Towards a metalanguage for multiliteracies education: Describing the meaning-making resources of language-image interaction

LEN UNSWORTH
University of New England, Australia.

ABSTRACT: The increasingly integrative use of images with language in many different types of texts in electronic and paper media has created an urgent need to go beyond logocentric accounts of literacy and literacy pedagogy. Correspondingly there is a need to augment the genre, grammar and discourse descriptions of verbal text as resources for literacy pedagogy to include descriptions of the meaning-making resources of images. Some augmentation along these lines has involved the articulation of Hallidayan systemic functional descriptions of language, mainly focussed on verbal grammar, with the social semiotic descriptions of the meaning-making resources of images described in a grammar of visual design proposed by Kress and van Leeuwen. However, current research indicates that articulating discrete visual and verbal grammars is not sufficient to account for meanings made at the intersection of language and image. This paper adopts a systemic functional semiotic perspective in outlining a range of different types of such meanings in different kinds of texts, suggesting the significance of such meanings in comprehending and composing contemporary multimodal texts, and the importance of developing an appropriate metalanguage to enable explicit discussion of these meaning-making resources by teachers and students.

KEYWORDS: Multiliteracies, new literacies, metalanguage, visual literacy, visual grammar, image-text relations, multimodality, systemic functional linguistics, social semiotics.

INTRODUCTION

It is now widely accepted that literacy and literacy pedagogy can no longer be confined to the realm of language alone, and that reconceptualizing literacy and literacy education needs to account for the role of images (as well as other modes of meaning-making) in paper (hard copy) and electronic media texts. In Australia, State English syllabi generally require students to learn about the role of images in their comprehension, and to a lesser extent, their composition of various kinds of texts. This appears to be largely uncontentious in contemporary English teaching. What is, and has long been contentious in dealing with language in English teaching in Australia, the United Kingdom and North America is the role of metalanguage – the type of grammar, its purpose in the curriculum and approaches to its teaching. Today, in the national curriculum for England and in English syllabi in Australian States, grammar is required to be taught. For the most part, traditional grammar terminology has been retained, although some Australian States also incorporate functional grammatical concepts from systemic functional linguistics (SFL), sometimes known as Hallidayan linguistics (Halliday & Matthiessen, 2004; Martin, 1992). Substantial curriculum support documents and appendices to syllabi routinely include quite

technical grammatical concepts to facilitate teachers' and students' explicit use of this metalanguage (Education, 1995; New South Wales Board of Studies, 1998).

No such comparable accounts of a metalanguage describing the meaning-making resources of images and image/text interaction accompany these government curriculum documents and syllabi. Faced with the requirement to address the multimodality of texts, the prescription of verbal grammar, and the absence in syllabi of comparably theorized resources for describing the meaning-making resources of images, some teacher educators and teachers have made use of the "grammar of visual design" developed by Kress and van Leeuwen (1996), extrapolating from SFL accounts of language. The commonality of the systemic functional theoretical approach to language and image as social semiotic systems facilitates an articulation of visual and verbal grammar as descriptive and analytical resources in developing students' comprehension and composition of multimodal texts. However, beyond accounting for the independent, albeit sometimes strategically aligned, contributions of language and image to the meaning of composite texts, is the challenge of systematically theorising and describing resources for the construction of meaning at the intersection of language and image.

The purpose of this paper is to outline recent work addressing this challenge, and in so doing to indicate the pedagogic utility of formulating such a metalanguage of multimodality for the development of the multiliteracies education needed by students to engage with contemporary multimodal texts and texts of electronic multimedia. Firstly, I shall invite readers to experience an introductory example of one type of meaning made at the intersection of language and image in Anthony Browne's (1994) picture book *Zoo*. In the next section of the paper I will outline the key tenets of systemic functional semiotic theory that facilitate its use in describing meaning-making resources within and across a variety of modes of meaning including language, images, music and gesture. The subsequent section, and main body of the paper, will outline recent research dealing with the development of descriptions of meaning-making resources of image-language interaction. Finally, I will suggest – on the basis of research reporting the pedagogic efficacy of the metalanguage of SFL, some work on the pedagogic use of the grammar of visual design, and the discussion in previous sections of the emerging research on descriptions of image-language interaction – that teachers, teacher-educators and researchers consider further the pedagogic potential of existing and emerging metalanguage drawing on systemic functional semiotic approaches to multimodal texts.

EXPERIENCING MEANING-MAKING AT THE INTERSECTION OF LANGUAGE AND IMAGE

A very clear example of meaning constructed at the intersection of image and language is provided in Anthony Browne's (1994) picturebook, *Zoo*. For readers who are not familiar with this story and are not able to readily locate a copy, it is possible to read the relevant excerpt via the story sample provided on Amazon.com or simply go to Amazon.com and search for *Zoo*. In this story, Mum and Dad and their pre-adolescent sons, the narrator and his brother Harry, go to the zoo. In the book, images of the family and other visitors to the zoo are on the left-hand side of the double page spreads and images of the zoo animals are on the right hand side. In the story

segment discussed here, the image on the left hand side shows a very low angle, medium close view of Dad from the waist up with two white clouds in the sky positioned to suggest they are horns protruding from each side of Dad's head. In the text below the image, Harry, the narrator, asks if they can eat the chocolate that Mum packed. Dad refuses, and when asked why, simply says, "Because I say so." The image on the right hand side shows the giraffes with no text. On the subsequent left hand page we see a rear distance view of Dad and the boys leaning over a fence. The text below concerns the tiger they are looking at and makes no mention of the chocolate or eating. However, in the image on the ground at Dad's feet it is possible to discern what looks like a discarded chocolate wrapper. After reading these pages the reader is in a position to suggest why Dad did not allow the boys to have the chocolate, but to do so s/he must make the inference on the basis of converging information from the image (showing the discarded chocolate wrapper) and the text that occurred two pages earlier.

KEY TENETS OF SYSTEMIC FUNCTIONAL SOCIAL SEMIOTIC THEORY

According to SFL, the structures of language have evolved (and continue to evolve) as a result of the meaning-making functions they serve within the social systems or cultures in which they are used. Language is considered as a meaning-making system where the options available to individuals to achieve their communicative goals are influenced by the nature of the social context and how individuals are positioned in relation to it. However, although Halliday focused on language, he was very clear that this was only one semiotic system among many other modes of meaning in any culture, which might include

> … both art forms such as painting, sculpture, music, the dance, and so forth, and other modes of cultural behaviour that are not classified under the heading of forms of art, such as modes of exchange, modes of dress, structures of the family, and so forth. These are all bearers of meaning in the culture. Indeed we can define a culture as a set of semiotic systems, as a set of systems of meaning, all of which interrelate (Halliday & Hasan, 1985, p. 4).

