

O'Halloran, K. L. (in press). Multimodal Analysis and Digital Technology. In A. Baldry and E. Montagna (Eds.), *Interdisciplinary Perspectives on Multimodality: Theory and Practice*. Campobasso: Palladino.

Multimodal Analysis and Digital Technology

Kay L. O'Halloran
Department of English Language and Literature
National University of Singapore

Multimodal Analysis Laboratory
Interactive Digital Media Institute (IDMI)
National University of Singapore

“Assessments of change, dynamics, and cause and effect are at the heart of thinking and explanation. To understand is to know *what cause provokes what effect, by what means, at what rate*. How then is such knowledge to be represented?” (Tufte, 1997: 10)

The aim of this paper is to explore possible directions for advancing multimodal analysis through the use of digital technology. Historically, the impact of technology on knowledge is evident from the introduction of the printing press in the late fifteenth century, which led to the close study and subsequent development of mathematics and science (Eisenstein, 1979). Today, “computerization turns media into computer data” (Manovich, 2001: 45). The recursive nature of the computerisation process, where “*images scatter into data, data gathers into images*”, (Galison, 2002: 322) permits the visual display of data flow which is experienced over time, “be it in science, statistics, architecture, design, digital art, or any combination of these” (Paul, 2003: 175). There are few constants in the digital environment where everything is variable and open to manipulation and recontextualisation. However,

does digital technology function to reproduce more effectively and efficiently knowledge which already exists, or does it produce new theories and approaches in disciplines such as mathematics and science? How does digital technology function to constrain knowledge? What are the implications for multimodal analysis? In what follows, these questions are explored. The theoretical basis for the discussion is Michael Halliday's social semiotic theory (Baldry & Thibault, 2006; Halliday, 1978, 1994 [1985]; Halliday and Matthiessen, 2004; van Leeuwen 2005).

The paper unfolds in the following manner. First, the social semiotic approach to multimodal discourse analysis (MDA) is introduced. Second, the relationship between technology, mathematics and science is explored, and the impact of semiotic transcoding (i.e. symbolic data ↔ visual images) in digital media is investigated. Last, the use of digital technology, visual reasoning and multimedia analysis for multimodal analysis are considered.

Social Semiotic Approach to Multimodal Discourse Analysis (MDA)

Social semiotics is concerned with “the way people use semiotic ‘resources’ both to produce communicative artefacts and events and to interpret them ... in the context of specific social situations and practices” (van Leeuwen, 2005: preface). The multimodal social semiotic approach draws upon Michael Halliday's (1978, 1994 [1985]; Halliday and Matthiessen, 2004) systemic functional (SF) theory to provide frameworks for conceptualising the complex array of semiotic resources which are used to create meaning (e.g. language, visual imagery, gesture, sound, music, three dimensional objects and architecture) and detailed practices for analysing the meaning arising from the integrated use of those resources in communicative artifacts (i.e. texts) and events.

The notion of ‘semiotic resource’ in multimodal social semiotics is critical. Van Leeuwen (2005: 3) explains “[i]t originated in the work of [Michael] Halliday who argued that the grammar of a language is not a code, not a set of rules for producing correct sentences, but a ‘resource for making meanings’ (1978: 192)”. Halliday (1994 [1985]) identifies four types of meaning potential, which he calls the metafunctions of language: (a) experiential meaning for constructing experience; (b) logical meaning for

establishing logical relations; (c) interpersonal meaning for enacting social relations; and (d) textual meaning for organising the message. Halliday's (1994 [1985]; Halliday and Matthiessen, 2004; Martin, 1992; Martin and Rose, 2003) SF theory provides a description of the metafunctional-based phonological, grammatical and discourse systems which are used for the analysis of linguistic texts and events.

The SF approach to MDA (henceforth SF-MDA), first introduced in Michael O'Toole's (1994) *The Language of Displayed Art* and Kress and Gunther Kress and Theo van Leeuwen's (1996) *Reading Images: The Grammar of Visual Design*, is a rapidly expanding area of research (e.g. this volume; Jewitt, forthcoming). Halliday's SF theory has been extended to semiotic resources which include visual images (O'Toole, 1994; Kress and van Leeuwen, 2006 [1996]), mathematical symbolism and images (O'Halloran, 2005), music and sound (van Leeuwen, 1999), movement and gesture (Martinec, 2000, 2001) and architecture and space (O'Toole, 1994; Pang, 2004; Stenglin, 2004). The multimodal analysis of print texts (e.g. Baldry and Thibault, 2006; Kress and van Leeuwen, 2006 [1996]; Lemke, 1998; Martin, 2002; O'Halloran, 2004a, 2005; Royce and Bowcher, 2006; van Leeuwen, 2005; Ventola et al., 2004), video texts and internet sites (e.g. Baldry and Thibault, 2006; Djonov, 2006, 2007; Iedema, 2001; Lemke, 2002; O'Halloran, 2004a) and 3-D sites (e.g. Pang, 2004; Ravelli, 2000; Stenglin, 2007) followed. The research field has collectively been called 'multimodality', where 'multimodal' typically refers to the multiple modes (e.g. spoken, written, printed and digital media, embodied action, and 3-D material objects and sites) through which social semiosis takes place. Halliday's SF theory provides a comprehensive framework for MDA because the metafunctional principle provides an integrating platform for multimodal theory and practice.

Djonov (2006) contains a comprehensive description and interpretation of the major developments in the evolution of the SF-MDA approach. Recent developments include multimodal transcription methodologies and 'cluster analysis' involving scalar hierarchies of multimodal items which interact across different levels to construct meaning (Baldry and Thibault, 2006); corpus-based approaches to multimodal analysis (Baldry, 2004; Baldry and Thibault, 2005; Baldry and Thibault, forthcoming; Bateman, et. al, 2004); the concept of semiotic metaphor and the formulation of metafunctionally-based inter-semiotic systems to help explain the expansion of

meaning which takes place in multimodal discourse (e.g. Lim, 2004; O'Halloran, 2005; 2007). Despite these advances, video texts and interactive digital sites remain a major challenge for SF-MDA discourse analysts (e.g. Djonov, 2006; Jones, 2007), largely due to the complexity of SF-MDA analysis, the dynamic nature of digital texts and the sheer amount of data which is generated. Technological advancement appears to have outpaced the development of SF-MDA theory and practice required for multimodal analysis of digital media texts and events.

