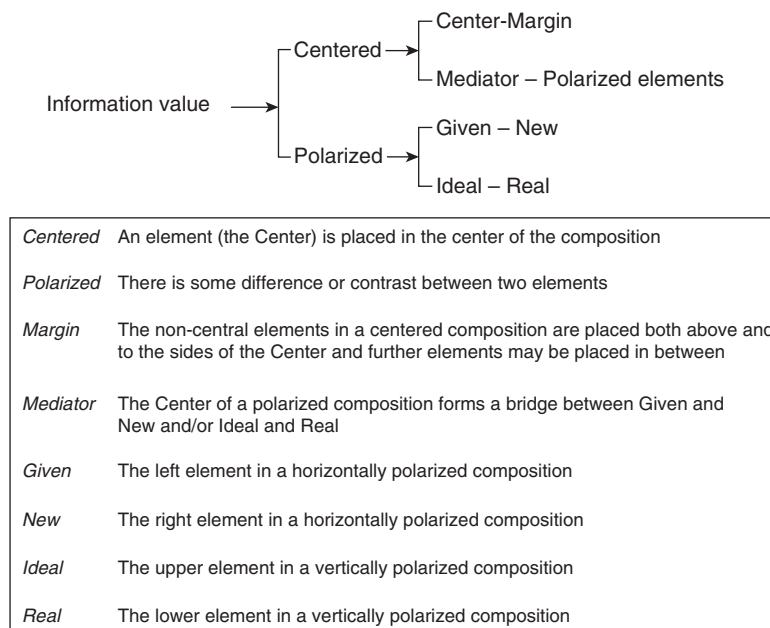
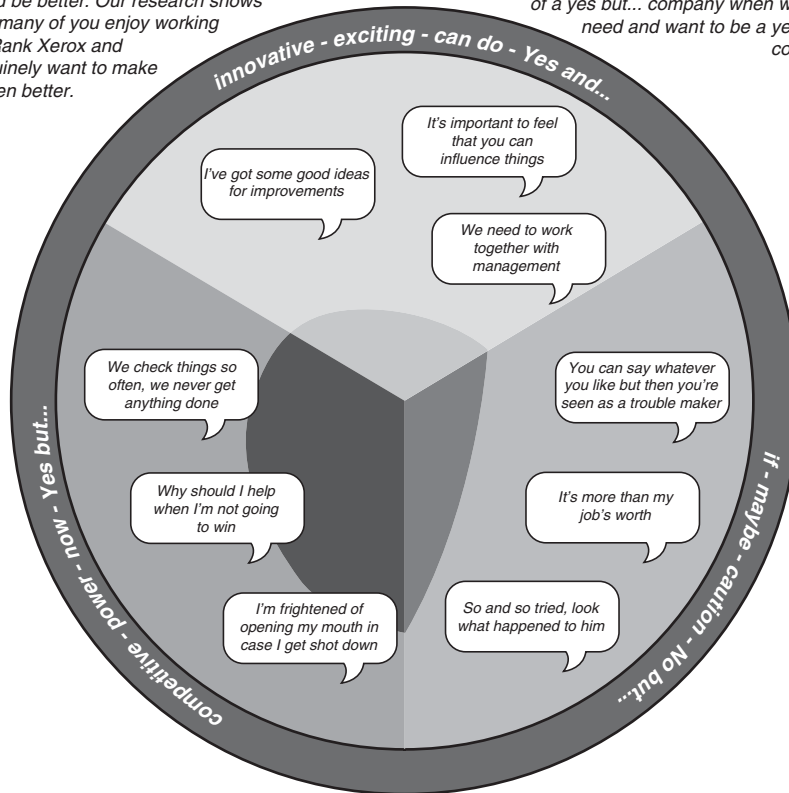


Figure 28.9 Network of information value

Where we are now

We've been talking with many of you to understand where we are now, and what it's like working here and what could be better. Our research shows that many of you enjoy working for Rank Xerox and genuinely want to make it even better.

One thing we're all clear about is that we want to change. You can see from the illustration that we've become more of a yes but... company when we really need and want to be a yes and... company.



"Why do we have to change?"

Figure 28.10 'The Three Voices of XBS' (Rank Xerox, 1997). © 1997 Xero28. Reproduced with permission

distribution and making the positive responses Ideal and the less positive ones Real.

Much information is now presented in this way, diagrammatically rather than discursively. Thanks to Excel, tables are particularly popular: for instance, in corporate documents such as 'action plans', where the horizontal dimensions will include such categories as 'action', 'person responsible', 'timeline', 'success criteria', etc. In the past, elements like 'action' and 'person responsible' would have been connected through the grammar of sentences. If I am the person

responsible and writing is my action, I write 'I write' and not as shown in Table 28.1.

