

Mikko Lehtonen

MULTIMODAL INTERPLAYS

The Verbal, Visual and Auditive in Culture

Late modern culture is saturated with hybridities – textual, cultural, ethnic and also national hybridities, to name but a few. In the light of all this hybridisation it is more than clear that we in the academic world have to do our best to get rid of the rigid borderlines between traditional academic disciplines. If we do not work in this direction, the university is not likely to play any substantial cultural role in tomorrow's world. Why so? Simply because the most substantial and urgent topics for research and teaching seem to lie more and more in the academic no man's lands, that is, somewhere between the borders of, say, the humanities and the social sciences or between the borders of studies of verbal and visual symbols. In addressing these topics it is, moreover, absolutely necessary not to draw solely from one's own disciplinary traditions, since all these traditions are still largely characterised by the modern drive towards purity. In studies of different symbolic forms this can be seen in the fact that disciplines are inclined to be monomodal. They are limited to one single form of media, be that speech, writing, television or digital media. To be sure, there are multimodal fields of research and teaching, such as communication studies, but they do not try to look at the simultaneity and interaction of different media forms. Instead, they are internally divided into sub-areas characterised by media purity. Even the new studies on digital

culture have in places been organised in accordance with new medial borders, even though the digital media erodes all media differentiations.

All this, of course, reflects the culture studied in which the most highly valued modes of communication (novels, dissertations, official documents etc. without illustrations) are strictly monomodal.

Multimodality, then, is an emerging area of research in textual, media and cultural studies. If modernity was a period of monomodality and purity, multimodality is a late modern phenomenon, closely related to the increasing hybridity of our world.

Anthony Smith, a media historian and former director of British Film Institute, guides us into the everyday of multimodality of culture. In *Books to Bytes* (1993:6) he writes:

Let us consider a not unusual career for a modern work of fiction. It may begin as a novel about which an individual writer has pondered for years, or it may originate as a commission conceived by an agent or a publisher and fostered upon a writer of recognised skill. If it seems likely to sustain the investment, the finished work may be promoted, and through dextrous manipulation of the apparatus of literary review and public discussion, forced through a series of different kinds of text distribution. It will come out in hardcover and paperback, in serial fiction and digest form, and then as an even cheaper paperback. But it may also be transmuted into a set of moving images, where its basic authorship will be further dehydrated and industrialised in complex ways. A film designated for cinema distribution may in fact be shown, in widescreen format, only for further promotional purposes; the 70mm image will be seen only by a small fraction of the emerging audience, as the work passes into 35mm and 16mm gauges for distribution in various specialist systems (such as the film society network or the college circuit). It will appear in a cassette form (all the framing of the original lost in the transformation to the smaller screen) and videodisk, on cable and pay TV, ending up on 'free' over-the-air television, public or commercial. At later stages in its career, the work may return to one or more of its earlier phases, but it will remain in public consciousness with greater permanency than that bulk of Victorian fiction which failed to become one of the tiny band of classics.

Smith's description of the multiplicity of one and the same texts could be expanded further by talking about soundtracks, novels based on films or television series or the re-publishing of novels in connection with their

dramatization. In contemporary culture, this kind of recycling is becoming increasingly common.

My own interest in the problematics of multimodality was indeed aroused through such connections. As I was browsing a list of the new offerings from a Finnish book club some years ago, I made the observation that each and every book offered to the reading public had one connection or another to some other medium – in this case, films or TV series. This simple observation included an equally simple amazement regarding the fact that not one discipline or individual researcher had, to my knowledge, seized on this matter, which is familiar to all in our everyday lives.

Contrary to this academic silence, language has throughout history existed as just one mode in the totality of modes involved in the production of any text, spoken or written. A spoken text, like this one, is not just verbal but also visual, combining with “non-verbal” modes of communication such as facial expression, gesture, posture and other forms of self-representation. A written text, similarly, involves more than language: it is written on something, on some material and it is written with something. And so are the pictures drawn or otherwise produced on some material with some materials.