Digital technology provides a common platform for semiotic resources to combine and unfold in new and innovative ways. Therefore, digital technology is multimodal social semiotic technology. However, how is digital technology actually impacting on disciplines which embrace the new technology (e.g. mathematics and science), and how can we use digital technology for SF-MDA, especially for the analysis of video texts and interactive digital sites? From this vantage point, SF-MDA offers a comprehensive approach to understanding and analysing the functionality and constraints of digital media technology.

In what follows, it is proposed that digital technology is leading to new theories and practices in mathematics and science, and that it has the potential to advance SF-MDA. At present, Anthony Baldry and Michele Beltrami's (2005) *Multimodal Corpus Authoring (MCA)* system is the only software application for the multimodal analysis of video texts (<http://mca.unipv.it>). From this promising start, the use of digital technology for SF-MDA is further explored in this paper.

Mathematics, Science and Digital Technology

The impact of technology on the development of mathematics and scientific knowledge has been documented (e.g. Colonna, 1994; Davis, 2006; Eisenstein, 1979; Galison, 2002; O'Halloran, 2005; Swetz, 1987). The printing press, for example, explains the increased popularity of the Hindu-Arabic numerical system in early printed arithmetic books (Swetz, 1987). Manual forms of computing (e.g. the abacus and counters) and visual representations of these computational strategies could not compete with the more efficient Hindu-Arabic system. The printed mathematical texts contributed to the close study and development of arithmetical algorithms and the

standardisation of mathematical procedures, techniques and symbols which, in effect, paved the way for the development of symbolic algebra. Eisenstein (1979: 467) explains, “[c]ounting on one’s fingers or even using an abacus did not encourage the invention of Cartesian coordinates”. Swetz (1987: 284) states “[p]rinting ... forced a standardization of mathematical terms, symbols, and concepts. The way was now opened for even greater computational advances and the movement from a rhetorical algebra to a symbolic one”. The integrative use of mathematical symbolism, visual representations and language provided the semiotic tools for the scientific revolution, where the world was rewritten in terms of patterns and order (Stewart, 1995).

It appears that digital technology is having a major impact on mathematics and science today. The influence of computers is such that they have given rise to “a new world view which regards the physical world not as a set of geometrical harmonies, nor as a machine, but as a *computational process*” (Davies, 1990: 23-24). One outstanding feature of the shift to computation is the ability of computers to generate, represent and manipulate numerical results as dynamic visual patterns which unfold over time (Rabinovich et al., 2000; Walgraef, 1997). This has led to the development of sophisticated visualisation processes and computer graphics software. These two developments are considered in turn.

Visualisation processes are concerned with “the transformation of numerical data from experiments or simulation [via mathematical models] into visual information” (Grave and Le Lous, 1994: 12). Visualisation processes play an important role in mathematics and science in areas which involve the study of complexity and non-linear dynamical systems, for example (Dalenoot, 1994; Mikhailov and Calenbuhr, 2002; Peak and Frame, 1994; Ragsdell and Wilby, 2001; Smith, 1998). The traditional scientific approach involves finding analytical (i.e. symbolic) solutions to the model equations constructed from experimental data in order to explain the behaviour of physical systems. However, simple chaotic systems are difficult to capture analytically. As a result, the computerised numerical approach entails generating and visualising numerical solutions. “Virtual experimentation, therefore, assumes that the underlying analytical models of the simulation are correct (much as we presume that a real object is ‘correct’) and explores the range of behaviours produced by that model ...” (Colonna, 1994: 183-4). The researcher can change the parameters or request

different views of the numerical results to explore the behaviour of the system. “The researcher, the numerical calculations, and the displayed images form a feedback loop through which the complex behaviour of the simulated system is explored” (Colonna, 1994: 184). The means for exploring the behaviour of the system remains multisemiotic (i.e. symbolic, visual and linguistic). However, the meaning potential of those semiotic resources are expanded in the dynamic realm of computer-based visualization (e.g. colour, texture, size, shape, point of view, and so forth).

In addition, the shift to computation has witnessed the rapid development of computer graphics applications which are concerned with “the pictorial synthesis of real or imaginary objects from their computer based models”. Computer graphics are used in areas which include virtual reality, animation, interactive user interfaces and ‘image processing’ which deals with “the management, coding, and manipulation of images, the analysis of scenes, and the reconstruction of 3D objects from their 2D projectional presentations” (Groß, 1994: 2). Colonna (1994) sees the pictorial synthesis of real and imagined objects as a scientific tool where global patterns, rather than traditional point-by-point descriptions, are generated. Complex forms become distinct, and hidden orders in the numerical results become ascertainable. Computer graphics are used for a range of functions; e.g. the visualisation of large data sets, the creation of three-dimensional objects, the construction of view-dependent visualisations, the production of multi-dimensional images of motion and the visualisation of reconstructions for research in applied mathematics, engineering, physics, geology, medicine, biology, architecture and design, to name but a few (Moorhead et al., 2002).

The use of visualisation processes and computer graphics in mathematics and science for the interpretation of complex data sets relate to human capabilities of seeing visual patterns. “Because visual analysis techniques are particularly well suited to the human cognitive capabilities, more emphasis has been placed on visual analysis tools for understanding computer simulations of complex phenomena” (Watson and Walatka, 1994: 7). Humans cannot process the information at the same rate when presented with the symbolic data generated by supercomputer simulations or high-powered scientific instruments. Colonna (1994: 184) explains that “[v]ision is the most highly developed of our human senses for reception, isolation and understanding

of information about our environment. Vision provides a global perception of coloured shapes against a changing, moving, and noise-filled background. The idea of using the eye as the main tool in the analysis of numerical results is therefore quite natural”.

“Scientific visualization is the art of making the unseen visible” (Colonna, 1994: 184).

This statement echoes Newton’s efforts at making the invisible ‘visible’ through mathematical description (Barry, 1996). For Newton, the semiotic resource which could be easily rearranged to solve problems in print format was symbolic mathematical notation. Mathematical and scientific graphs and diagrams were expensive and time-consuming to produce in print format, thus limiting their usefulness. Today, computer-generated dynamic visual images can be re-arranged, juxtaposed, connected and transformed with minimal effort. The visual image is the semiotic resource increasingly being used to establish results, as observed by mathematician Philip Davis (2006: 147-161) in *The Decline and Resurgence of the Visual in Mathematics*. Systems of meaning such as colour, shading, brightness and texture are increasingly playing an important role in mathematics and science (Danaher, 2001; Levkowitz, 1997). “Like the various components of a musical orchestra which all come together to create a useful whole, the scientist, numerical calculation, and picture synthesis all work together to form a scientific instrument...” (Colonna, 1994: 184-5).