The strength of SFL in contributing to frameworks for the development of intersemiotic theory emanates from its conceptualization of language as one of many different interrelated semiotic systems, and hence the assumption that the forms of all semiotic systems are related to the meaning-making functions they serve within social contexts. SFL proposes that these meaning-making functions can be grouped into three main categories, or metafunctions. These are the three types of meaning-making that are inherent in all instances of communication, regardless of whether the communication is via language, image, music, sculpture or some other semiotic mode. The three kinds of meaning-making or metafunctions are related to three corresponding situational variables that operate in all communicative contexts.

Any communicative context can be described in terms of these three main variables that are important in influencing the semiotic choices that are made. The first of these, FIELD, is concerned with the social activity, its content or topic; the second, TENOR, is the nature of the relationships among the people involved in the communication; and the third, MODE, is the medium and channel of communication. In relation to language, MODE is concerned with the role of language in the situation – whether

spoken or written, accompanying or constitutive of the activity, and the ways in which relative information value is conveyed. These situational variables are related to three overarching areas of meaning, or metafunctions: "ideational", "interpersonal" and "textual". For example, if I say, "My daughter is coming home this weekend", ideationally this involves an event, a participant and the circumstances of time and place associated with it. Interpersonally it constructs me as a giver of information and the reader/listener as a receiver (as well as perhaps suggesting I have at least some acquaintance with the listener). Textually, it locates "my daughter" as the "Theme" or orientation or point of departure for the interaction, simultaneously suggesting that "my daughter" is given information that we both know about ("Given") and the new information is that she is coming home "this weekend" ("New"). If I say, "Is my daughter coming home this weekend?" the ideational meanings remain the same – the event, the participant, the circumstances have not changed. But the interpersonal meanings have certainly changed. Now I am demanding information, not giving it (and there may be some suggestion of estrangement between the listener and me). Similarly, if I say, "This weekend my daughter is coming home", the ideational meanings are still the same, but this time the textual meanings have changed. Now the orientation (Theme) is the weekend and this is the given or shared information. What is new or unknown concerns what my daughter is doing. So the different structures reflect different kinds of meaning, which in turn reflect different aspects of the context. The metalanguage of systemic functional grammar derives from this linking of language structure, meaning and context.

It is this metafunctional aspect of SFL and its link to the situational variables of social contexts that has provided a common theoretical basis for the development of similar "grammatical" descriptions of the meaning-making resources of other semiotic modes. For example, Kress and van Leeuwen (1996) proposed that images, like language, also always simultaneously realize three different kinds of meanings. Images construct not only representations of *material reality* but also the interpersonal interaction of *social reality* (such as relations between viewers and what is viewed). In addition images cohere into textual compositions in different ways and so realize *semiotic reality*. More technically, the "grammar of visual design" formulated by Kress and van Leeuwen (1996) adopted from SFL the metafunctional organization of meaning-making resources:

- *representational/ideational* structures verbally and visually construct the nature of events, the objects and participants involved, and the circumstances in which they occur.
- *interactive/interpersonal* verbal and visual resources construct the nature of relationships among speakers/listeners, writers/readers, and viewers and what is viewed.
- *Compositional/textual* meanings are concerned with the distribution of the information value or relative emphasis among elements of the text and image.

Many researchers exploring image/text relations explicitly acknowledge the grounding of their work in the SFL metafunctional hypothesis (Baldry, 2000; Lemke, 1998a, 1998b, 2002; Macken-Horarik, 2003a, 2004; Martin, 2002; O'Halloran, 2004; Royce, 1998, 2002). Similar extrapolations from the metafunctional basis of SFL have provided social semiotic descriptions of "displayed art" (O'Toole, 1994), music and sound (van Leeuwen, 1999) and action (Martinec, 1999, 2000a, 2000b).

TOWARDS A METALANGUAGE OF IMAGE-LANGUAGE INTERACTION

A metafunctional orientation to describing inter-modal meaning

The formulation of a metalanguage for multiliteracies needs to entail both the description of the specific characteristics of each participatory semiotic mode and also the more broadly encompassing semiotic characteristics that enable it to be related to the meaning-making contributions of other modes in multimodal texts (Kress, 2000a, 2003a; Macken-Horarik, 2003a; Martin, 2003). In working towards this formulation Kress (2000a; 2003a) cautions against too much reliance on descriptions deriving from language-based theories of communication and meaning.

Acknowledging the concern by Kress that over-reliance on language-based theories of meaning and communication would obviate an adequate and integrated description of multimodal textual objects, it would nevertheless seem that he and other SFL-influenced social semioticians have effectively established a mapping of the SFL metafunctions across modalities. With slight differences in nomenclature, the equivalent of Halliday's metafunctions have been readily applied to social semiotic accounts of images as summarized in Table 1 adapted from Martin (2002).

metafunction: modalities:	naturalizing reality	enacting social relations	organizing text
verbiage			
Halliday (1994)	ideational	interpersonal	textual
image			
Kress & van Leeuwen (1996)	representation	interaction/modality	composition
O'Toole (1994)	representational	modal	compositional
Lemke (1998b)	presentational	orientational	organizational

Table 1. Metafunctions in verbiage and image (after Martin, 2002, p. 1)

Martin (2002) further pointed out that these same modes of meaning had, to some extent, been deployed for analyzing relations across modalities in multimodal texts. For example, he cited the work of Kress and van Leeuwen (1996) on textual/compositional meaning, which both adopted and adapted the SFL notion of "information focus". One aspect of this is the distinction between "Given" and "New" information. In language, typically information that is already known or familiar to the reader, or "Given", is located at the beginning of the clause (mapped onto the Theme), while information that is "New" is located at the end of the clause (mapped onto the Rheme) (Halliday, 1994).

The visual analogue of this proposed by Kress and van Leeuwen (1996) is that typically in images and image/text compositions, for those in Western cultures, the Given information is located on the left and New information is located on the right. But Kress and van Leewen (1996) provided further descriptions of additional parameters of information focus distinctive to images and image/text compositions. For example, they distinguished the top half of such compositions as typically the location of the "Ideal" while the lower half was the location of the "Real". In

advertisements, the top part typically indicates the promise of the product – its imagined or ideal effects, while the bottom part of the layout indicates more concrete information about the product itself. In textbooks, the top part deals with the more generalized, abstract, conceptual information, while the bottom part deals with the specific, concrete, observable information. It may be, then, that Given or New are realized by an image in some instances and language in others. Similarly, in some texts the Ideal is realized by an image and the Real by language or vice versa.

The challenging task of formulating a metalanguage of multiliteracies can be very usefully informed by the initiatives of social semiotic researchers in this direction involving adoption and adaptation of, and innovation on, generative semiotic theory with an SFL lineage. The subsequent sub-sections review recent examples of such work dealing in turn with research focusing on each of the three metafunctions.