Today, a new form of writing fuses language and visual communication. The elements are still linguistic, words in boxes, but they are incomplete – nominal groups rather than fully

Table 28.1 Example of language and visual fusion

Action	Person responsible
Writing	Theo van Leeuwen

formed clauses. But the ‘verbs’ that connect them into complete ‘propositions’, the equivalent of clauses, are visual, spatial structures, diagrammatic templates. As a result, the words in the boxes can easily be replaced by pictures – for example, in the above example, a writing icon and a picture of Theo van Leeuwen.

Martinec and Van Leeuwen (2008) have described a number of such diagrammatic ways of arranging information and shown how they underlie the structure of contemporary multimodal texts such as websites. The ‘star’ arrangement (center with marginal elements around it) and the table we have already mentioned, but there is also the taxonomy (Figure 28.11a), which links elements with a ‘is a kind of’ or ‘is a part of’ relation, and the network (Figure 28.11b), which can link elements along a variety of semantic dimensions: for example, ‘a is like b’, ‘a combines with b’, ‘a co-occurs with b’, ‘a does something to b’, and so on.

Diagrams provide only a limited vocabulary of visual ‘verbs’ and often use the same form for several different meanings, as with the two examples above, or with the arrows

in flowcharts, which sometimes indicate actions, sometimes transport, sometimes sequentiality and sometimes causality. What for instance is the link between ‘ownership’ and the ‘X-team’ in Figure 28.12? Does ‘ownership’ *relate to*, or *feed into* or *cause* the X-team? Diagrammatic representations of this kind can be much less precise than they seem. Perhaps a more subtle and complex language of boxes and arrows needs to be developed. But, in the meantime, important communicative work is done through diagrams that often obscure or contradict the verbal information that accompanies them. Multimodal analysis can therefore make an important contribution by analyzing the communicative effectiveness of diagrams and making suggestions for their improvement, as is done, for instance, in Martinec and Van Leeuwen (2008).

FRAMING

The brochure shown in Figure 28.13 was distributed to every household in Australia in 2007 to alert parents to the dangers of the Internet. Frame lines separate the zone of the Given (which shows children using the Internet) from the zone of the New (which contains advice for parents), and also separate the three areas demarcating different age groups (under 8 year olds; 8–12 year olds; 13–17 year olds). In addition, the photos on the left are all dominated by the color beige, while the text boxes on the right are in different shades of mauve. All this creates degrees of connectivity and separation between the six key elements of this page. The beige/mauve color contrast makes for a strong divide between the Given and the New (children and their parents), while the color difference between the boxes on the right is much smaller, suggesting that these three elements are more similar and present different varieties of the same kind of content.

I use the term ‘framing’ here to refer to visual devices that can create degrees of

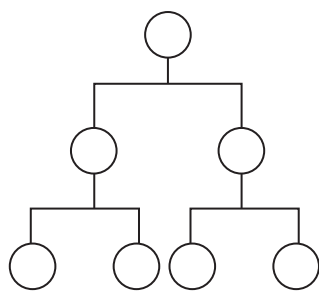


Figure 28.11a Taxonomy

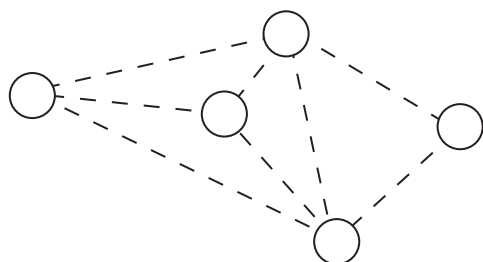


Figure 28.11b Network

What makes an X Team...