Now, if the fundamental symbolic forms – speech and other sounds, writing and pictures – are “always already” multimodal, then multimodality inevitably also covers the more complex symbolic forms developed on the basis of the three. Hence we could say that *multimodality characterises all symbolic forms utilized by humans*. (Kress & van Leeuwen 1996:39–40.)

People interact with the world and other people with their entire bodies and all of their senses. We can here speak of different perceptory channels and sense modalities, such as the seeing eye and the hearing ear. In Western societies it is commonplace to think that we have five senses, but this is a cultural, not a physiological fact. It is also a cultural and not a natural phenomenon that one of these sense modalities, the sight, in modern societies is given priority over the others.

If the concept of sense modalities refers to physico-cultural perceptory channels, the concept of symbolic forms refers to the general human modes of communication – speech, writing, music, pictures and combinations of these. Sense modalities and symbolic forms are linked closely to each other. The

plurality of the sense modalities is constitutive to the plurality of symbolic modes and hence to the multimodality of culture.

Yet multimodality does not stop at the level of symbolic forms. All human societies use a variety of modes of representation. Cultures are never constructed by relying solely on one form of representation. Even the oral societies have other symbolic forms at their disposal. Each of these modes has different representational potentials, different potentials for meaning-making. We can for instance say or write that the popularity of President Bush has increased or decreased, but it would be immensely difficult to communicate the same in music (even though it might be done with music and image, showing President Bush's happy or sad face and playing some hilarious or melancholic tunes). Furthermore, each mode of representation has a specific social valuation in particular social contexts. I shall soon discuss the special value attached to writing in modern Western societies. These various modes of representation are, finally, not autonomous communicational resources in a culture, nor are they deployed separately, either in representation or in communication; rather, they intermesh and interact at all times.

Each mode of representation has its potentials and limitations. These potentials of different modes of representation can be called "affordances", that is, what a certain mode can and cannot do (Kress & van Leeuwen 2001). This is also connected to the fact that some things are more easily communicated in one mode than another. For instance, speech is a mode of time and sequence whereas images most often include elements that stand simultaneously in space. In verbal language there is compact syntax but open lexis, whereas in visual language the syntax is quite open but the lexis is compact. Hence when we translate from one mode to another we have to add something that was not there but also take something away from what the first mode included but which cannot be represented in another mode.

Different forms of representation cannot be separated from each other either on the level of individual consciousness or on the level of the entire culture, on the contrary, they have an effect on each other at all times. It can be said that the forms of representation in use at a given time form a certain network that is constructed from differences and similarities.

An important part of the relations between different forms of media is the fact that the different forms of media at times conquer domains from one

another, and at times part with them. A good example of this would be what has happened to the epic narratives in Finland in the recent past. Now, modern Finnish literature, at least the most read literature, from Aleksis Kivi to Väinö Linna and Kalle Päätalo, used to be realistic epic. In the last fifteen years, however, we have seen fewer and fewer new epic novels defining who “we”, the Finns, are. But epic narration is not extinct in Finland. Literature has more or less abandoned “us” as a theme and is concentrating on different “I’s”. At the same time the question of “us” has moved into two other media forms: new epic films and equally popular epic TV-series, all chronicling distant as well as recent histories and hence taking part into the ongoing defining of “Finnishness”.

There is yet another fundamental matter entailed in the interaction of different forms of representation. Besides the fact that a certain division of labour occurs between different forms of presentations, they also constantly affect each other’s *content*. For instance the cinema, which in its time was a new mode of representation, probably did indeed in the course of time conquer terrain previously dominated by the two earlier modes of representation, the novel and the drama. Simultaneously with this, however, both the novel and drama significantly influenced the forms the cinema acquired, while the cinema left its own mark on the development of the novel and the drama (which in this regard we can call a contentual influence). It is not possible to go into greater detail about the matter here, but it is well known that the person and family-centred perspective of the classical realistic novel had its own effect on the cinema and subsequently also on TV. Indeed, there are strong similarities between the hero-centred classical realism and the language of close-ups of the television. (Lehtonen 2001a, 2001b.)