Has digital technology led to new theories and approaches in mathematics and science? Davis (2006) explains “[i]t may be that the important question is not whether visuals lead to new results in traditional symbolically presented mathematics, but whether visuals lead to a corpus of significant mathematical experiences that carry their semiotic content and may or may not be connected with traditional material”. In response, it appears that digital technology *has* created new paradigms of research in mathematics and science. Galison (1997: 689), for example, explains “[w]ithout the computer-based simulation, the material culture of late-twentieth-century microphysics is not merely inconvenienced - it does not exist”. The same can be said for other new paradigms of research, such as dynamical systems theory in applied mathematics and physics. Mathematics and science are multisemiotic enterprises; they *do* rely on symbolic, visual and linguistic forms of semiosis. However, ‘the scientist, numerical calculation, and picture synthesis’ (Colonna, 1994: 184-5) work together to

create new theories and approaches which clearly extend beyond those which exist in print format. The move to the computational realm has had a profound effect on mathematics and science.

Mathematics and scientific knowledge is tied to technology, which historically has been limited to the pen, paper, the printing press and 3-D mathematical models. Today, digital technology has advanced knowledge in mathematics and science. The challenge remains for us to advance multimodal analysis beyond the limitations of page-based approaches. If “[a]ssessments of change, dynamics, and cause and effect are at the heart of thinking and explanation” (Tufte, 1997: 10), then dynamic visual imagery which can be juxtaposed, transformed and generally manipulated in infinite manner of ways would appear to have much to offer in terms of modelling the temporal unfolding of the integrative meta-functionality of semiotic choices in dynamic multimodal texts in order to extract the semantic patterns which subsequently arise. The limitations of current approaches to SF-MDA are discussed before investigating the use of digital technology for advancing multimodal analysis.

Contemporary Approaches to SF-MDA

The humanities are far behind mathematics and the sciences with regards to advancing knowledge through the use of digital technology. SF-MDA is no exception, despite the fact that digital technology provides the means to move beyond the limitations of page-based approaches because it provides a common dynamic platform for the integration of audio, visual and motion analyses. Ironically, interest in multimodal analysis is accelerating amongst computer scientists who are attempting to develop new techniques for archiving, searching and retrieving information from visual images, the internet and multimedia data bases. Indeed, research teams are desperately trying to develop effective means for managing and extracting information from the digital texts which are proliferating through the rapid advance of digital technology. Multimodal social semiotic theory provides a rich theoretical basis for these investigations which are concerned with multimedia representation and data analysis, storage and retrieval. Therefore, the time seems right to merge multimodal social semiotic theory with computer-based multimedia techniques of analysis. The

limitations of existing approaches to SF-MDA are explored below in order to substantiate this proposal.

The representation of SF-MDA system networks has taken several forms. The most typical form is the classificatory-type modelling approach developed for language (e.g. Baldry and Thibault, 2006: 196; Kress and van Leeuwen, 1996: 107; Van Leeuwen, 2005: 13), where system networks consist of a series of discrete options formulated as 'x and y' and 'x₁, x₂, ... x_n' and 'y₁, y₂, ... y_n'. Lines which are used to indicate clines (Kress and van Leeuwen, 1996: 165) are integrated within system networks to represent degrees of variation (e.g. Cheong, 2004: 193; Muntigl, 2004: 44; Pang, 2004: 41). In addition, system choices have been represented using various forms of symbolic notation (e.g. Baldry, 2004: 86; Baldry and Thibault, 2006: 215). While these forms of representations have proved productive, it is recognised that they are not always appropriate for modelling the systems which come into play in multimodal discourse. For example, systems such as colour, size, angle, framing, perspective and light are difficult to classify as discrete options and linear clines. The representation of SF-MDA system networks requires further development to effectively portray the options which are available and the choices which are made in multimodal texts and events, a problem which is further considered below.

SF- MDA is concerned with displaying and analysing the integrative meanings arising from semiotic choices which combine in dynamic ever-changing patterns. These patterns are conceptualised as continuous spatial-temporal-type relations, unfolding much like an orchestral score (van Leeuwen, 2005) with phases and transitions (Baldry, 2004; Baldry and Thibault, 2006; Tan, 2005; Thibault, 2000; van Leeuwen, 2005). However, the practices for representing, transcribing and analysing dynamic multimodal texts tend to rely on static images which have been annotated with vectors and linguistic descriptions to display interactions between items (e.g. Baldry and Thibault, 2006; van Leeuwen 1996) and movement (e.g. Kendon, 2004). In addition, SF-MDA transcription methods for video texts include sophisticated transcription tables with visual frames extracted from the video text. The analysis is coded using a combination of linguistic descriptors, symbolic notations and abbreviations (e.g. Baldry and Thibault, 2006; Djonov, 2006; Jones, 2007; van Leeuwen, 2005; Tan, 2005).

Page-based methodologies, however, severely limit the analyst's ability to display, describe and analyse the complexity of the multifunctional interplay of semiotic choices, especially in the case of the video texts and interactive digital sites. The data is too complex, and furthermore, the data is dynamic which makes page-based transcription methods extremely difficult and time-consuming to produce. Moreover, the shift from dynamic forms of semiosis to static forms of SF-MDA analysis may (necessarily) be reductive with regards to the complexity of the integrative semantics which take place in the dynamic multimodal environments. While SF-MDA analysts openly acknowledge the continuous dynamic nature of multimodal social semiosis and the inherent difficulties of page-based transcription methods, a fully-fledged methodology is yet to be developed in a dynamic multimodal environment. Such an approach requires the use of digital technology which replicates the meaning potential of what is being analysed. The challenge of developing digital technology for SF-MDA has been faced by Anthony Baldry, Michelle Beltrami, Paul Thibault and a team of Italian researchers who developed the *Multimodal Concordancer Authoring (MCA)* software program, a research initiative which is outstanding in terms of the foresight it demonstrates.

Baldry and Beltrami's (2005; Baldry, 2004; Baldry and Thibault, 2006) MCA is an on-line multimodal concordancer which permits the video texts to be played while the SF-MDA is coded. The analytical procedures in MCA involve selecting an option from a text-based grammar, or alternatively, entering a linguistic description of what is seen in the video text. The resulting relational databases may be searched to analyse patterns across multimodal corpora, leading to the development of multimodal corpus-based approaches to text analysis (Baldry and Thibault, 2001; Baldry and Thibault, forthcoming). At this stage of innovation, however, MCA is based on traditional text-based coding devices, which take the form of discrete system networks and text-based descriptors (see Figure 1).