Describing resources for the inter-modal construction of ideational meaning

What is being investigated here is the space of integration between language and image as social semiotic systems in order to provide a theoretical description of the dynamics of interaction between language and image in meaning-making (Lim, 2004). In terms of ideational meaning, this interaction may be characterized as ideational *concurrence* (Gill, 2002), *complementarity* or *connection.*

Ideational Concurrence
Ideational concurrence was described by Gill (2002) in a study of image/text relations in picture storybooks for young children. Concurrence referred to ideational equivalence between image and text. This was operationalized as the image and text having an equivalent participant-process-phenomenon configuration. For example, the first image in Anthony Browne's well-known picturebook, *Gorilla* (Browne, 1983), can be transcoded as. "Hannah is reading a book about gorillas while sitting on the floor." This concurs with the verbal text: "She read books about gorillas.". Concurrence may entail some form of redundancy across modes, but this is not a simple inter-modal duplication of meaning. Although Martinec and Salway (2005) do not use the category *concurrence*, they describe one such type of image-text relations as "exposition" – "where the image and the text are of the same level of generality"(Martinec & Salway, 2005, p. 350). Their example is the relation between an image and its caption "Light micrograph of a bone" (Martinec & Salway, 2005, p. 362).

In examples from children's literature like the one from *Gorilla* (Browne, 1983), the image-text relation is one of instantiation. The language conveys the habitual nature of the activity while the image indicates one instance, adding to the meaning of the language version that, at least on some occasions, this reading was done while sitting on the floor in the house. The degree of redundancy is variable depending on the context of the process or activity common to the language and image. For example, image three in *Gorilla* depicts the father walking along the street with a briefcase. This concurs with, and provides an instantiation of the text: "He went to work every day." The image clearly suggests additional meanings such as what kind of work he did and to some extent how he got to work. A similar category from the Martinec and Salway (2005) work on news websites, textbooks and advertisements is that of "exemplification". This relation obtains when either the image or the text is more

general. The former is illustrated by a skull and crossbones image with the caption "Kills by biting prey with jagged teeth". The latter is illustrated by an image of children playing and rolling down a hill in a cardboard box accompanied by the caption "Remember when total freedom came in a box". These could also be described in terms of instantiation. The children playing in the box is an instance of the generality of the caption and the death symbol of the skull and crossbones is instantiated by the accompanying text. In both cases, like the children's literature examples, significant additional meanings are added by the language or the image.

A further means by which ideational concurrence is achieved inter-modally is perhaps the most immediately arresting to the reader/viewer. This is the phenomenon of "homospatiality", discussed by Lim (2004), and refers to texts where two different semiotic modes co-occur in one spatially bonded homogenous entity. One example shows the linguistic representation, "snaaap", which visually appears with the "sna" segment forming one arm of an inverted "v"shape and the "aap" segment forming the other arm, so that it appears that the word itself has "snapped", as indicated in Figure 1.
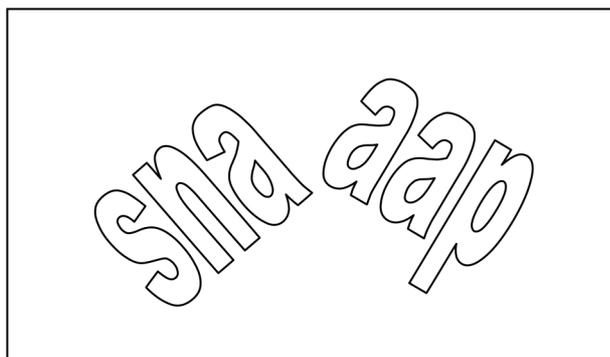

Figure 1. "homospatiality"

Another example shows an image of a campfire with the heat arising from the fire represented by curved lines, which can be read to spell the word "hot".

Ideational concurrence then, is consistent with Lemke's (2002) notion of the multiplicative nature of the meaning-making capacity of multimodal texts being the logical product of the capacities of the constituent semiotic systems. In other words the visual-verbal interface is synergistic, producing a total effect that is greater than the sum of the contributions of each modality (Royce, 1998). At this point we could summarize our partial framework for understanding the construction of ideational meaning at the intersection of language and image as a set of semantic options for intermodal relations as indicated in Figure 2.
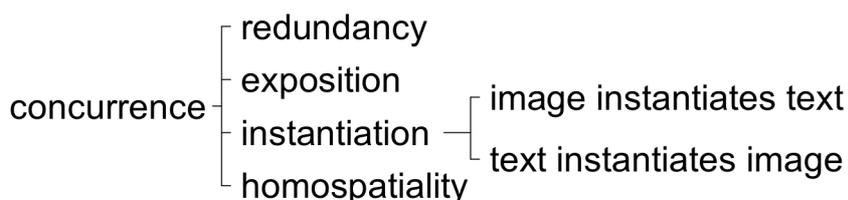
```
                     ┌ redundancy
                     ├ exposition              ┌ image instantiates text
concurrence ─┤ instantiation ─┤
                     └ homospatiality          └ text instantiates image
```

Figure 2. Ideational Concurrence

Ideational Complementarity

Ideational complementarity refers to the situation in multimodal texts where what is represented in images and what is represented in language may be different but complementary and joint contributors to an overall meaning that is more than the meanings conveyed by the separate modes. Quite independently, Kress in the UK (Kress, 1997, 2000b, 2003a, 2003b) and Lemke in the US (Lemke, 1998b) have explicated what is referred to as the "functional specialization" of language and image. According to this specialization principle, the resources of language are most apposite to the representation of sequential relations and the making of categorical distinctions, while the resources of images are most apposite to the representation of spatial relations and for formulating relationships such as those of degree, gradation, continuous co-variation and dynamic emergence (Lemke, 1998b). Language and images are not restricted to the areas of representation indicated by the functional specialization principle, but as images are becoming more frequent in a wide range of texts, functional specialization is likely to characterize the ideational complementarity of these two modes.

One type of ideational complementarity is *augmentation* – where each of the modes provides meanings additional to and consistent with those provided in the other mode. Martinec and Salway (2005) refer to this as "extension", but provide only examples that indicate the text adding to the meaning of the images. In a study comparing school science explanations in books, on CD ROMs and on the World Wide Web (Unsworth, 2004) in terms of the relationships between illustrations and the main text, data for the image-text relation of extension included instances where the image extended the meanings of the text. For example, the explanation of the water cycle on the Classroom of the Future website included evaporation from the soil and the movement of clouds in its diagram but did not mention these in the main text. See also Unsworth (in press)for instance of images in advertisements extending the meanings of the text, such as the advertisement for Mercedes-Benz E-class sports pack cars where it is only in the image that we are informed that this vehicle is available in sedan and wagon models.

