

# **On the Image- Language Co-occurrence and Interrelationship in Multimodal Texts, with Particular Reference to Halliday's & Barthes' Views**

Dr. Abdul-Jabbar Musaafir  
University of Philadelphia  
Jordan

## **1. Introduction**

In the literature of this 12<sup>th</sup> Philadelphia University symposium we read that our present age is the age of image. It would not be far from the truth to say, however, that one of the most outstanding features of this age is that it is an age of signs. There is a proliferation of signs, not only in the usual broad sense of, for instance, "Danger" or "Personnel Only", but in the broad semiotic sense. There are linguistic signs of different forms; orthographic or written, and verbal or spoken signs. There are also many types of visual signs; gestures, photography, painting, cinema, theatre, television, sculpture, fashion design, dance, advertising images, icons on computer screens, architectural design, and many other images. All of these signs are used to influence our behaviour in one way or another: they are used to guide, to instruct, to constrain or to allow, to protect or to warn, etc.

These linguistic and visual signs may be used in isolation, or they may be used in combination. This potential for a multiplicity and variety of signs, in combination with the growth and role of the internet and multimedia in what has been characterised as the "information revolution", has raised consciousness of the visual sign as a conveyor of meaning, so much so that there is an increasing interest amongst educationists of a possible need for developing students' abilities in visual literacy, rather than simply literacy. The co-occurrence of and interrelationship between visual and linguistic signs in multimodal texts is the issue that is addressed in this paper.

It is generally agreed upon that in most developed societies or cultures, language, in its spoken and written forms, has been generally viewed as the dominant, and indeed, the superior mode of communication. On the other hand, image or the visual sign is generally seen as being subordinate to the spoken or written sign. The dominance of the linguistic, and the dominance of the written over the spoken as the more "advanced" mode, is realised across various cultures in and through educational practices and attendant value systems (see McLuhan 1964; Ong 1982).

However, the rate and degree of change that has been brought about by the increasingly computerised and multimedia-based modes of communication in this age of signs may well form a challenge to the traditional dominance of language over image. This may be supported by the generally accepted observation that the young generation of recent times does not read but watch. In addition, workers in the world of film industry are advised not to "tell them but show them." This is, one may readily note, a growing challenge. For, the fact remains that the linguistic mode dominates; it dominates in the educational systems, and it dominates in the print media, whether it be academic or journalistic, whether the texts therein utilise only the linguistic or written mode, or whether they combine it with some kind of visual

representation. However, the question of where one mode is more dominant than the other, or whether a particular verbal text can be understood with or without the visual mode, will not be investigated deeply in this paper.

## **2. Linguistics Perspective**

While the central focus of the discipline of linguistics throughout the twentieth century has been on the study of natural language, either in structural, sociolinguistic, psycholinguistic or functional terms, some of the trends of the latter decades of the century have reflected an increasing interest in forms of communication other than the linguistic ones and the ways that they project and organise their meanings. This interest has been supported not only by the continual development and enrichment of semiotic theory, but also by the rapid changes in the modes of communication brought on by the revolution in communication technology. The result is an increase in the publication of studies investigating, for example, meaning projections in digital imagery (Marchese 1995), the development of automated presentation design, intelligent multimedia interfaces and associated architectural and modelling issues (e.g. Maybury 1993). Jenks (1995) is an instance of an expanding debate about the 'visual' character of contemporary culture.

However, it has also been recognised that developments in general linguistic theory could inform the interpretation of other communication modes besides language. Researchers utilising concepts from linguistic theory to examine non-linguistic modes have included the Prague School of the 1930's and 1940's, the Paris School of the 1960's and 1970's (see Nöth 1990 for a good overview), and more recently what has been loosely characterised by Kress and van Leeuwen as "the fledgling movement" of "Social Semiotics" (1996:5). This latter paradigm utilises Systemic Functional Grammar theory, developed by M.A.K. Halliday (1978, 1985, and 1994).

Halliday's linguistic theory interprets language as a socially-based semiotic system, and applications of it to non-linguistic forms of communication have in recent times been gaining prominence through research on the visual semiotics of displayed art forms such as sculpture, architecture and painting by O'Toole (1994, 1995), and the proposal of a 'grammar' of visual design in images in general as well as in educational contexts by Kress and van Leeuwen (1990, 1996).