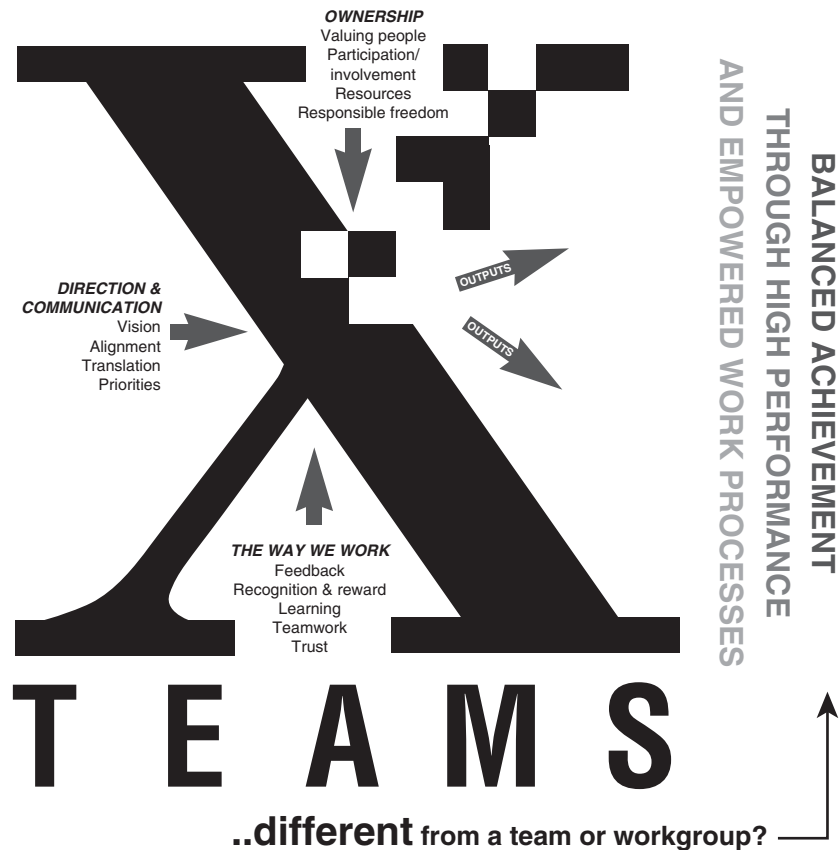


Figure 28.12 'What makes an X-team?' (Rank Xerox, 1997). © 1997 Xero28. Reproduced with permission

connection or disconnection between the elements of a composition—the visual equivalents of punctuation, indentation and interlineal spacing, you might say. Disconnection devices include actual frame lines, which can of course vary in thickness, empty space between elements, and contrasts in tone or color. Connection devices include similarities in color (or other visual aspects) and connecting 'vectors'—lines that lead from one element to another. Different kinds of framing may subtly differ in meaning. Frame lines, for instance, create radically different territories for the elements they divide, while the use of empty space places them in the same textual territory, but at a distance from

each other, in the way that people sitting at different tables in a café share a space and do the same kind of thing without immediately interacting.

Like composition, framing is a key multimodal device, existing, one way or another, in every communicative mode, as every communicative mode has to have resources for creating 'boundaries' between units of meaning, so creating the equivalent of sentences, paragraphs, sections, etc. Van Leeuwen (2005: Chapter 1) applies the concept of framing to classrooms and offices, which, today, tend to diminish framing ('open' classrooms and offices), although people often introduce ad hoc 'frames' by dragging in various kinds of partitions.

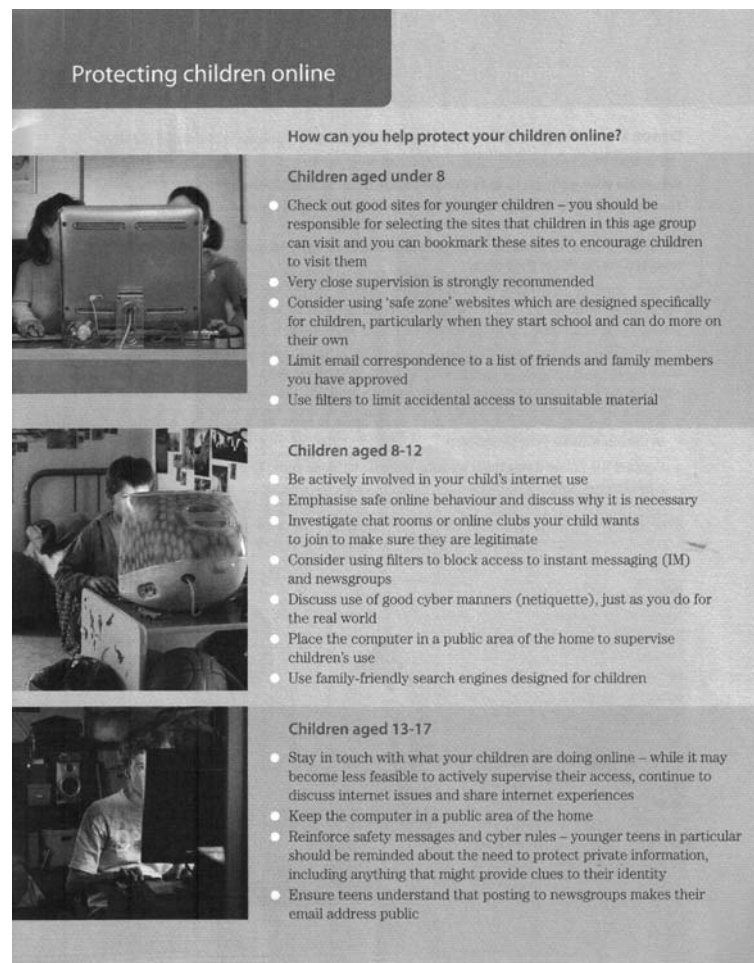


Figure 28.13 Page from Australian Government brochure (Nairn, 2007). © 2007 Commonwealth of Australia. Reproduced with permission