The augmentation of the text by images is fundamental to the construction of interpretive possibilities in literary picture-books for children. This can be seen in examples of such picture-books where significant segments of the narrative are conveyed by several pages that consist of images alone. In *Where the Wild Things Are* (Sendak, 1962), for example, the conduct of the "wild rumpus" is conveyed by images alone in three consecutive double page spreads. But also, where images and text are co-present, significant elements of the action of the story frequently occur within the images only. For example, on page nine of Anthony Browne's *Gorilla* (Browne, 1983), the text foreshadows subsequent events: "In the night something amazing happened." Then the images on this page are exclusively responsible for

conveying just what the amazing event was. It is these images only that depict Hannah's toy gorilla growing into a real gorilla (Gill, 2002). Juxtaposed images and text in picture-books have also been shown to jointly construct activity sequences. Gill, extending her work on ideational concurrence in picture-books, described the nature of this joint image/text construction of meaning as "distribution". Distribution, however, seems to be appropriately discussed within ideational complementarity. According to Gill (2002), there are two types of distribution. Intra-process distribution refers to the portrayal by images and text of different aspects of a shared process. For example, the image(s) might depict the end result of a process described in the verbal text. This occurs in *Gorilla* when the text indicates that Hannah and the gorilla crept downstairs and Hannah put on her coat and the gorilla put on her father's hat and coat. The image shows them standing in the doorway so dressed. Inter-process distribution occurs when images fill a gap in the ideational flow of meaning in the verbal text. For example, later in the story of *Gorilla*, the text indicates that it is time to go home and then indicates that they danced on the lawn, which is clearly in front of Hannah's home. But the text makes no reference to their actually going home. This is conveyed by the image of the gorilla walking along the street with Hannah on his shoulders.

Another form of *Ideational Complementarity* is *Ideational Divergence*, where the ideational content of text and image are opposed. Ideational divergence does not seem to have figured in the research dealing with inter-semiotic concurrence and complementarity, and it is not mentioned in the system for image-text relations proposed by Martenic and Salway (2005). Nevertheless, it is clearly important in children's literary picture books. For example, in the "Shirley" books by John Burningham (1977; Burningham, 1978), the text and images of Shirley's parents convey a narrative of a typical beach visit or of a child taking a bath, while the images of Shirley depict her as participating in exciting adventures such as her encounter with pirates. Similarly, McCloud (1994) has drawn attention to the role of ideational divergence in the narrative art of comic books. In his category of image/text relations, he uses the term "parallel combinations" to denote instances where "words and pictures seem to follow very different courses – without intersecting"(McCloud, 1994, p. 154).

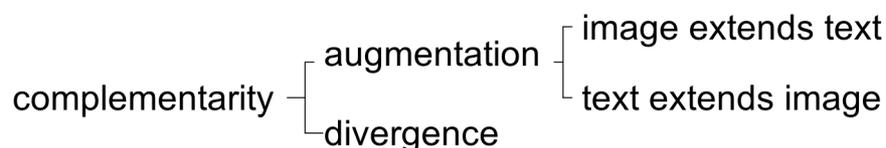A simple framework summarizing these types of image-text ideational complementarity is shown in Figure 3.

Figure 3. Ideational Complementarity

Connection

There are two types of *connection* between images and text. The first of these is known as *projection* and most commonly involves the quoting or reporting of speech or thoughts. The second type of connection involves the *conjunctive relations* of time,

place and cause.   Martinec and Salway (2005) treat this second type as a further category (enhancement) in their description (using different terminology) of ideational concurrence and complementarity, and distinguish this whole group of three categories (which they call *expansion*) from projection.   This is consistent with SFL descriptions of logical relations.   My departure from SFL here is exploratory only, in the light of the advice from Kress noted earlier and in seeking a felicitous account of image-text inter-modal meaning-making.

Projection in the Martinec and Salway (2005) system refers to either a "locution", which is the quoting or reporting of wording, or an idea, which is the quoting or reporting of thought.   They cite the speech or thought bubbles in cartoons as the typical realizations.   But a further realization of projection in language/image interaction occurs where a verb in the text "projects" or quotes what a character says or is thinking and the verbal or mental quotation is realised by images rather than language.   It also refers to the juxtaposition of quoted speech and a participant in the image represented as the obvious source of the quote.   The latter form of projection commonly occurs in magazine advertisements where the participant looks directly at the viewer from a social or close-up position, thereby making contact that "demands" a pseudo interpersonal interaction.   The juxtaposed quote then is very strongly assumed to be attributed to this represented participant.   One example of such an advertisement is provided by Cheong (2004).   It shows a "demand" image of a smiling young woman at a medium close-up position holding a poster with the logo of the "M1" company that offers an attractive, weekly, "off peak" discount for energy consumption.   The quote spans the width of the advertisement and is located just above the head of the woman: "I get the feeling that M1 wants me to enjoy value – and enjoy life.   Everything they offer is brighter, nicer and more fun!"

The *Economist* magazine advertisement analysed by Royce (1998) shows a monochrome photograph with a medium to close-up, eye-level view of a young woman whose gaze is directed at the viewer, and whose frontal plane is parallel with that of the viewer.   These visual features realize a pseudo interpersonal relation of direct involvement at a personal level with a demand for a response.   Positioned immediately above this image is the following question in the largest font on the page: "Does your environmental policy meet your granddaughter's expectations?"   The implicitness of the attribution of the quote combined with the interactive role of the choice of image demonstrates the powerful engagement of projection achieved through the intersection of language and image.

Image projection from a verb in the text occurs in the picture book *Hyram and B* (Caswell & Ottley, 2003).   In this story Hyram and B are two bears who have lived on the shelf in a second-hand shop longer than any other toys.   They have shared memories of their traumatic days of being discarded and their understandings of loneliness.   Eventually a young war orphan named Catherine takes B home, turning the world of the two bears upside down. But later, Catherine returns to the shop and collects Hyram.   Two, consecutive, double-page spreads deal with Hyram's recounting his earlier life to B.   In the first of the double page spreads, B says: "Hyram sleeps a lot.   He told me about it once."   The verb "told" projects what is realized both verbally and visually in the next double page spread as Hyram recounts his experience; and this past experience is also recalled visually in the illustrations.   On a later, double-page spread the verb "remember" on the right hand page projects

B's memories represented only by the images on the left hand side of that double page and also on the subsequent double page spread.

A further realization of projection is proposed by Martinec and Salway (2005) to account for occurrences of a diagram which recapitulates the ideational content of a juxtaposed segment of main text. However, they do not make it clear whether they regard this as a locution or an idea. Here, *projection* will be described as *verbal* when the quoting or reporting is concerned with wording, and as *mental* when the concern is with thought.

Conjunction refers to the *connection* of images and text in terms of causal, temporal or spatial relations. The third-last, double-page spread in *Hyram and B* (Caswell & Ottley, 2003) constructs causality at the intersection of image and text. The left-hand third of this double spread shows the following text (in column format) with a rear view image of Catherine holding B in her arms with "action" lines suggesting that she is trembling or that she is rocking B.

> Catherine loves me.
> Catherine understands the secret language of bears.
> She understands what it means to be lonely
> (Caswell & Ottley, 2003, no page numbers).

The remaining two thirds of the left page and the entire right hand page show an explosive warfare scene with fire and a helicopter in the background, artillery and a damaged tank in the foreground as well as a the rear-view image of a red-headed girl in an almost parallel pose to the separate image of Catherine on the far left page onto which this warfare image is partially superimposed. This parallelism suggests why "Catherine understands the secret language of bears" and why she "understands what it means to be lonely".