Various definitions of 'text' have been provided by applied and theoretical linguistic researchers (see for example Stubbs 1983, Brown and Yule 1983, van Dijk 1977, Halliday & Hasan 1976 etc.). For the purposes of this paper, however, we need only refer to a 'text' in the term 'multimodal text' according to Halliday and Hasan. They define it as "a semantic unit: not of form, but of meaning" (1976:1-2), and that it is functional, or "language that is doing some job in some context" (1985:10). "Cohesion" of a text is also viewed in terms of meaning, being the "relations of meanings that exist within the text, and that define it as a text" (1976:4). Although Halliday and Hasan are referring primarily to language in its spoken or written forms, they remind us of the fact that their view of text as contextualised meaning and function permits the consideration of other modes of meaning-making. Thus, a text "may be either spoken or written, or indeed any other medium of expression that we may like to think of." (1985:10).

### **3. Barthes' Views on Image-Language Relations:**

Barthes' work on the photographic image, fashion, architecture, painting, film, and advertising has made him the most prominent and cited member of the Paris school (Nöth: 1990). Barthes' contribution bears some consideration, since his analysis of image-text relations is directly relevant to the focus of this paper.

Roland Barthes [1915-1980], as a follower of Saussure's semiological thought and a leading structuralist is recognised for his contribution to text semiotic studies of myth, theology, literature and narrative. He became famous also for his contribution to the semiotics of various types of visual communication, and for his systematic model of *signification*, a Saussurean term which Barthes defines as "a process: it is the act which binds the signifier and signified, an act whose product is the sign" (1964/67:48). Barthes set up two orders of *signification*, which he referred to as *denotation* and *connotation*. His particular usage of these terms derives directly from Saussure's dichotomy (signifier-signified) model of the sign, and as Nöth (1990:310) quite rightly notes from a "greatly simplified version of the glossematic sign model" developed by Hjelmslev. They refer to the first and second levels or systems of meaning in a sign. Denotation is the literal (or cognitive) meaning of a sign, in terms of what is "objectively" observable or which is easily recognised and identified. For example, an image of an eagle denotes a bird, or type of bird. Connotation, on the other hand involves meanings which are dependent on the denotative level. It is the symbolic level. A sign is interpreted by a viewer or reader connotatively when the interpretation goes beyond the literal (denotative) interpretation via the activation of context-dependent conventions or codes. An image of an eagle (a symbol), depending on the context, can connote a football team (Manly Sea Eagles rugby club), danger or a need to warn or protect (as in advising tourists/campers in a national park), or indeed, spiritual superiority as in many ancient religions and cultures (Saladin, Arabism?). Barthes explains the operation of these concepts by stating that "the first system becomes the plane of expression or signifier of the second system ..... the signifiers of connotation ..... are made up of signs of the denoted system" (1967:91). Viewers of signs can make sense of or integrate these two levels only by making reference to their background knowledge of the culturally-based codes and various associative meanings; the system of *connotation* is dependent on this background knowledge for its existence. Thus, in Saussurean terms the eagle example above is a *signified* which can become the *signifier* (the means or vehicle) of a further *signified* at another level.

One must note the importance of context, and therefore ideology here, for when readers or viewers engage with verbal or visual signs, their interpretation is the result of the background cultural codes they introduce to unify the *signifier* and *signified*. Barthes clarifies these concepts in his treatment of the semiotics of publicity images, where he discusses the levels of meanings in advertising images. The *denotative* level of meaning in advertising images is referred to as a noncoded iconic message, while the *connotative* level is termed a coded iconic or symbolic message, which is grounded on underlying pragmatic, cultural, patriotic, historic or aesthetic background knowledge. For Barthes, the advertising photograph (and image) exists in an ideological world, a view which Dyer (1982:129-130) elaborates:

Ads, as a means of representation and meaning, construct ideology within themselves through the intervention of external codes which are located within society. The ad will use images, notions, concepts, myths, etc. already available in the culture. An ad does not simply reflect ideology; it reworks it, thus producing new meanings. It uses objects which are signifieds of ideological systems and thought that already exist and then makes them signifiers of another structure (the ad). Its connotational process depends on our knowledge of the forms of ideology that advertisements employ.

In his study of press photographs and advertisements Barthes (1977) discusses a third level of meaning, besides the denotative and connotative. This level of meaning is of significance to the study of multimodal texts because it is connected to the way that ideology operates through the linguistic message that may or may not accompany the image. Barthes raises questions about the nature of the functions of the linguistic message with regard to the denotative and connotative aspects of the iconic message. He states that there are two such

functions: *anchorage* and *relay*, a dichotomy developed by Barthes in an attempt to address the question of the contextual relationship between images and verbal text (op.cit:38).