COLOR

Color is usually viewed as expressing emotive meanings. In public and commercial buildings it is said to have a direct effect: 'bright reds are energizing and good for offices in the banking or entertainment fields' (*Guardian*, 3 September 2001: 5); 'pink, properly applied, relaxes hostile and aggressive individuals within 15 minutes' (according to the Naval Correctional Center in Seattle, as cited in Lacy, 1996: 89). In the private domain it is said to express personality traits. 'A yellow entrance hall usually indicates a person who

has ideas and a wide field of interests... as this color is associated with the intellect, ideas and a searching mind' (Lacy, 1996: 29). This approach to color has been widely adopted by psychologists, with the result that color is often treated as a natural phenomenon. Yet color meanings are neither universal, nor of all times. In the Middle Ages, for instance, color denoted abstract concepts such as 'faith', 'hope' and 'charity', and the rank and status of knights, and the current discourse of color as emotive and psychological originated in the Romantic era, largely inspired by Goethe's remarkable book *Theory of Colors* [1810]

(1970). The changing fashions in color, too, should alert us to the fact that color preferences are not, or not just, a matter of individual personality, but are also cultural. It is therefore important not to rely too much on the decontextualized lists of emotive color meanings that can so readily be found in publications by color experts and on the Internet. Color codes do exist, but they usually apply to specific, narrow fields such as traffic and safety, or uniforms. In New South Wales, for instance, yellow traffic signs convey information ('there are pedestrians crossing here'), whereas red is used for commands such as 'Stop!' or 'No parking!'. At the University of Newcastle (NSW), Doctors of Business have the hood of their academic gown lined with 'deep Indian red' silk, Doctors of Education with jade silk, Doctors of Music with lilac silk, and so on. But in many domains of contemporary culture, color meaning is not so definite, cannot so easily be captured in a lexicon-like format, and comes about in context-specific ways, through connotation and metaphor (Kress and Van Leeuwen, 2002).

In the case of connotation, a color is imported from a specific domain (e.g. the Army) into a new domain (e.g. street fashion), where it can then convey meanings and values linked to the domain from which the color is imported, or some fictionalized version thereof (e.g. the qualities of action heroes). At my university, the Faculty of Humanities chose a brownish orange as its color because it is the color of the 'red earth' in Australia's dry and arid center, and the Faculty wanted to express its interest in practicing a form of humanities with a distinct Australian accent. This color was therefore imported from the Australian landscape into the domain of corporate identity, to signify identification with a certain kind of Australian-ness.

Hue, the aspect of color that makes us recognize colors as 'red' or 'green' is only one of the significant dimensions of color. Actual colors – for example, an actual red – are composites of a number of features: any red is either light or dark; pure or mixed in with blue or brown; intense and highly saturated or

pale and diluted; and so on. These features all contribute to the communicative potential of the color, because what they literally and visibly *are* can become what they, more figuratively, *mean*. A given red may *be* 'bright' and also *mean* bright. The context will then narrow this 'brightness' down – it could, for instance, be 'brightness' in the sense of 'illumination', as when we say that someone 'sees the light'; 'brightness' in the sense of fame, as when we say that someone is a 'shining star'; and so on. The yellow in Figure 28.13 is bright and highly saturated, and as it is here chosen to represent 'optimism', and a 'positive outlook', the metaphoric meaning of brightness is narrowed down to 'optimism', as when we say 'look on the bright side', etc. The following features can all contribute to the communicative potential of a given color – or a given color scheme, because features such as 'brightness' and 'saturation' may also form the unifying factor in color schemes:

Brightness

Brightness (also known as 'value') is the scale from light to dark. I have already commented on its communicative potential.

Saturation

Saturation is the scale from full color to pale, pastel versions of the same color. Its communicative potential is based on the idea of 'fully expressed', 'intense' versus 'toned down', 'restrained'.

Purity

Purity is the scale from maximum purity to maximum 'mixed-ness' or 'hybridity', or, if you like, 'impurity'. Its communicative potential flows from this characteristic.

Modulation

Modulation is the scale from a flat version of a color to maximally varied shading and nuance,

and can therefore express the contrast between 'bold', 'straightforward', and 'subtle', 'nuanced' (or, more negatively, 'simplistic' versus 'fussy'). Flat colors are often more 'conceptual' (green as *the* color of grass), and modulated colors more 'naturalistic' (photographs of grass will show many shades of green)

Luminosity

Luminosity is the range from 'dull', 'matt' to 'reflective' (e.g. silver and gold) and 'luminous' colors (e.g. those of stained glass through which light shines from behind).