▶	Text 20: Vauxhall (b)	665	695
DRIVER		Text 20: YES WOMAN DRIVER	
STORYTELLER		Text 20: YES participant's voice male voice at the end	
▶	Text 21: Vauxhall (c)	696	725
DRIVER		TEXT 21: YES MALE DRIVER	
STORYTELLER		Text 21: YES+sound track+male voice at the end	
▶	Text 22: Toyota (b)	726	755
DRIVER		TEXT 22: YES IMPLIED DRIVER man seen before and after drive phase	
STORYTELLER		Text 22: YES+sound track+male voice	
▶	Text 23: Proton	756	785
DRIVER		TEXT 23: YES IMPLIED DRIVER	
STORYTELLER		TEXT 23: YES male voice soundtrack	
▶	Text 24: Volvo S60	785	815
DRIVER		TEXT 24: YES IMPLIED DRIVER	
STORYTELLER		Text 24: YES soundtrack female voice at the end	
▶	Text 25: Fiat Punto	815	855
DRIVER		Text 25: YES WOMAN DRIVER then MALE DRIVER	
STORYTELLER		TEXT 25: YES participants (male + female) voices + female voice at the end	

Figure 1 Relational Data Base Search using MCA (Baldry, 2004: 98)

The potential for developing interactive digital platforms for mapping, analysing and retrieving the patterns of inter-semiotic relations in the dynamic environment of digital technology requires further exploration. For example, SF-MDA may involve the visual display of patterns which track the metafunctional orchestration of semiotic choices over time. Interactive digital platforms consisting of multiple tracks and overlays can be annotated in various ways to code the analysis. Such an approach has been attempted using commercially available video-editing software *Adobe Premiere* (O'Halloran, 2004b). Figure 2 is a screen shot of the film tracks and overlays used to analyse a short movie extract. However, it is apparent that the complexity of SF-MDA requires dedicated platforms to map the metafunctional flow of meaning in dynamic texts. Drawing manual overlays to map the patterns using commercially available software is simply too difficult and time-consuming; it's not an option. In addition, such an approach does not permit patterns across corpora to be discerned, which is a defining feature of Baldry and Beltrami's MCA system.

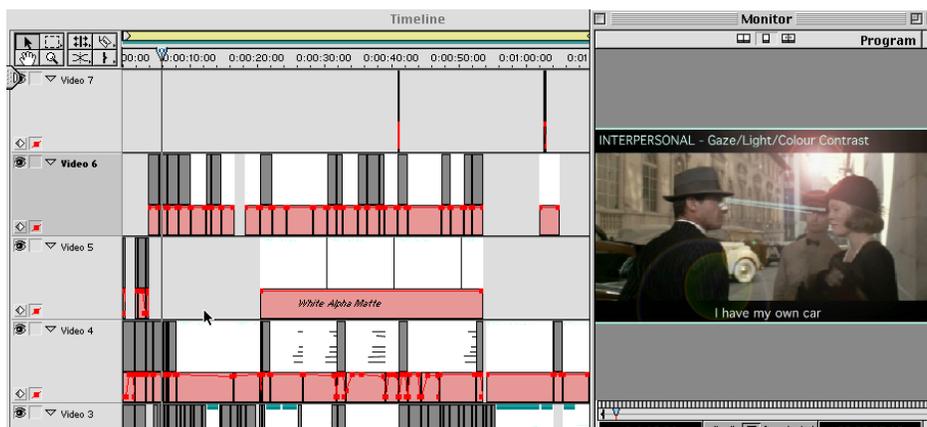


Figure 2 Visual Tracking using *Adobe Premiere* (from O'Halloran, 2004b)

Prototype software for SF-MDA which utilises computer-based multimedia techniques of analysis and statistical data may consist of multiple tracks and overlays which can be annotated and manipulated in various ways to analyse and compare intersemiotic patterns of meaning across different corpora. The integrative modelling will involve the visual display of patterns which track the functionality of semiotic choices over time. For example, changes in the facial expressions may be mapped and coded using 2D/3D graphics for real-time techniques for face detection, face recognition, face tracking and 3-D face modelling. Similarly, gesture may be mapped using 2D/3D human tracking, image based modelling, simulations of virtual human motion and 3-D reconstruction methods. Interactive dynamic modelling of complex relations has become a possibility, given the latest advances in computer-based techniques for multimedia analysis. The use of digital technology and visual forms of reasoning to advance SF-MDA are further explored below.

Digital Technology and Visual Reasoning

The use of digital technology for SF-MDA offers the opportunity to develop visual systems of reasoning, to function alongside linguistic and symbolic descriptions. Visual reasoning has the potential to extend beyond that which is currently found in print media (see Tufte, 1997, 2001). For example, the traditional analysis of non-verbal behaviour (e.g. Kendon, 2004; Pease, 1994) and facial expressions (e.g. Ekman and

Friesen, 1978) involves visual descriptions and vectors to indicate movement. However, the display and easy manipulation of data in digital media means that new systems for visual reasoning may be developed using a combination of multimedia techniques of analysis, visualisation processes and computer graphics software. Existing computer-based approaches may be incorporated and further developed in partnership with social semiotic theory for the purposes of SF-MDA.

Furthermore, visual reasoning for SF-MDA has the potential to extend beyond that found in mathematics and science, where data is first filtered through the lens of numerical quantification before being distilled into generalised mathematical models which are visualised. This practice decontextualises and reduces the complexity of the phenomenon under study. Funkenstein (1986: 75) explains “[t]he strength and novelty of seventeenth century science, both theoretical and experimental, was in its capacity to take things out of context and analyze their relations in ideal isolation” The complexity of contextualised data, however, may be maintained in digital media texts and events. The challenge is to model the metafunctional orchestration of system choices across semiotic resources and to contextualise the multimodal analysis at higher levels, i.e. the context of situation and culture.

Moreover, reasoning in mathematics and science is traditionally formalised using mathematical symbolism and language, rather than visual imagery. One reason is linked to technology; visual representations were static, time-consuming and expensive to produce in print media. However, visual representations which are dynamic, quick and inexpensive to produce have become a reality, thanks to the recursive semiotic transcoding between data and images which takes place in digital media. Despite the growing importance of visual representation, tension over the respective roles of visual imagery and analytical description in mathematics and science continues to exist. “After tracking the endless drive back and forth between images and data, it becomes clear that the powerful drive towards images and the equally forceful pressure towards analysis never completely stabilized scientific practice. Quite the contrary: neither the ‘pictorial-representative’ nor the ‘analytical-logical’ exist as fixed positions. Instead ... the image itself is constantly in the process of fragmenting and re-configuring ... image to non-image to image” (Galison, 2002: 322).