Causal conjunction is illustrated by Martinec and Salway (2005) by means of an image showing what appears to be people walking around a line up of "body bags" or "coffins accompanied by the caption: "Police believe a short circuit set fire to the hall's thatch roof." The authors claim here that "the image enhances the text. The dead bodies lying on the floor are the result *of a short circuit set fire to the hall's thatch roof*" (Martinec & Salway, 2005, p. 351). However, one might also reason that the text enhances the image since the short circuit… was the cause of the line-up of dead bodies.

Temporal relations between images and text can also be seen in picture-books where juxtaposed images and text jointly construct activity sequences. Gill (2002) described the nature of this joint image/text construction of meaning as "distribution". According to Gill (2002), there are two types of distribution. Intra-process distribution refers to the portrayal by images and text of different aspects of a shared process. For example the image(s) might depict the end result of a process described in the verbal text. This occurs in the picture book *Gorilla* (Browne, 1983), when the text indicates that Hannah and the gorilla crept downstairs and Hannah put on her coat and the gorilla put on her father's hat and coat. The image shows them standing in the doorway so dressed. Inter-process distribution occurs when images fill a gap in the ideational flow of meaning in the verbal text. For example, later in the story of

Gorilla, the text indicates that it is time to go home and then indicates that they danced on the lawn, which is clearly in front of Hannah's home. But the text makes no reference to their actually going home. This is conveyed by the juxtaposed image of the gorilla walking along the street with Hannah on his shoulders.

Martinec and Salway (2005) illustrate temporal relations between image and text with a segment showing an image of one of Max Beckmann's paintings and an accompanying main text which begins:

> Beckmann worked for the German army's medical corps during the war, sketching the horrors of what he saw. Following a nervous breakdown, his paintings became harsher… (Martinec & Salway, 2005, p. 351).

The authors indicate that "Following a nervous breakdown" situates in time the example of Beckmann's paintings provided by the image. Martinec and Salway (2005) illustrate enhancement by place with an image of Newcastle Airport (with the name and location of the airport on a sign in the image) and the following accompanying caption: The woman arrived too late to board the flight to Paris (Martinec & Salway, 2005, p. 350).

The framework describing meaning made at the intersection of image and language through *connection* can be summarized in Figure 4.
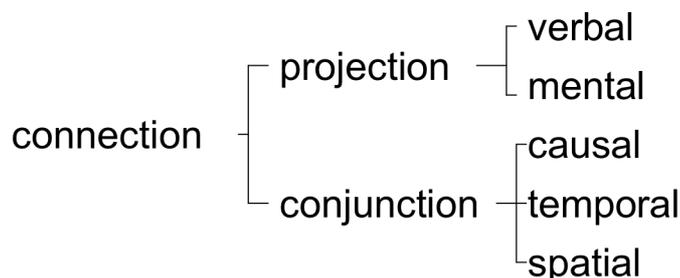

Figure 4. Connection

An overall framework describing ideational meaning-making at the intersection of image and language can be summarized as indicated in Figure 5.

Image-text relations in the construction of interpersonal meaning

Interactive and Evaluative Meaning
Interpersonal meaning in SFL includes interactive and evaluative meaning. Interactive meaning refers to the roles of interactants in giving information (making statements) or providing goods and services (making offers) or demanding information (asking questions) or ordering goods and services (giving commands). These interactive roles are realized grammatically by the mood system (Halliday, 1994). The grammar of visual design proposed by Kress and van Leeuwen (1996) indicates that visually only two interactive roles can be portrayed: a "demand" image has the gaze of one or more represented participants directed to the viewer and hence "demands" some kind of response in terms of the viewer entering into some kind of pseudo-interactive relation with the represented participant; an "offer" does not have

the gaze of any represented participant directed to the viewer and hence provides a portrayal for the viewer's contemplation.
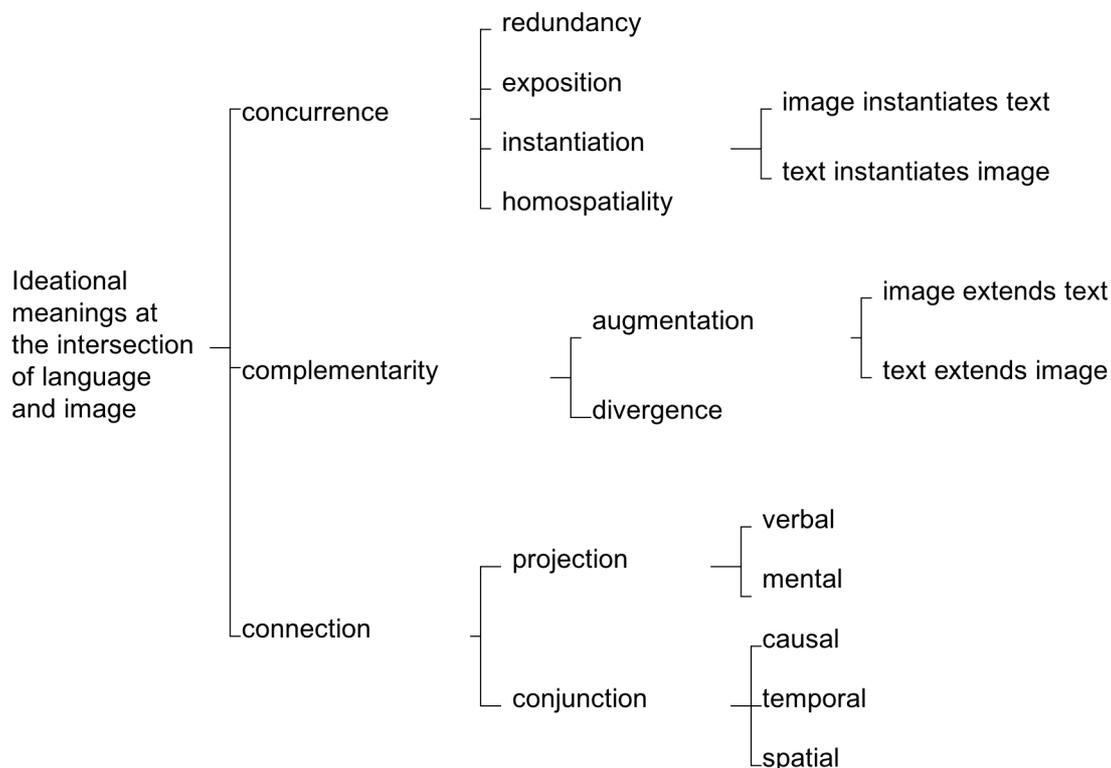
```
                                        ┌ redundancy
                                        │
                            ┌concurrence┤ exposition
                            │           │                    ┌ image instantiates text
                            │           │ instantiation ──────┤
                            │           │                    └ text instantiates image
                            │           └ homospatiality
Ideational                  │
meanings at                 │
the intersection ───────────┤                          ┌ image extends text
of language                 │           augmentation ───┤
and image                   │complementarity│            └ text extends image
                            │           │
                            │           └divergence
                            │
                            │                        ┌ verbal
                            │            projection ──┤
                            │           │            └ mental
                            └connection ┤                      ┌causal
                                        │            │         │
                                        └ conjunction ──────────┤temporal
                                                              └spatial
```