Hue

Hue, finally, is the quality of color that makes it blue, green, red and so on. As a scale, it is also the scale from 'warm' to 'cool', and many of its associations derive from this characteristic.

Today, colors and color schemes are also used to create textual coherence. In Figure 28.14, a page from *Cosmopolitan* magazine, the words 'love' and 'lust' are printed in pink, a color with a clear connotation of femininity. But metaphor plays a role as well: as we move from 'love' to 'lust' the color becomes progressively darker and more saturated: 'lust' is here a 'darker' and more 'intense' variety of 'love'. The same color is also used for the words 'how to'; for the headings of the individual pieces of advice beneath the picture ('Dish some insight', 'Inspire a repeat performance', 'Pay attention'); for the text box on the left; for the lips of the model in the photograph; for the telephone in the picture at right bottom; and for the words at the bottom of the page which point towards the next page. It therefore forms a thread through the page, tying together the key elements that structure it and make it hang together meaningfully: the title, the nature of the article as an 'advice' genre ('how to'), the actual items of advice, and the key theme that permeates the whole page: communication (lips, a telephone).

Clearly, color is communicative: it can convey ideas (e.g. 'femininity' and the 'darkness' and 'intensity' of 'lust'); it can affect people emotively; and it can provide textual cohesion, linking the elements of the multimodal page together into a meaningful structure. Color is also essentially multimodal. It cannot exist on its own (color field paintings remain a relatively limited genre) and must necessarily work together with images, letter forms, clothes, interior decoration, objects, architecture, and so on. It can only ever be 'multi-'.

TYPOGRAPHY

Traditionally, typography was a self-effacing craft. Legibility was its main concern, and typographers were adamant that letter forms should not draw attention to themselves. More recently, typography has become a communicative mode in its own right, and many printed texts now communicate not just through the linguistic meaning of the words but also through their *typographic meaning*.

As with color, typographic meaning can either derive from connotation or from metaphors, and always needs the context to narrow it down. In the case of connotation, the reader recognizes the context from which a letter form comes, and can therefore link that form to meanings and values associated with the context it comes from. The typewriter font used in the Amnesty poster (Figure 28.5) evokes low technology. It looks as if it has been written with an old, battered portable typewriter, by someone without access to a computer, somewhere in the Third World, and it therefore has the quality of an authentic document. But metaphor also plays a role – there is deliberate irregularity in the way the letter forms are spaced and placed on the baseline. Typographers often employ irregularity, sometimes to signify an unwillingness to produce neat, regular forms, which, in context, can become 'unconventionality', 'playfulness', 'rebellion', 'lack of discipline' and so on, sometimes to signify an impossibility of producing regularity, perhaps because of surging emotions, an injury or

COUPLES BY BROOKE LE POER TRENCH

LOVE AND LUST

HOW TO...
Bring Out His Romantic Side

"I knew you had it in you."

SPAT OF THE MONTH
"He asks for my advice but never takes it."

"Be flattered that he trusts you enough to want your opinion," says Janet Lever, PhD, professor of sociology at California State University at Los Angeles. And remember, you're giving an opinion, not a command he must obey. "When people are seeking advice, they usually poll and consider a few points of view before making up their mind," says Lever. So while he may not have acted on your words directly, you've contributed to the solution.



Are you wasting your breath?

BLOW-HIS-MIND TIP
Have Phone Sex

Get yourself in a comfy position and dial your dude for some long-distance lovin'. Speak a little softer than you normally would, and once you've said your hellos (make sure it's not his roommate), parlay into some dirty talk. Describe—in detail—what you wish he were doing to you right now. As his breathing gets heavier, ask him what he wants to do with your naked bod...and make your hands obey.

he doesn't get the hint. He'll be encouraged to repeat these actions once he sees how happy you are and realizes his chances of getting lucky have increased...dramatically!

Pay attention. You may have very different ideas of what romance actually is. "For men, it's often looking after their partner." So while you're pining for flowers, he's lovingly filling your car with fuel. The key: Tell him what you want but also appreciate the sweet gestures he comes up with on his own.



Be sure to ignore your call-waiting.

UNLOCK HIS EMOTIONS BY ASKING...IF YOU COULD SPEND A DAY WITH ONE LIVING CELEB, WHO WOULD IT BE?

NOVEMBER 2004 COSMOPOLITAN 83

Figure 28.14 Page from UK *Cosmopolitan* magazine (November 2004)

disability, or, in this case, the battered state of an old typewriter. Metaphors like this are again based either on *what we can see in the letter forms*, on their actual visible characteristics, such as that they are round or angular, curved or straight and so on, or on experience, on *what we know about the production of the letter forms*, for example that irregularity can be caused by poor equipment – leaking pens, or battered typewriters. As in the case of color, such features combine to form the communicative potential of letter forms (Van Leeuwen, 2002). They include:

Weight

Weight is the difference between bold and less bold typefaces. Boldness increases salience and can therefore mean ‘assertive’, ‘solid’, ‘important’ (or, more negatively, ‘domineering’, ‘overbearing’ and so on). Lack of boldness can signify the opposite.