In the modern world visuality has been the privileged symbolic form, and among the visual signs writing has been most valued. Speech is thought to be natural whereas writing is considered a cultural achievement. Why so?

In case of spoken language it is easy to think that the speaker and the speech form a certain unity, that the speech is a creation of the speaker and in that sense identical with her. In the case of written and printed language the situation is different since the written language is a certain objectification which may be deemed independent of its producer. As the unity of the language and

its producer is split, the producer can examine her language and thereby herself as someone else. Hence writing enables a new kind of introspection.

But this is not the only cultural effect of writing. Learning to read requires great concentration and self-discipline. Appreciation of writing indicates a culture requiring certain amount of regulation of instincts, a society where there is a certain amount of self-repression. What else is reading in this sense but intense self-control, postponement of immediate pleasure in the hope of a future enhanced pleasure? In this sense reading might be seen as a certain technology of the self in a Foucauldian sense. (Lehtonen 2001a.)

All thinking entails observations, suppositions, conclusions, generalisations, descriptions and evaluations. Writing, however, may be considered a more or less conscious representation of these activities. We make suppositions and draw conclusions constantly. In literary language suppositions may be seen to be suppositions and conclusions to be conclusions. In this sense we may think that literary language is synonymous with being conscious of language and conscious of ourselves. Hence literary language can be seen as a form of observation directed simultaneously towards both language and ourselves. And hence literary language may be seen to be involved in producing a certain kind of a subject that is constantly observing and assessing itself.

The logic of oral cultures is largely based on structures of speech that are characterised by sequences, repetition, re-formulations, slow explication and slow elaboration. The logic of writing, on the other hand, is the logic of hierarchy, condensation and complex syntactic relations. Writing is a visual medium. In it the time and temporality of speech are replaced by space and spatiality. The logic of temporal sequence is replaced by abstract spatial hierarchies. The logic of writing is rather nominal than verbal, it is a logic of objects, not a logic of processes, a logic of abstraction, not that of specificity.

In our society writing is a medium of public social and political life, whereas speech is a medium of privacy. When a public person speaks, he has to adapt to the forms of writing, as I am of course doing here, too. For those in power there is only one language, written language.

Our culture values reading enormously. For the last 300–400 years Western scholars have equated literacy with progress. When we speak of reading we speak of modern states and modern citizenship. Since the beginning of the 19th century the cultural or aesthetic formation of citizens has been a crucial

means of producing subjects for modern nation states. The concepts of “taste” and “aesthetics” imply an idea of culture as a process of self-cultivation where such ethical subjects are moulded that have the capacity to exceed their own limitations and to see things from the perspective of a certain totality – that of the state.

Since Kant and Schiller culture has been seen as an area where we as subjects can leave behind our haphazard specific features. Culture has been seen as some kind of universal subjectivity, as a mark of the values common to all human beings on the basis of our humanity. If arts have been valued, this has been because they have been seen to represent human values in a concise form. When we read, watch and listen, it has been thought, we can leave behind our empirical selves with their social, sexual, ethnic and other variables and become universal subjects, pure humanity, so to speak.

Why are such subjects needed? The feudal, autocratic states did not require their subjects to identify themselves with the state. The modern states, however, claim that their authority originates with the people. Hence Immanuel Kant, Friedrich Schiller or Matthew Arnold did not theorise only on culture but also on such subjects that could act not only in the institutions of state but also in the civil society in the way the modern state requires.

The modern theories of culture, especially those produced within the humanities, were largely built on the opposition between culture and society. Contrary to this I would stress that culture is not in an opposition towards the society but that culture is rather the crossroads of state and civil society where modern citizenship has been formed. (See also Lloyd & Thomas 1998.) And if culture has in this sense been the religion of modern nation states, literacy has been the means by which that religion has been preached. A citizen is one who can read. Reading and writing form the basis of our culture and society by offering a model for our subjecthood. A reading subject develops in herself both sensibility and understanding, both emotionality and reason. (Lehtonen 2001a.)