In his explanation of *anchorage*, Barthes is of the opinion that there is a need for the meaning of images to be always related to, or in some way dependent on the linguistic text. Images without the support of language are seen as being too open to a variety of meanings, having an inherent indefiniteness which necessitates a reliance on language to provide a more substantive interpretation. The main reason behind that is that all images are polysemous; they imply, underlining their signifiers, a 'floating chain' of signifieds, the reader able to choose some and ignore others. .... Hence in every society various techniques are developed intended to fix the floating chain of signifieds in such a way as to counter the terror of uncertain signs; the linguistic message is one of these techniques (op.cit:38-39).

The function of the linguistic message then, whether it is in the form of a caption, heading, headline or accompanying reportage or prose, is to fix or 'anchor' the various possible meanings, directing the reader's interpretations and settling possible visual ambiguities and contradictions. The verbal text does this at the denotative level by answering "the question: *what is it?* The text helps to identify purely and simply the elements of the scene and the scene itself; it is a matter of a denoted description of the image." (op.cit:39). The verbal text also does this at the connotative, symbolic level where ideology is operating in a significant way. Here the text *directs* the reader through the signifieds of the image, causing him to avoid some and receive others ..... it remote-controls him towards a meaning chosen in advance. In all these cases of anchorage, language clearly has a function of elucidation, but this elucidation is selective, a metalanguage applied not to the totality of the iconic message but only to certain of its signs. The text is indeed the creator's (and hence society's) right of inspection over the image; anchorage is a control, bearing a responsibility ..... for the use of the message. (op.cit:40)

Therefore, *anchorage*, in terms of an interpretation of the total message of the language-image text, involves a situation of visual-verbal dependency, whereby the verbal text links the image and the situation in space and time, a relation which cannot be established through purely visual means of expression. In addition, the verbal text selects for the viewer or the reader one of several possible interpretations of the image, thereby fixing "the floating chain of signifieds" (op.cit:39).

Barthes' view of *relay* however is that "text and image stand in a complementary relationship; the words, in the same way as the images, are fragments of a more general syntagm and the unity of the message is realised at a higher level" (op.cit:41). Thus *relay*, in terms of an interpretation of the total message of the language- image text, involves a complementary relationship between both the verbal and image constituents — it denotes a reciprocal association, such that each mode (linguistic and visual) contributes to the overall message projected. Barthes claims that this is rare in fixed images, but most common in modes such as in the speech balloons used in comics, and those involving moving images as in TV or film where the (linguistic) dialogue works in concert with the images. However, the two language-picture relations should not be viewed as mutually exclusive. There are obviously multimodal texts containing both images and verbal text where either the primary relationship is one of anchorage, with some element of relay, or one of relay with some element of anchorage (cf: Vestergaard and Schröder, 1985:33-36).

Nöth (1990:454) develops this point in a brief, general overview of research into picture-text relationships in which the anchorage-relay dichotomy does not obtain, or in which anchorage and/or relay could be considered as sub-types of other higher-order relations. These relationships are listed below with some examples, and it can be suggested that these may apply equally well to other forms of images such as those that are used in graphs, tables and charts, as well as the photographic image or picture:

1. **ILLUSTRATION:** Here the image is subordinated in terms of its semiotic function to the verbal text which it illustrates. This is exemplified most commonly in the illustrations used in fiction - some editions of literary works are produced with illustrations, while others of the same work are not.
2. **PICTORIAL EXEMPLIFICATION:** This is a sub-type of relay, and unlike illustrations, the images in pictorial exemplification “retain a higher degree of informational autonomy in relation to the verbal message they exemplify visually”. A common example would include the photographs provided in encyclopedias to exemplify the type of phenomena being described or discussed.
3. **LABELING:** This involves the affixing of a label to a picture, and may be an instance of anchorage (as in the title/name of a portrait or work of art), or as a type of relay (as in the name given for a person in a portrait or photograph).
4. **MUTUAL DETERMINATION:** This is a mixture of relay and anchorage; the interpretation of the picture is dependent upon the verbal text, but the picture is equally as necessary for the interpretation of the verbal aspect. This is very often seen in advertisements, where the two modes are combined via a verbally posed question of some sort.
5. **CONTRADICTION:** This is rare, but involves the presentation of opposite or contrasting messages by both the language and the picture. This has perhaps become more common in recent times with radical graphic design techniques becoming more in vogue, as evidenced by magazines which are constructed in ways which purposely breach compositional conventions for the effects it can create. For example the pop magazine RAYGUN refers to itself as: “THE BIBLE OF MUSIC + STYLE AND THE END OF PRINT” (sic).