The continuous oscillation between visual and symbolic forms of display described by Galison (2002) may be the prelude to the development of new grammatical systems for visual reasoning. If so, the visual semiotic will possess an important advantage over the traditional semiotic resources of language and mathematical symbolism, given our ability to perceive patterns in visual data. “Just as the microscope showed us the ‘infinitely small’ and the telescope showed us the ‘infinitely large’, so the computer will enable us to regard our world in a new, richer way” (Colonna, 1994: 191). The ‘newer richer way’ suggested by Colonna (1994) arises through the capacity to model, experiment, manipulate, juxtapose, interconnect and synthesise the dimensions of experience. Furthermore, computer technology has the potential to incorporate other dimensions of experience, such as the touch, smell and taste of the ‘signifying body’ (Thibault, 2004).

The modelling process in digital environments permits noise, complexity and context to be incorporated to provide a more comprehensive picture of reality. Indeed ‘reality’ undergoes a transformation from *real*, that which can be perceived, to *abstract*, that which can be imagined. Software applications which recreate space and time are available (e.g. Danaher 2001). The spatial world and the multi-dimensional temporal world combine to create a dynamic rendition of real and virtual worlds, which can be modelled, experienced and analysed. The implications and applications of the semiotic re-ordering of reality made possible through digital technology are significant for developing new approaches to SF-MDA. Tufte (1997) provides a starting point for this venture.

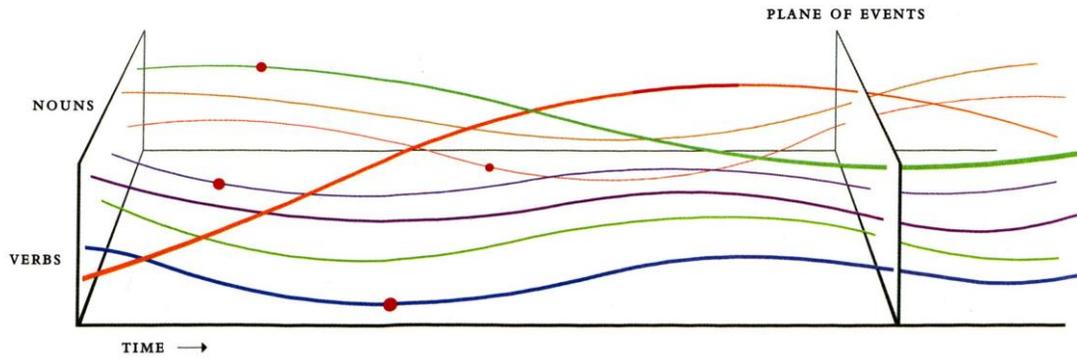


Figure 3(a) Visual Confection (Tufte, 1997: 121)

Tufte (1997: 121) visualises time flow from left to right in Figure 3(a), where the coloured strands represent stories in the ‘Ocean of the Streams of Story’. The red dots represent an instance in the story and sections of the coloured strands represent a short story or image sequence in Figure 3(a). The “assembly of many visual elements, selected from various Streams of Story, then brought together and juxtaposed on the still flatland of paper” results in a *visual confection* in the ‘Plane of Events’ in Figure 3(a). Tufte (1997: 121) explains that through the “multiplicity of image-events, confections illustrate an argument, present and enforce visual comparisons, combine the real and the imagined, and tell us yet another story”. Tufte (1997: 121-151) gives several examples of visual confections which are typically collections of images with some linguistic text.

The use of digital technology for SF-MDA has the potential to create a new form of visual confection, one in which strands come together from tracing the inter-weaving of meaning arising from semiotic choices over time. The two-dimensional plane of events captured on ‘the still flatland of paper’ may become a three-dimensional construction in the multidimensional world of digital technology captured in Figure 3(b). Multimodal social semiosis is reassembled in order to tell the story of how it was constructed in the first place, the nature of the story which results, and the other possibilities which may have been. Tufte’s (1997: 121) visual confection provides the basis for the new approach to SF-MDA, one in which the red dots and the plane of events are replaced with multi-functional streams of semiosis and multi-dimensional

renditions of experience. In other words, the paradigmatic shift for SF-MDA involves the visual representation of the social semiotic process-sequence for real and imagined realities.

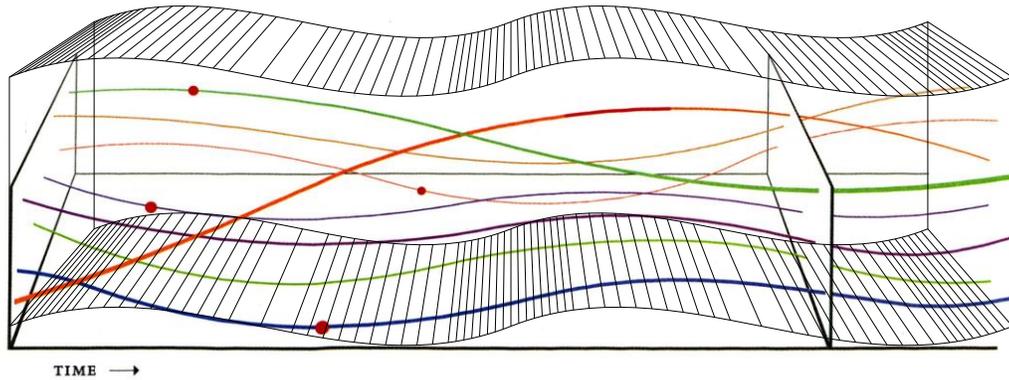


Figure 3(b) MDA Confection (adapted from Tufte, 1997: 121)

Multimodal Social Semiotics and Multimedia Analysis

Advancing SF-MDA through interactive digital technology is innovative because it combines social semiotic approaches to multimodal analysis with computer-based multimedia techniques of analysis, visualisation processes and computer graphics applications for the analysis of communicative artefacts and events. The approach involves the operationalisation and visualisation of Michael Halliday's metafunctional principle to track the flow of meaning as it unfolds in visual images, video texts and interactive digital sites. The meaning potential of the phenomenon under investigation (i.e. digital texts and events) is matched with the meaning potential of the multimodal tools of analysis.