Television has for some time now been the dominant medium in the late modern societies, but reading and the printed word are still valued by those in power. The anxiety concerning the future of reading is, from this point of view, an anxiety concerning the destinies of the Freudian reality principle under the pressures of the pleasure principle preached by commercial media. What is

at stake is an opposition between a citizen and a consumer, the former steered by the reality principle, the latter by the pleasure principle, the former thinking constantly of us and our future, the latter looking only at the fulfilment of its own desires in the eternal “now”.

As a concept and a field of research in textual, media and cultural studies multimodality, as said, is relatively new. The changes in late modern culture have only recently made us more receptive to this area. These cultural changes are seen perhaps most clearly in that the printed word is no longer the self-evident lord of the media landscape. The visualisation or pictorialisation of culture, increasing production of multimodal texts as well as the increase of intermedial relations between different media forms all contribute to the formation of a new landscape. Perhaps the best way to depict this change is to say we are indeed witnessing a transition from the page to the screen as the dominant form of our culture.

Yet, from the point of view of various studies on cultural texts, monomodality is still largely preferred. This, of course, reflects the culture studied in which the most highly valued modes of communication (novels, dissertations, official documents etc. that are without illustrations) are strictly monomodal. (Kress & van Leeuwen 2001.)

And still, both the textual and the cultural dimension of multimodality call into question the modern notions of the identities of texts and media forms. For a long time it was common to see texts and different media forms as identical with themselves, autonomous and full in their own terms. The concept of multimodality, however, stresses the *heteronomous* nature of texts and media forms, their dependence on stuff that is not derived from their “proper” area.

All and all, we are in a need of a different logic for looking at the whole of different media forms. We need to replace the current stress on differences and lack of connections by a careful consideration of both differences and similarities, both autonomies and interdependencies of different media forms. This does not, however, mean that multimodality is replacing identity totally with difference, that it seeks to shed any identity whatsoever of different media forms. It means rather to ask how the identities of different texts and media forms are built from and on differences.

There is, however, one major problem in outlining the phenomena of multimodality. This problem is that most theories of symbols and signs are based on linguistics and hence concentrate on speech or the written word. Consequently, there is not much theorisation on the varying possibilities and limits of different symbolic and media forms. Even less is there work on what happens when words, pictures and voices are mixed with each other. For example in film studies, the study of sound is relatively new and still largely marginal.

One of the consequences of the increasing interest in multimodality is that the concept of “medium” becomes decidedly problematic. On the one hand, the singular term, “medium”, stresses the specificity of each medium and hence foregrounds *differences* between different media forms. On the other hand, the formally plural but nowadays virtually singular “media” hints at the important dimension of *similarities* and *interaction* of different media. It is, then, a puzzling paradox that at the nominal level in cultural and media studies the plural “media” is used more often than the singular “medium”, but this preference for similarities and interaction seems to apply merely to the nominal level since in critical practice preference seems to be given to differences, not connections and similarities. In other words, while we at the nominal level speak of “media”, we nevertheless tend to put “medium” first in what we actually do.

But what kind of representation is the word “medium”? The concept of “medium” represents certain human social and cultural practices. But “medium” is, in a sense, a very curious representation of these practices. It depicts its object in a specific way, not so much as an activity or practice but as an autonomous agency, technology or tool. (See Williams 1976, s.v. “medium”.) Yet, if “medium” is seen to be an autonomous agency, technology or tool, it is easy to forget that what it performs – “mediation” – is indeed an active *relationship* that never can be reduced to the neutral transmission of messages.

In cultural praxis, however, we never see “media” as such, as they were independent of their contents. Hence, what is represented in the concept of “medium” is in actual fact always conscious practice or “practical consciousness” where both “form” and “content” are involved along with other factors such as social relations. What constitutes a “medium” is a set of relations that is necessarily unstable and prone to change. Moreover, as a result of a set of

relations, what is represented as a “medium” is never autonomous, but always heteronomous, never self-sustaining but dependent on other “media” and other social and cultural relations. Hence the concept of “medium” is always an abstraction that covers only some aspects of what is at stake in human signification. No medium is merely a vehicle through which meanings simply “flow”. Various media are socially shared and reciprocal activities, already embedded in active relationships and themselves also relationships. And a crucial part of these relationships is the relation of each media form to other forms of media.