Figure 5. Ideational meanings at the intersection of language and image

Evaluative meaning in SFL has traditionally been confined to commentary on the truth of what is represented linguistically. This is realized by polarity (yes or no) and by the system of modality, which realizes possibilities between positive and negative polarity, such as degrees of certainty and probability (perhaps/of course; possibly/probably/certainly), and degrees of usuality and frequency (sometimes/ usually/always). In the grammar of visual design, evaluation also focuses on the truth or credibility of images, also referred to as modality (Kress & van Leeuwen, 1996). Modality value, however, is related to "coding orientation". Within a naturalistic coding orientation, high modality is a reflection of the fidelity of the representation with the natural world, such as that achieved in a high-quality, colour photograph. Within a scientific coding orientation, fidelity may be calibrated more in relation to the representation of conceptual clarity rather than naturalistic reality.

Martin has extended SFL perspectives on evaluation by proposing an "appraisal network" including three main systems – *attitude, engagement* and *graduation* (Martin, 2000; Martin & Rose, 2003). This work has also been made available in a form easily accessible to teachers and students (Droga & Humphrey, 2002). Here I will deal with the category of attitude only. Within attitude there are a number of sub-categories: Affect refers to the expression of feelings, which can be positive or negative, and may be descriptions of emotional states (for example, happy) or behaviours that indicate an emotional state (for example, "crying"). Sub-categories of Affect are "happiness", "security" and "satisfaction". Appreciation relates to

evaluations of objects, events or states of affairs and can also relate to the characteristics of people but not their behaviour. Appreciation is further subdivided into "reaction", "composition" and "valuation". Reaction involves the emotional impact of the phenomenon (for example, thrilling, boring, enchanting, depressing). Composition refers to the form of an object (for example, coherent, balanced, haphazard) and valuation refers to the significance of the phenomenon (for example, groundbreaking, inconsequential). Judgment can refer to assessments of someone's capacities (brilliant, slow), their dependability (tireless, courageous, rash) or their relative normality (regular, weird). Judgment can also refer to someone's truthfulness (frank, manipulative) and ethics (just, cruel, corrupt). Recent research on interpersonal meaning in image/text relations has noted the joint construction of interaction, but the main impact of these inter-modal relations from an interpersonal perspective seems to be oriented to the construction of evaluative stance in multimodal texts.

Portraying interpersonal interaction through image-text relations

The *Economist* magazine advertisement analysed by Royce (1998) shows a monochrome photograph with a medium to close-up, eye-level view of a young woman whose gaze is directed at the viewer, and whose frontal plane is parallel with that of the viewer. These visual features realize a pseudo interpersonal relation of direct involvement at a personal level with a demand for a response. Positioned immediately above this image is the following question in the largest font on the page: "Does your environmental policy meet your granddaughter's expectations?" Royce points out the ways in which this question, with its second person address and similar features in the subsequent text, effects a joint image/text initiation of interaction, which he refers to as "Reinforcement of Address". Similar work by Cheong (2004) shows how the medium to close-up, eye-level demand image of a smiling young woman whose frontal plane parallels that of the viewer is juxtaposed with the written text positively evaluating the products of the M1 telecommunications company, so that she appears to be the speaker of the quotation: "I get the feeling that M1 wants me to enjoy value – and enjoy life. Everything they offer is brighter, nicer and more fun!" In texts of this kind, the image/text relations are jointly constructing evaluative stance as well as interaction.

Communicating evaluative stance through image-text relations

Gill found that interpersonal alignment could occur across image/verbiage juxtapositions, which she described as a resonance of interpersonal meaning. For example, on pages 27-28 of Anthony Browne's (1983), *Gorilla*, the ideational content does not concur. However, there is a resonance between the image and text construction of the affect portrayed between Hannah and her father. In the text the father says: "Happy birthday, love", and the image shows Hannah with her father putting his hands on Hannah's shoulders. Gill's analysis showed many examples of resonance of appraisal content, such as Affect. For example, on page eleven of *Gorilla*, the text indicates: "Hannah was frightened", corresponding to the image of a frightened Hannah with the bedclothes drawn up over part of her face. On pages 17-18, where Hannah and the gorilla visit the orang-utan and the chimpanzee in the zoo, the text indicates: "She thought they were beautiful. But sad," – corresponding visually to the expression of the orang-utan and to a lesser extent that of the chimpanzee. Similarly, instances of interpersonal resonance with appraisal content were found in the picture book, *the baby who wouldn't go to bed* (Cooper, 1996). For

example, the images consistently depict the other participants as looking down on the baby from a high vertical angle, positioning them as having power over her. This concurs with the mood structure of the text, where the other participants make statements that serve as indirect, disciplinary comments about the baby's behaviour.

As far as interpersonal meaning is concerned, verbiage/image relations in multimodal texts, according to Martin (2002), are more concerned with appraisal than with mood or modality. He argues that a key function of images is to co-articulate attitude (including Affect, Judgment and Appreciation). In doing so, images operate in a similar way to imagery, provoking an evaluative reaction in readers, and the images are typically positioned to do this so that they preview or foreshadow the value positions to be constructed in the subsequent verbiage. One example is taken from Nelson Mandela's *The illustrated long walk to freedom* (Mandela, 1996). In the section dealing with the 1976 Soweto uprising, the well-known photo of the body of thirteen-year-old Hector Pieterson being carried from the fray is positioned as a full-page image on page 147, preceded by its caption in the right hand margin of the previous page. The main text dealing with the Soweto uprising then appears overleaf on page 148. The photo previews and amplifies the reaction induced by Mandela's verbal imagery. In SFL terms, Martin suggests that the photo functions as an evaluative interpersonal Theme, naturalizing the stance from which the remaining verbiage can be read. Additional examples are provided by Martin's (2002) analyses of other sections of this text and further examples from his analyses of the Australian Government Report (1997) *Bringing them home* on the generations of Aboriginal children taken from their families and placed in alternative care. This report similarly deploys images and imagery to establish evaluative orientations to the ensuring text. On the basis of this work, Martin (2002) suggests that for multimodal texts the Given/New elements of the compositional meaning-making resources of images, extrapolated by Kress and van Leeuwen (1996) from SFL, need to be augmented to include a visual version of the SFL concept of Interpersonal Theme. As Martin reasons:

> The left is not simply Given, but has a positive forward looking function, instigating an naturalizing a reading position for the evaluation of verbiage/image texture that ensues (Martin, 2002, p. 334).