The SF-MDA approach promises to advance existing computer-based approaches to multimedia analysis which depend on low-level feature recognition algorithms based on pixel values for image analysis (e.g. scenes, objects, structure, and spatial arrangements) and video understanding (e.g. motion-based recognition of actions and events, gestures and facial expressions). Many objects, actions and events can be

expressed as simple functions of low-level features (e.g. colour and texture measures, location of corresponding image regions in multiple frames, and spatio-temporal measures). Low level features which may be automatically detected include video cuts, shot changes, video captions, graphics, motion, subject positioning, angle shot, camera focus, lighting and mattes, grey scale, special sound effects and word recognition (Shah and Kumar, 2003; Smith and Kanade, 2005; Rosenfeld et. al., 2003). However, fully-automated image and video analysis has largely been successful only within domains where marked patterns of low-level features repeatedly occur (e.g. news, sports highlights, sports broadcasts, movie trailers and previews). It is proposed that SF-MDA provides a new theoretical platform for multimedia analysis, even if the analysis has to initially be manually and semi-automatically coded. The linguistic, visual and sound track analyses are integrated to provide a coherent and encompassing description of contextualised meaning. The operationalisation and visualisation of the SF metafunctional principle using a pre-programmed grammar (which can be modified) promises to advance multimedia event analysis.

The interdisciplinary approach promises to contribute to our understanding of the relationship between interactive digital technology and patterns of thinking. The relationship between technology and the creation of new knowledge has been documented (Galison, 2002). The significance of this relationship cannot be overstated. Digital technology impinges on our ability to use semiotic resources as tools for making meaning, and therefore digital technology is an agent for change. However, the potential of digital technology for creating new ways of thinking is yet to be harnessed. We remain in the pre-digital twilight zone where search engines are text-based and internet sites tend to rely on page-based conventions. The interdisciplinary approach advocated in this paper will explore the current limitations of creating meaning using digital technology with view to the future developments of semiotic forms with an expanded potential for knowledge creation.

The use of digital technology to model, describe, analyse and predict semiotic patterns involving the integrative dynamic codeployment of multiple semiotic resources will result in a new research paradigm which will extend linguistic and symbolic descriptions to dynamic visual representations. One precursor to this endeavour is to understand the paradigm shifts which have taken place in other

disciplines, including mathematics, science and computer science, and to develop, extend and modify those theories and techniques to fulfil the aims of SF-MDA. This requires a collaborative research effort between SF-MDA analysts, mathematicians, scientists and computer scientists to develop the interface (i.e. the interactive user platform), the system (i.e. the operational capabilities) and the program (i.e. the program algorithms). The respective roles of the interdisciplinary research team are displayed below.

LAYERS	RESEARCH TEAM
THE INTERFACE: User Driven (user interface design and usability for multimodal social semiotic analysis)	<ul style="list-style-type: none"> • Multimodal Discourse Analysts • Computer Scientists/Programmers • Researchers in other Disciplines
THE SYSTEM: Bridging technology and user considerations	<ul style="list-style-type: none"> • Multimodal Discourse Analysts • Computer Scientists/Programmers • Researchers in other Disciplines
THE PROGRAM: Technology Driven	<ul style="list-style-type: none"> • Computer Scientists/Programmers

Table 1 Interdisciplinary Collaboration: The SF-MDA Research Team

Conclusion

Multimodal analysis involves theorising the functionality and systems which constitute the grammar of semiotic resources. The modelling of semiotic resources other than language may involve a different paradigm from the one used for language where system networks are described using discrete categories. For example, grammars for gesture, facial expression, action and movement may be developed using computer-based multimedia techniques of analysis. Facial expressions may be dynamically modelled and interactively coded using visualisation techniques for real-time face detection and recognition, face tracking, 3-D face modelling and interactive shaping from shading. Gesture and human movement may be modelled using technologies for

simulation of human motion and tracking, 3-D reconstruction, image-based modelling and rendering, appearance modelling and image/video editing. In a similar manner, modelling and analysing speech, sound and music require the dynamic approaches made possible through digital technology.

Second, multimodal analysis requires analytical techniques which are capable of tracing and analysing the integrative patterns of inter-semiotic relations. This may involve visually modelling and annotating the metafunctional flow and orchestration of semiotic choices in dynamic texts. This requires interactive platforms for image and sound tracks which permit annotation of semiotic choices across layers which may be rendered to model the resultant meanings. This will require digital technology which integrates image, audio and motion analysis, and platforms to map phases and transitions which constitute the event sequence of dynamic texts.

Third, multimodal analysis involves the analysis, search and retrieval of semiotic patterns across multimedia data bases to investigate the relations between semiosis, text and context. Such patterns explain the nature of realities which are constructed, the interests served by such constructions, and the nature of alternative constructions. Such patterns reveal individual, community, cultural and global patterns of meaning making. The historical evolution of social semiosis may also be investigated with the aim of investigating the semiotic landscape and its relations with technology. Multimodal analysis can enhance data storage and retrieval techniques and software development for the analysis of visual images, video texts and interactive digital sites, a major problem facing industry and government.

Digital technology offers a platform for real time interactive and realistic visualisation. Scientists and engineers use supercomputing and scientific visualisation to presents solutions to better understand complex problems in the physical world such as motion, physical and chemical reactions, medical diagnoses and surgical operations, manufacturing processes and design optimisation. Researchers apply scientific visualisations to analyse, display and animate simulations. In a similar manner, digital technology may be used to help solve the multiple problems facing SF-MDA analysts. Significantly, these are problems which have major significance for industry, business and government with regards to the analysis, storage and retrieval

of information from the visual images and video texts proliferating through the rapid advance of technology in today's information age. The problems have significance for understanding the world in which we now live.

Acknowledgements

Figure 1 reprinted from 'Phase and Transition Type and Instance: Patterns in Media Texts as Seen Through a Multimodal Concordancer,' from *Multimodal Discourse Analysis* edited by Kay O'Halloran (2004) with kind permission of Continuum.

Figure 3(a) and 3(b) reprinted by permission, Tufte, Edward R. (1997: 121). *Visual Explanations: Images and Quantities, Evidence and Narrative*. Cheshire, Connecticut: Graphics Press. With special thanks to Professor Edward Tufte for waiving the copyright fee for this publication.