We should not, however, confuse ontology and epistemology here. On the one hand, multimodality as a phenomenon has existed as long as there have been human cultures. It is not, then, *ontologically* a new phenomenon. On the other hand, however, *epistemologically* it is a new concept. And yet we should remember that multimodality is indeed a *concept*, that is, a way of getting a grasp on cultural reality, of constructing a new object for our studies: multimodality. From this point of view the concept of multimodality is an abstraction, but it is a real abstraction, so to speak, something that has a reference point in textual and media reality.

It is useful also in the case of multimodality indeed to think of concepts *as* abstractions. Here it is good as well to keep in mind that the word “abstraction” means literally ‘to draw something out from something’. Hence multimodality is a phenomenon that is drawn out from texts, media and cultures. It cannot be seen anywhere as such, but comes into being only after somebody gives it a name, produces the abstraction and concept of multimodality. This should also remind us that the concept of multimodality is a certain *representation*. In other words, it is a way of presenting something as something, a way of depicting the textual and media reality as a multimodal phenomenon. As such the concept is a beam of light that lets us see the object, that allows us to abstract it temporarily from its real relations for analytical purposes. Hence the concept of multimodality is not so much an answer as a question or a means of asking new kinds of questions about cultural reality.

Multimodality challenges the existing disciplinary and other borders in all studies concerning human symbolic forms and calls for interdisciplinary work. Multimodality as a new transdisciplinary research field has become visible in

studies of, for example, multimedia, the visual forms of culture, media convergence and cross-media products. In these and other fields there is a growing interest in sense modalities, symbolic modes and media forms and their differences and similarities.

What we need, then, is a unified and unifying semiotics. It is high time to develop research that takes seriously, both practically and theoretically, the new media landscape of late modernity. The search for unifying semiotics can be conducted in at least two different ways. It is possible, first, to work out detailed grammars for each and every semiotic mode, detailed accounts of what can be said with what mode and how, and then see where these different grammars overlap and where they do not. Or, secondly, we can outline a multimodal theory of communication based on an analysis of the specificities and common traits of semiotic modes which takes account of their social, cultural and historical production, of when and how the modes of production are specialised or multi-skilled, hierarchical or team-based, of when and how technologies are specialised or multi-purpose, and so on. The second variant of multimodality especially challenges the existing disciplinary and other borders in studies of language and media and calls for serious interdisciplinary work.

Last, but not least, the notion of multimodality also challenges the ways all our educational systems are working. Not only higher education but also primary and secondary education are still monomodal. Here in Finland as in all the other countries I know of we have separate classes for language and literature, other classes for visual communication and yet other classes for music. In multimodal reality there is an acute need for a reform where all these are brought together into to a new subject, multimodal communication. All the present talk about media education is a welcome *ad hoc* solution towards this reform, but what is really needed is a profound overhaul where the school becomes not a means of transferring the monomodal tradition of the 20th century to the youngsters but where it prepares the pupils for the multimodal realities of today and tomorrow.

This is why in the last instance we need go no further than to look at the mirror. The burning issue of these times is: Do we educate the future educators to educate tomorrow's adults for a monomodal or a multimodal world?

Literature

- Kress, Gunther & van Leeuwen, Theo (1996), *Reading Images*. London: Routledge.
- Kress, Gunther & van Leeuwen, Theo (2001), *Multimodal Discourse*. London: Arnold.
- Lehtonen, Mikko (2001a), *Post scriptum*. Tampere: Vastapaino.
- Lehtonen, Mikko (2001b), "On No Man's Land. Theses on Intermediality." In *Nordicom Review* 1/2001, pp. 71–83.
- Lloyd, David & Thomas, Paul (1998), *Culture and the State*. New York: Routledge.
- Smith, Anthony (1993), *Books to Bytes*. London: BFI.
- Williams, Raymond (1976), *Keywords*. London: Fontana.