Expansion

Expansion is the difference between condensed, narrow letter forms and expanded, wide letter forms. Narrow typefaces make efficient use of limited space while wide typefaces ‘spread themselves around’. This visible characteristic can then provide the potential for metaphoric meanings such as ‘economical’, ‘efficient’, or, more negatively, ‘cramped’, ‘overcrowded’, etc.

Slope

Slope refers to the difference between cursive, ‘script’-like typefaces and upright typefaces. Cursive writing usually connotes handwriting, and the qualities we might associate with handwriting, qualities such as ‘informal’, ‘personal’, etc.

Curvature

A letter form can stress angularity or curvature, with all the associations that may follow

from this. Curvature may also be realized by the difference between rounded ascenders and descenders (or ascenders and descenders which use loops and flicks) and straight ascenders and descenders.

Connectivity

Letter forms can be connected to each other, as in running script, or lack connection, so that they become quite separate and self-contained. This again ‘means what it is’ – ‘connectedness’ and ‘disconnection’, and as most forms of handwriting are ‘connected’, connectivity also suggests the qualities we associate with handwriting.

Orientation

Typefaces may be oriented towards the horizontal axis by being comparatively ‘flattened’ or oriented towards the vertical axis. The meaning potential of horizontality and verticality derives from our experience of gravity and walking upright, so that horizontality could mean ‘heaviness’ or ‘solidity’ (but also, e.g. ‘inertia’), and verticality ‘upwards aspiration’ (but also, e.g. ‘instability’).

Regularity

Many typefaces have deliberate irregularities, through random variations in size, weight or slope, and through irregular placement on, above or below the baseline. As mentioned, this can signify rebelliousness, playfulness, etc., (whatever contrasts with the strict discipline of regularity) or suggest a technical or physical cause for the irregularity.

Typography is not just about letter forms. Other factors also play a role. Letter forms can be colored, as in Figure 28.14; textured (rough or smooth, shiny or dull, etc.); three-dimensional or flat; still or animated; abstract or semi-pictorial. The spacing between

letters, words and lines plays a key role, as does the use of margins. In the case of Figure 28.14, the words 'love', 'and' and 'lust' overlap, suggesting that, in a magazine like *Cosmopolitan*, 'love' and 'lust' are not distinct, but shade into each other. Examples like this bring out that typography is not just a matter of aesthetics, but makes critical contributions to the message conveyed by the multimodal text as a whole.

Decoration

In many contemporary documents, pictures are no longer discrete elements, but function as a setting, a background for written text. In ads, movie titles, book covers, etc., words have been superimposed on pictures for a long time, but today superimposition has moved into many other domains of writing. The vice-president of my university, for instance, presented his annual budget in a PowerPoint® presentation, superimposing the figures on a calm, blue ocean and a blue sky with small, wispy clouds. In doing so he perhaps sought to present himself as a visionary rather than a dry accountant (he also inserted quotes from visionary business gurus between the slides with figures). In other words, through his choice of a visual background (which, of course, remained constant throughout his presentation) he expressed his identity as a speaker. In other cases, people may use visual backgrounds to express the genre of a presentation or document, or some essential characteristic of its topic.

Traditionally, such backgrounds would have been seen as embellishment, decoration, a kind of wallpaper. But decorative motifs have always been culturally important, whether in script-based Islamic decoration, or in the nature-based motifs of nineteenth-century decoration (e.g. William Morris' flowery wallpapers), and visual analysts should pay close attention to them. Figure 28.15 shows the cover of the annual report of my university's library. The back-

ground contains both words and abstract graphic motifs (lines, grids, circles), some of them illuminated by a kind of glow. The words ('access', 'absorb', 'resource', 'intellect', 'accumulate', 'research') convey key characteristics and values of contemporary libraries. But so do the graphic elements – a rigid grid of rectangles on which larger, more strongly framed zones are superimposed, as well as points and circles that light up like the activated parts in a brain scan. The library is represented here as a rigid, rational system of classification, a grid, but within that grid, there are zones of special strength and value, and moments of illumination.

Decorative patterns have two key characteristics – they repeat themselves and they are endless—so that we can never see all of them in a single glance. This allows us to take them for granted. We know they are there, even when we do not pay attention, and we know they will continue unchanged even when we cannot see them. Thus, they provide constancy and stability to culturally salient meanings, and to the identity of speakers, text genres and specific texts.