Reference List

(a) Primary sources for text analyses

(b) Referenced websites

<http://mca.unipv.it> (MCA multimodal concordancer: Software developed by Anthony Baldry and Michele Beltrami)

(c) Software and web-based tools

<http://mca.unipv.it> (MCA multimodal concordancer: Software developed by Anthony Baldry and Michele Beltrami)

(d) Scientific works

Baldry, Anthony (2004). "Phase and transition, type and instance: patterns in media texts as seen through a multimodal concordancer". In Kay O'Halloran (ed.),

- Multimodal Discourse Analysis: Systemic functional perspectives*. London and New York: Continuum, pp. 83-108.
- Baldry, Anthony, Beltrami, Michele, (2005). "The MCA Project: concepts and tools in multimodal corpus linguistics". In Maj Asplund Carlsson, Anne Løvland and Gun Malmgren (eds.), *Multimodality: Text, culture and use: Proceedings of the Second International Conference on Multimodality*. Kristiansand: Agder University College and Norwegian Academic Press, pp. 79-108.
- Baldry, Anthony, Thibault, Paul J. (2001). "Towards multimodal corpora". In Guy Aston, Lou Burnard (eds.), *Corpora in the description and teaching of English*. Bologna: CLUEB, pp. 87-102
- Baldry, Anthony, Thibault, Paul J. (2005). "Multimodal corpus linguistics". In Geoff Thompson and Susan Hunston (eds.), *System and Corpus: Exploring Connections*. London and New York: Equinox, pp. 164-183.
- Baldry, Anthony, Thibault, Paul J. (2006). *Multimodal Transcription and Text Analysis*. London: Equinox.
- Baldry, Anthony, Thibault, Paul J. (forthcoming). *Multimodal Corpus Linguistics*. London: Routledge.
- Barry, James (1996). *Measures of Science: Theological and Technical Impulses in Early Modern Thought*. Illinois: Northwestern University Press.
- Bateman, John, Delin, Judy, Henschel, Renate (2004). "Multimodality and Empiricism: Preparing for a corpus-based approach to the study of multimodal meaning-making". In Eija Ventola, Cassily Charles, Martin Kaltenbacher (eds.), *Perspectives on Multimodality*. Amsterdam and Philadelphia: John Benjamins.
- Colonna, Jean-Francois (1994). "Scientific Display: A Means of Reconciling Artists and Scientist", in Clifford A. Pickover, Stuart K. Tewksbury (eds.), *Frontiers of Scientific Visualization*. New York: John Wiley & Sons, pp. 181-212.
- Dalenoort, Gerhard J. (ed.) (1994). *The Paradigm of Self-Organization II*. New York: Gordon and Breach Science Publishers.
- Danaher, Simon (2001). *Digital 3d Design*. London: Cassell and Co.
- Davis, Philip J. (2006). "The Decline and Resurgence of the Visual in Mathematics". in Philip J. Davis, *Mathematics and Commonsense: A Case of Creative Tension* Wellesley, MA: A K Peters, Ltd, pp. 147-161.
- Davies, Paul C. W. (1990). "Why is the World Knowable", in Ronald E. Mickens (ed.), *Mathematics and Science*. Singapore: World Scientific, pp. 14-54.

- Djonov, Emilia (2006). *Analysing the Organisation of Information in Websites: From Hypermedia Design to Systemic Functional Hypermedia Discourse Analysis*. PhD Thesis, School of English and School of Modern Language Studies, University of New South Wales.
- Djonov, Emilia (2007). "Website hierarchy and the interaction between content organization, webpage and navigation design: A systemic functional hypermedia discourse analysis perspective", in print for *Information Design Journal* 15(2), Benjamins.
- Eisenstein, Elizabeth L. (1979). *The Printing Press as an Agent of Change*. Cambridge: Cambridge University Press.
- Ekman, Paul, Wallace V. Friesen (1978). *Facial Action Coding System*. Palo Alto, California: Consulting Psychologist Press Inc.
- Funkenstein, Amos (1986). *Theology and the Scientific Imagination from the Middle Ages to the Seventeenth Century*. New Jersey: Princeton University Press.
- Galison, Peter (1997). *Image and Logic: A Material Culture of Microphysics*. Chicago: University of Chicago Press.
- Galison, Peter (2002). "Images Scatter into Data, Data Gathers into Images". In Latour, B. and Weibel, P. (eds.), *Iconoclasm: Beyond the Image Wars in Science, Religion, and Art*. Cambridge, MA: MIT Press, pp. 300-323.
- Grave, Michel, Le Lous, Yvon (1994). "Visualization Services in Large Scientific Computing Centres", in Michel Grave, Yvon Le Lous and W. T. Hewitt (eds.), *Visualization in Scientific Computing*. Berlin: Springer-Verlag, pp. 10-28.
- Groß, M. (1994). *Visual Computing: The Integration of Computer Graphics, Visual Perception and Imaging*. Berlin: Springer-Verlag.
- Halliday, M. A. K. (1978). *Language as Social Semiotic: The Social Interpretation of Language and Meaning*. London: Edward Arnold.
- Halliday, M.A.K. (1994 [1985]). *An Introduction to Functional Grammar*. Second edition. London and Melbourne: Arnold.
- Halliday, M.A.K., Matthiessen, Christian (2004) *An Introduction to Functional Grammar*. 3rd Edition, London: Arnold
- Iedema, Rick (2001). "Analysing Film and Television: A Social Semiotic Account of Hospital: An Unhealthy Business". In Theo van Leeuwen and Carey Jewitt (eds.), *Handbook of Visual Analysis*. London: SAGE, pp. 183-204.

- Jewitt, Carey (forthcoming). *Routledge Handbook of Multimodal Analysis*. London and New York: Routledge.
- Jones, Janet (2007). *Multiliteracies for Academic Purposes: A Metafunction Exploration of Intersemiosis and Multimodality in University Textbook and Computer-Based Learning Resources in Science*. EdD Thesis, University of Sydney
- Kendon, Adam (2004). *Gesture: Visual Action as Utterance*. New York: Cambridge University Press.
- Kress, Gunther, Van Leeuwen, Theo (2006 [1996]). *Reading Images: The Grammar of Visual Design*. 2nd Edition. London and New York: Routledge.
- Kress, Gunther, Van Leeuwen, Theo (1996). *Reading Images: The Grammar of Visual Design*. 1st Edition. London and New York: Routledge.
- Lemke, Jay L. (1998). "Multiplying meaning: visual and verbal semiotics in scientific text." In James R. Martin and Robert Veel (eds.), *Reading Science: Critical and Functional Perspectives on Discourses of Science*. London and New York: Routledge, pp. 87-113.
- Lemke, Jay L. (2002). "Travels in hypermodality". *Visual Communication*. 1,3: 299-325.
- Levkowitz, Haim (1997). *Colour Theory and Modeling for Computer Graphics, Visualization, and Multimedia Applications*. Dodrecht: Kluwer Academic Publishers.
- Manovich, Lev. (2001). *The Language of New Media*. Cambridge, MA: The MIT Press.
- Martin, James R. (1992). *English Text: System and structure*. Amsterdam and Philadelphia: John Benjamins.
- Martin, James R. (2002). "Fair Trade: Negotiating Meaning in Multimodal Texts". In Patrick Coppock (ed.), *The Semiotics of Writing: Transdisciplinary Perspectives on the Technology of Writing*. Turnhout, Belgium: Brepols, pp. 311-338.
- Martin, James R. and Rose, David (2003). *Working with Discourse: Meaning Beyond the Clause*. London: Continuum.
- Martinec, Radan (2000). "Types of Processes in Action". *Semiotica*, 130-3/4, 243-268.
- Martinec, Radan (2001). "Interpersonal Resources in Action". *Semiotica*, 135-1/4, 117-145.
- Mikhailov, Alexander S., Calenbuhr, Vera (2002). *From Cells to Societies: Models of Complex Coherent Action*. Berlin: Springer.