It can be seen from the above that Barthes was concerned with answering the question: “Does the image duplicate certain of the information given in the text by a phenomenon of redundancy or does the text add a fresh information to the image?” (1977:38). However, this is a simplification which Nöth suggests does not capture the fact that “the juxtaposition of picture and word usually results in a new holistic interpretation of the scripto-pictorial or the audio-visual message” (1990:453). Here it is not simply a question of a text-image relation of addition and duplication, but of the ways in which both of the modes work together to produce a unified, coherent, multimodal text. It is thus a question of the nature of intersemiotic complementarity.

#### **4. Conclusion**

In our daily life language and image fight to get our attention. In some situations both, language and image, share to a certain degree, the field. In many other situations, such as in sport reports and advertisement, however, language seems to have lost grounds to image. Language seems no longer the suitable means to attract our attention and ultimately, to influence our reaction and behaviour. Thus, the saying that a picture paints a thousand words is accepted without counter argumentation. The question is where language would work better than image in all its types and kinds. This question may be answered by examining certain situations where the linguistic text rather than image that has the dominance and can work in isolation. Such examination may show that, through the use of certain syntactic ‘moves’, language can have a great impact on changing human behaviour as in the writing of certain brochures and pamphlets. In such cases, this change of behaviour may well, otherwise, need thousands of images to realise.

#### **Bibliography**

- Barthes, R. (1964/67) Elements of Semiology. London: Cape.  
 Barthes, R. (1977) Image, Music, Text. New York: The Noonday Press.  
 Brown, G. and Yule, G. (1983) Discourse Analysis. New York: Cambridge University Press  
 van Dijk, T. A. (1977) Text and Context. London: Longman

- Dyer, G. (1982) Advertising as Communication. London: Methuen.
- Halliday, M.A.K. (1976) System and Function in Language. (ed. G. R. Kress) London: Oxford University Press.
- Halliday, M.A.K. (1978) Language as Social Semiotic. London: Edward Arnold.
- Halliday, M.A.K. (1985) An Introduction to Functional Grammar. London: Edward Arnold.
- Halliday, M.A.K. (1994) An Introduction to Functional Grammar. 2nd ed. London: Edward Arnold.
- Halliday, M.A.K. and R. Hasan. (1976) Cohesion in English. London: Longman.
- Halliday, M.A.K. & R. Hasan, (1985) Language, Context and Text: Aspects of language in a social-semiotic perspective. Oxford: Oxford University Press.
- Jenks, C. ed. (1995) Visual Culture. London: Routledge.
- Kress, G. and R. Hodge (1979) Language as Ideology. London: Routledge and Kegan Paul.
- Kress, G and T. van Leeuwen, (1990) Reading Images. Geelong, Victoria: Deakin University Press.
- Kress, G and T. van Leeuwen, (1996) Reading Images: The Grammar of Visual Design. London: Routledge.
- McLuhan, M (1964) Understanding Media: The extensions of man. New York: Signet Books.
- Marchese, F. T. ed. (1995) Understanding Images: Finding meaning in digital imagery. Santa Clara, California:TELOS - The Electronic Library of Science, Springer-Verlag Publishers.
- Maybury, M. T. ed. (1993) Intelligent Multimedia Interfaces. London: AAAI Press/The MIT Press.
- Nöth, W (1995) Handbook of Semiotics. Bloomington: Indiana University Press.
- Ong, W. (1982) Orality and Literacy: The Technologizing of the Word. (New Accents Series) London: Methuen.
- O'Toole, M (1994) The Language of Displayed Art. London: Leicester University Press.
- O'Toole, M. (1995) "A Systemic-Functional Semiotics of Art", in Peter Fries and Michael Gregory, Discourse in Society: Systemic functional perspectives. Vol. L in the series Advances in Discourse. Norwood, New Jersey: Ablex Publishing Corporation, 159-179.
- Saussure, Ferdinand de. (1916/1966) Course in General Linguistics. New York: McGraw-Hill Paperbacks.
- Stubbs, M (1983) Discourse Analysis. Chicago: University of Chicago Press.
- Vestergaard, T. and K. Schrøder (1985) The Language of Advertising. Oxford: Blackwell.