CONCLUSION

This chapter has moved from the analysis of text–image relations to the analysis of composition, framing, color, typography and the use of visual backgrounds in what I have called the new writing.

The new writing merges aspects of yesterday's text media and yesterday's image media. The screens of websites and PowerPoint® slides are far more word-oriented and far less image-oriented than the screens of film and television. At the same time they are far more visual than the pages of books. They use the 'landscape' format familiar from film and television and they introduce visual background and filmic transitions such as fades and dissolves into

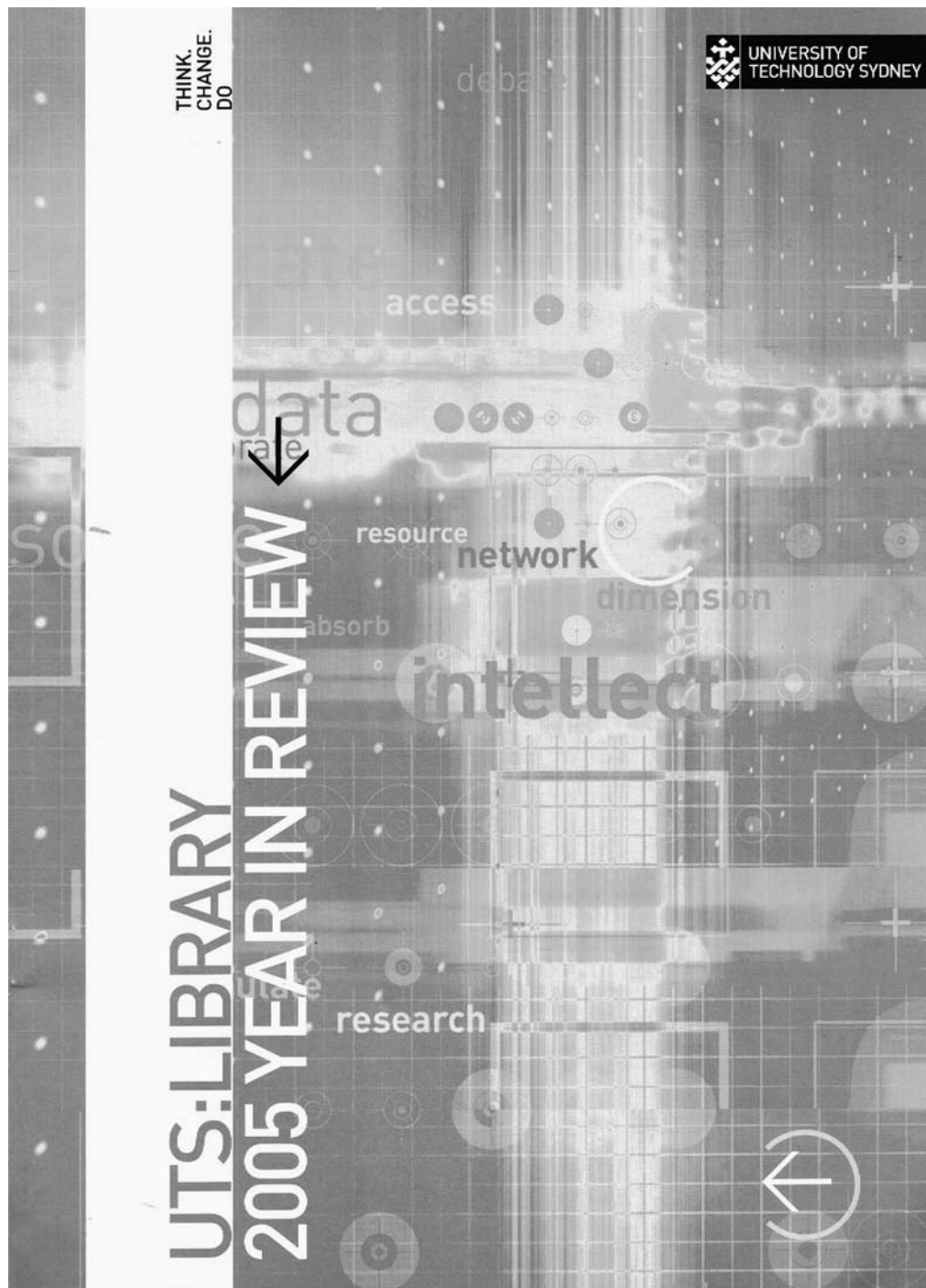


Figure 28.15 Annual report. © 2005 Library, University of Technology, Sydney. Reproduced with permission

written text. And they hang together, not as webs of words, but as multimodal compositions, through composition, framing, color and typography.

As a result, the new writing can no longer be read aloud. It would not make sense, for instance, to read aloud a web page like the one shown in Figure 28.6. The link between speech and writing introduced, long ago, by the alphabet, is severed here, and alphabetic writing is mixed with an increasing range of pictographic signs.