Textual/Compositional Meanings in Image/Text relations

The descriptions of compositional meanings in images by Kress and van Leeuwen (1996) have been extensively applied by them to image/text relations. Further studies of school science books have shown how layout resources of Given/New, Ideal/Real, and Framing are deployed to structure pedagogic texts (Veel, 1998). Typically, what is likely to be familiar to students, whether in the form of language or image, is placed in the Given position on the left and that which deals either visually or verbally with unfamiliar, technical information is placed in the New position on the right. While these Given/New structures are consistent with the usual left to right progression in reading, the Ideal/Real structures in school curriculum texts do not necessarily map strategically onto our practice in working from top to bottom of the text. Students might be advised to examine the specific, concrete information of the Real positioned at the bottom of the layout before addressing the more abstract, conceptual, and generalized information of the Ideal positioned at the top. Often the salience of

concrete images in the Real will influence students to adopt such a reading path (Unsworth, 2001).

It has been noted, in Martin's (2002) work above, that the descriptions of the compositional meanings in multimodal texts need to be extended to take account of the role of images as Interpersonal Theme. Further extensions are suggested in the work by Jewitt (2002) dealing with the compositional resources for constructing character in the Novel as CD ROM version of Steinbeck's *Of mice and men* (SteinbeckSeries, 1996), and by issues of "framing" raised in Macken-Horarik's (2003b) study of texts which were central to the children overboard affair.

From her *Of mice and men* study of the Novel as CD ROM, Jewitt (2002) hs suggested that the spatial relationship between image and verbiage on each of the screens is itself a meaning-making resource. She argued that writing serves as a visual element, a block of "space" that makes textual meaning beyond its content. Jewitt indicated that on the CD screens, the blocks of writing were positioned in different places: the left or right side, along the bottom or top length of the screen, or in the top or bottom corner. The size and position of the block and its location combined to reveal or conceal different parts of the image layered "beneath it".   In this way, a block of writing emphasizes different aspects of the image on screen. According to Jewitt, the image at times cuts across the lexis and grammar of the written text to create a visual mood and rhythm, which she illustrated with one image of George and Lenny that runs across four screens of changing text:

> In the first screen, the block of writing sits above George's head as he talks to Lennie about what he could do if he left him. In the second screen Lennie is visually obliterated by George's angry talk of leaving, visually foregrounding George. In the third screen, as George's anger subsides, the block of writing is placed on the screen so that both George and Lennie are visible (Jewitt, 2002, p. 184).

Jewitt suggested that it is through the visual arrangement of image and writing on screen that the narrative construct of character indicated intensity of emotion to suggest the alignment of the viewer with George's point of view, and to emphasise the agency/passivity of the characters in the novel. Whether negotiating the meaning of newspaper stories, or literary narratives (in book or electronic media), layout features such as framing are crucial elements in the interpretation of the meanings at stake and in establishing the evaluative stance of the writer in relation to those meanings (Macken-Horarik, 2003b; Unsworth, 2006a).

Although the ideational, interpersonal and compositional perspectives on the meaning-making resources of image/language interaction have been discussed separately here, it must be remembered that in reality these meanings are always made simultaneously in all texts, and critical understanding of the interpretive possibilities of texts needs to be based on an integrative view of all three perspectives.

CONCLUSION: THE PEDAGOGIC POTENTIAL OF A SYSTEMIC FUNCTIONAL PERSPECTIVE ON METALANGUAGE AND MULTIMODALITY

There seems to be significant support for the view that the need to redefine literacy in the electronic age entails the development of a metalanguage that will facilitate metatextual awareness of image/text relations (Kamil *et al.,* 2000; Kress, 2003b; Macken-Horarik, 2004; Richards, 2001; Russell, 2000).   Metalanguage entails systematic, technical knowledge of the ways in which the resources of language and images (and other semiotic systems) are deployed in meaning making. English syllabi currently require a significant commitment by teachers and students to understanding and using metalanguage. Such an investment in teaching and learning can be productive if the metalanguage functions as a tool to enhance the development of critical social literacies.

For this to happen, the metalanguage must be based on systematic accounts of the meaning-making potential of the multimodal nature of contemporary texts and also be capable of expansion/modification in response to the expansion of meaning-making potential with the ongoing emergence of new forms of communication. This paper suggests that systemic functional semiotic theory has much to offer in this respect. However, the work on grammars for exploring the co-articulation of image and verbiage is in its infancy (Kress, 2001; Macken-Horarik, 2003a). Little classroom research has been done on the pedagogic use of such emerging grammars, although there is some evidence that young children can learn and productively use aspects of Kress and van Leeuwen's visual grammar in work with picture-books and with multimedia CD ROMs in curriculum area learning (Callow & Zammit, 2002; Howley, 1996). There is also a good deal of evidence for the efficacy of the metalanguage of SFL in literacy development and learning in primary/elementary and secondary/high school contexts (Quinn, 2004; Schleppegrell, 2004; Schleppegrell *et al.*, 2004; Torr & Harman, 1997; Williams, 1999, 2000). What is suggested here is that the theoretical bases of the social semiotic research arising from SFL are providing a generative and inclusive framework for the transdisciplinary development of a metalanguage of multiliteracies.

While the research on an evolving metalanguage of multimodality is in the very early stages and emerging descriptions remain quite tentative, there are at least two, firm, practical implications for English teachers. The first is the robustness, broad application, and practical usefulness of the metafunctional principle deriving from SFL. That is the principle that all texts, visual and verbal, separately and in combination, always simultaneously entail ideational, interpersonal and textual/compositional meanings. This principle is frequently reflected in the rationale of the English syllabi of different school systems. For example, the current English 1-10 English Curriculum for Queensland Schools in Australia indicates in its rationale:

> We use language purposefully to represent experiences of real and imagined worlds, to interact with others, and to create coherent and cohesive texts (QueenslandStudiesAuthority, 2005, p. 1).

The metafunctional principle is widely accepted as central to our understanding of contemporary and emerging forms of multimodal texts and provides a sound and accessible basis for English teachers to further examine the pedagogic potential of metalanguage. Detailed accounts of the ways in which SFL and the grammar of visual design can be used together in the English classroom are now well documented (Callow, 1999; Christie, 2005; Christie & Unsworth, 2005; Cope & Kalantzis, 2000; Goodman & Graddol, 1996; Jewitt, 2005; Jewitt & Kress, 2003; Kress, 2003b; Unsworth, 2001, 2006b; Unsworth *et al.*, 2005).

The second practical implication for teachers is their essential role as participants in the collaborative researching, theorising and re-formulating of our "metasemiotic" understanding of emergent, multimodal text-forms and the concomitant derivation of an evolving metalanguage of multimodality to enhance practical pedagogy. The interface of teaching and research has long been an essential characteristic of SFL work (Christie, 2005; Christie & Unsworth, 2005), and this continues to be the case in a great deal of the current research contributing to multimodal text description (Macken-Horarik, 1996, 1998, 2003a, 2003b, 2004). It is hoped that this article, by stimulating critically constructive responses to, and envisioning beyond, what is presented here, will encourage further collaborative work among teachers, teacher educators and researchers in other disciplines in exploring the nature and role of a metalanguage that will facilitate development of the multiliteracies pedagogies appropriate to the multimedia world of our children in the Twenty-First Century.