- Moorhead, Robert, Gross, Markus, Joy, Kenneth I. (eds.) (2002). *IEEE Visualization 2002: Proceedings*. Boston: Institute of Electrical and Electronics Engineers, Inc.
- Muntigl, Peter (2004). "Modelling Multiple Semiotic Systems: The Case of Gesture and Speech". In Eija Ventola, Cassily Charles and Martin Kaltenbacher (eds) (2004), *Perspectives on Multimodality*. Amsterdam: John Benjamins, pp. 31-49.
- O'Halloran, Kay L. (ed.) (2004a). *Multimodal Discourse Analysis: Systemic functional perspectives*. London and New York: Continuum.
- O'Halloran, Kay L. (2004b). "Visual semiosis in film". In Kay L. O'Halloran (ed.), *Multimodal Discourse Analysis: Systemic functional perspectives*. London and New York: Continuum, pp. 109-13.
- O'Halloran, Kay L. (2005). *Mathematical Discourse: Language, Symbolism and Visual Images*. London and New York: Continuum
- O'Halloran, Kay L. (2007). "Systemic Functional Multimodal Discourse Analysis (SF-MDA) Approach to Mathematics, Grammar and Literacy", in Anne McCabe, Michael O'Donnell, and Rachel Whittaker (eds.), *Advances in Language and Education*. London and New York: Continuum.
- O'Halloran, Kay L., Judd, Kevin (2002). *Systemics 1.0. CD-ROM* Singapore: Singapore University Press.
- O'Toole, Michael (1994). *The Language of Displayed Art*. London: Leicester University Press.
- O'Toole, M. (1999), *Engaging with Art (CD ROM)*. Perth: Murdoch University
- Pang, Kah Meng Alfred (2004). "Making History in From Colony to Nation: A Multimodal Analysis of a Museum Exhibition in Singapore". In Kay L. O'Halloran (ed.), *Multimodal Discourse Analysis*. London and New York: Continuum, pp. 28-54.
- Paul, Christiane (2003). *Digital Art*. London: Thames and Hudson.
- Peak, David, Frame, Michael (1994). *Chaos Under Control: The Art and Science of Complexity*. New York: W. H. Freeman and Company.
- Pease, Allan (1994). *Body Language: How to Read Others' Thoughts by Gestures*. Singapore: Heineman Asia.
- Rabinovich, Mikhail I., Ezersky, A. B. and Weidman, P. D. (2000). *The Dynamics of Patterns*. Singapore: World Scientific.
- Ragsdell, Gillian, Wilby, Jennifer (eds.). (2001). *Understanding Complexity*. New York: Kluwer Academic.

- Ravelli, Louise J. (2000). "Beyond Shopping: Constructing the Sydney Olympics in Three Dimensional Text". *Text*, 20(4), 489-515.
- Rosenfeld, Azriel, Doermann, David, DeMenthon, Daniel (eds.) (2003). *Video Mining*. Boston, Dordrecht, London: Kluwer Academic Publishers.
- Royce, Terry, Bowcher, Wendy (eds.) (2006). *New Directions in the Analysis of Multimodal Discourse*. New Jersey: Lawrence Erlbaum Associates.
- Shah, Mubarak, Kumar, Rakesh (eds) (2003). *Video Registration*. Boston, Dordrecht, London: Kluwer Academic Publishers.
- Smith, Michael A., Kanade, Takeo (2005). *Multimodal Video Characterization and Summarization*. Boston, Dordrecht, London: Kluwer Academic Publishers.
- Smith, Peter (1998). *Explaining Chaos*. Cambridge, UK: Cambridge University Press.
- Stenglin, Maree (2004). *Packing Curiosities: Towards a Grammar of Three Dimensional Space*. PhD Thesis. Linguistics Department, University of Sydney.
- Stenglin, Maree (2007). "Space Odyssey: A Guided Tour Through the Semiosis of Three-Dimensional Space". *Visual Communication*. In press.
- Stewart, Ian (1995). *Nature's Numbers: Discovering Order and Pattern in the Universe*. London: Weidenfeld and Nicolson.
- Swetz, Frank J. (1987). *Capitalism and Arithmetic*. La Salle, Illinois: Open Court.
- Tan, Sabine (2005). *A Systemic Functional Approach to the Analysis of Corporate Television Advertisements*. MA Thesis, Department of English Language and Literature, National University of Singapore.
- Thibault, Paul J. (2000). "The Multimodal Transcription of a Television Advertisement: Theory and Practice", in Anthony Baldry (ed.) *Multimodality and Multimediality in the Distance Learning Age*. Campobasso, Italy: Palladino Editore, pp. 311-385.
- Thibault, Paul J. (2004). *Brain, Mind, and the Signifying Body: An Ecosocial Semiotic Theory*. London and New York: Continuum.
- Tufte, Edward R. (1997). *Visual Explanations: Images and Quantities, Evidence and Narrative*. Cheshire, Connecticut: Graphics Press.
- Tufte, Edward R. (2001). *The Visual Display of Quantitative Information*. Cheshire, Connecticut: Graphics Press.
- van Leeuwen, Theo (1999). *Speech, Music, Sound*. London: Macmillan.
- van Leeuwen, Theo (2005). *Introducing Social Semiotics*. London: Routledge.
- Ventola, Eija, Charles, Cassily, Kaltenbacher, Martin (eds.), (2004). *Perspectives on Multimodality*. Amsterdam and Philadelphia: John Benjamins.

- Walgraef, Daniel (1997). *Spatio-Temporal Pattern Formation: With Examples from Physics, Chemistry, and Materials Science*. New York: Springer.
- Watson, Velvin, Walatka, Pamela P. (1994). "Visual Analysis of Fluid Dynamics", in David F. Rogers, Rae A. Earnshaw, R. A. (eds), *State of the Art in Computer Graphics: Aspects of Visualization*. New York: Springer-Verlag, pp. 7-18.