The new writing is also read differently. Because it is no longer 'linear', no longer organized sequentially, we no longer read it from beginning to end, gradually piecing together a sense of overall structure, gradually integrating the meaning of the parts into an overall understanding. The overall structure is now spatial, visible at a glance, and allows us to move down to more detailed levels of reading according to our interests and needs, just as we used to do with images.

Both forms of visual analysis continue to be important, because, for the time being, text and image media will continue to exist alongside the new writing. Newspapers, for instance, appear both in their traditional form and in 'new writing' form, on the Internet (Knox, 2007), and television programs, too, have their online counterparts. But the new writing is rapidly gaining importance, rapidly becoming the dominant form of multimodal communication. It is therefore a crucial task for visual analysis to develop tools for analyzing and interpreting it.

REFERENCES

- Baldry, A. and Thibault, P.J. (2006) *Multimodal Transcription and Text Analysis*. London: Equinox.
- Barthes, R. (1967) *Elements of Semiology*. London: Cape.
- Barthes, R. (1977) *Image-Music-Text*. London: Fontana.
- Barthes, R. (1983) *The System of Fashion*. New York: Hill and Wang.
- Birdwhistell, R. (1973) *Kinesics in Context*. Harmondsworth: Penguin.
- Goethe, J. W. von [1810] (1970) *Theory of Colors*. Cambridge, MA: MIT Press.
- Halliday, M. A. K. (1979) *Language as Social Semiotic*. London: Arnold.
- Halliday, M. A. K. (1985) *An Introduction to Functional Grammar*. London: Arnold.
- Jewitt, C. (ed.) (2009) *The Routledge Handbook of Multimodal Analysis*. London: Routledge.
- Knox, J. (2007) 'Visual-verbal communication on online newspaper home pages', *Visual Communication*, 6(1): 19–55.
- Kress, G. (1997) *Before Writing – Rethinking the Paths to Literacy*. London: Routledge.
- Kress, G. and Van Leeuwen, T. (1996) *Reading Images: The Grammar of Visual Design*. London: Routledge.
- Kress, G. and Van Leeuwen, T. (2001) *Multimodal Discourse – The Modes and Media of Contemporary Communication*. London: Arnold.
- Kress, G. and Van Leeuwen, T. (2002) 'Color as a semiotic mode: Notes for a grammar of color', *Visual Communication*, 1(3): 343–368.
- Kress, G., Jewitt, C., Bourne, J., Franks, A., Hardcastle, J., Jones, K. and Reid, E. (2005) *English in Urban Classrooms – A Multimodal Perspective on Teaching and Learning*. London: Routledge.
- Lacy, M. L. (1996) *The Power of Color to Heal the Environment*. London: Rainbow Bridge Publications.
- Lupton, E. (2004) *Thinking with Type – A Critical Guide for Designers, Writers, Editors & Students*. New York: Princeton Architectural Press.
- Martinec, R. and Salway, A. (2005) 'A system for image-text relations in new (and old) media', *Visual Communication*, 4(3): 337–372.
- Martinec, R. and Van Leeuwen, T. (2008) *The Language of New Media Design*. London: Routledge.
- Matejka, L. and Titunik, I. R. (eds) (1976) *Semiotics of Art: Prague School Contributions*. Cambridge, MA: MIT Press.
- Nairn, G. (2007) *NetAlert – Protecting Australian Families Online*. Canberra: Department of Communications, Information Technology and the Arts. © 2007 Commonwealth of Australia.
- Norris, S. (2004) *Analyzing Multimodal Interaction: A Methodological Framework*. London: Routledge.
- Norris, S. and Jones, R.H. (2005) *Discourse in Action – Introducing Mediated Discourse Analysis*. London: Routledge.

- O'Halloran, K. (ed.) (2004) *Multimodal Discourse Analysis: Systemic Functional Perspectives*. London: Continuum.
- O'Toole, M. (1994) *The Language of Displayed Art*. London: Leicester University Press.
- Pittenger, R. E., Hockett, C. F. and Danehy, J. J. (1960) *The First Five Minutes*. Ithaca, NY: P. Martineau.
- Stork, D. G. (ed.) (1997) *HAL's Legacy*. Cambridge, MA: MIT Press.
- Van Leeuwen, T. (1999) *Speech, Music, Sound*. London: Palgrave.
- Van Leeuwen, T. (2002) 'Towards a semiotics of typography', *Information Design Journal*, 14(2): 139–156.
- Van Leeuwen, T. (2005) *Introducing Social Semiotics*. London: Routledge.
- Ventola, E., Charles, C. and Kaltenbacher, N. M. (eds) (2004) *Perspectives on Multimodality*. Amsterdam: John Benjamins.