Are there any sustainable arguments for a positive relationship between knowledge about language (however understood) and increased effectiveness in some aspect of textual practice (reading/viewing or production)?

REFERENCES

Baldry, A. (Ed.). (2000). *Multimodality and multimediality in the distance learning age*. Campobasso, Italy: Palladino Editore.

Browne, A. (1983). *Gorilla*. London: Julia MacRae.

Browne, A. (1994). *Zoo*. London: Random House.

Burningham, J. (1977). *Come away from the water, Shirley*. London: Cape.

Burningham, J. (1978). *Time to get out of the bath, Shirley*. London: Cape.

Callow, J. (Ed.). (1999). *Image matters: Visual texts in the classroom*. Sydney: Primary English Teaching Association.

Callow, J., & Zammit, K. (2002). Visual literacy: From picture books to electronic texts. In M. Monteith (Ed.), *Teaching primary literacy with ICT* (pp. 188-201). Buckingham: Open University Press.

Caswell, B., & Ottley, M. (2003). *Hyram and b*. Sydney: Hodder Headline.

Cheong, Y. (2004). The construal of ideational meaning in print advertisements. In K. O'Halloran (Ed.), *Multimodal discourse analysis: Systemic functional perspectives* (pp. 163-195). London and New York: Continuum.

Christie, F. (2005). *Language education in the primary years*. Sydney: University of New South Wales Press.

Christie, F., & Unsworth, L. (2005). Developing dimensions of an educational linguistics. In J. Webster, C. Matthiessen & R. Hasan (Eds.), *Continuing*

discourse on language: A functional perspective (Vol. 1, pp. 217-250). London: Equinox.

Commission, H. R. & E. O. (1997). *Bringing them home: National inquiry into the separation of Aboriginal and Torres Strait Islander children from their families.* Sydney: Human Rights and Equal Opportunities Commission.

Cooper, H. (1996). *The baby who wouldn't go to bed.* London: Doubleday/Picture Corgi Books.

Cope, B., & Kalantzis, M. (Eds.). (2000). *Multiliteracies: Literacy learning and the design of social futures.* Melbourne: Macmillan.

Droga, L., & Humphrey, S. (2002). *Getting started with functional grammar.* Marrickville, Australia: Target Texts.

Education, Q. D. of (1995). *English 1-10 syllabus: A guide to analysing texts.* Brisbane: Queensland Government Printing Office.

Gill, T. (2002). *Visual and verbal playmates: An exploration of visual and verbal modalities in children's picture books.* Unpublished B.A. (Honours), University of Sydney.

Goodman, S., & Graddol, D. (1996). *Redesigning English: New texts, new identities.* London: Routledge.

Halliday, M. (1994). *An introduction to functional grammar* (2 ed.). London: Edward Arnold.

Halliday, M., & Hasan, R. (1985). *Language, context and text: Aspects of language in a social-semiotic perspective.* Geelong: Deakin University Press.

Halliday, M., & Matthiessen, C. (2004). *An introduction to functional grammar* (3rd ed.). London: Arnold.

Howley, P. (1996). *Visual literacy: Semiotic theory, primary school syllabus documents and classroom practice.* Unpublished Bachelor of Education Honours thesis, University of Sydney, Sydney.

Jewitt, C. (2002). The move from page to screen: The multimodal reshaping of school English. *Visual Communication, 1*(2), 171-196.

Jewitt, C. (2005). *Technology, literacy, learning.* London: Routledge.

Jewitt, C., & Kress, G. (Eds.). (2003). *Multimodal literacy.* New York: Peter Lang.

Kamil, M., Intrator, S., & Kim, H. (2000). The effects of other technologies on literacy and learning. In M. Kamil, P. Mosenthal, P. Pearson & R. Barr (Eds.), *Handbook of reading research* (Vol. 3, pp. 771-788). Mahwah, New Jersey: Erlbaum.

Kress, G. (1997). Visual and verbal modes of representation in electronically mediated communication: The potentials of new forms of text. In I. Snyder (Ed.), *Page to screen: Taking literacy into the electronic era* (pp. 53-79). Sydney: Allen and Unwin.

Kress, G. (2000a). Design and transformation: New theories of meaning. In B. Cope & M. Kalantzis (Eds.), *Multiliteracies: Learning literacy and the design of social futures* (pp. 153-161). Melbourne: Macmillan.

Kress, G. (2000b). Multimodality. In B. Cope & M. Kalantzis (Eds.), *Multiliteracies: Literacy learning and the design of social futures* (pp. 182-202). Melbourne: Macmillan.

Kress, G. (2001). Sociolinguistics and social semiotics. In P. Cobley (Ed.), *Semiotics and linguistics* (pp. 66-82). London: Routledge.

Kress, G. (2003a). Genres and the multimodal production of "scientificness". In C. Jewitt & G. Kress (Eds.), *Multimodal literacy* (pp. 173-186). New York: Peter Lang.

Kress, G. (2003b). *Literacy in the new media age*. London: Routledge.

Kress, G., & van Leeuwen, T. (1996). *Reading images: A grammar of visual design*. London: Routledge.

Lemke, J. (1998a). Metamedia literacy: Transforming meanings and media. In D. Reinking, M. McKenna, L. Labbo & R. Kieffer (Eds.), *Handbook of literacy and technology: Transformations in a post-typographic world* (pp. 283-302). New Jersey: Erlbaum.

Lemke, J. (1998b). Multiplying meaning: Visual and verbal semiotics in scientific text. In J. R. Martin & R. Veel (Eds.), *Reading science: Critical and functional perspectives on discourses of science* (pp. 87-113). London: Routledge.

Lemke, J. (2002). Travels in hypermodality. *Visual Communication, 1*(3), 299-325.

Lim, V. F. (2004). Developing an integrative multi-semiotic model. In K. O'Halloran (Ed.), *Multimodal discourse analysis: Systemic functional perspectives* (pp. 220-246). London and New York: Continuum.

Macken-Horarik, M. (1996). Literacy and learning across the curriculum: Towards a model of register for secondary school teachers. In R. Hasan & G. Williams (Eds.), *Literacy in society* (pp. 232-278). Harlow: Addison Wesley Longman.

Macken-Horarik, M. (1998). Exploring the requirements of critical literacy: A view from two classrooms. In F. Christie & R. Misson (Eds.), *Literacy and schooling* (pp. 74-103). London: Routledge.

Macken-Horarik, M. (2003a). A telling symbiosis in the discourse of hatred: Multimodal news texts about the "children overboard" affair. *Australian Review of Applied Linguistics, 26*(2), 1-16.

Macken-Horarik, M. (2003b). Working the borders in racist discourse: The challenge of the "children overboard affair" in news media texts. *Social Semiotics, 13*(3), 283-303.

Macken-Horarik, M. (2004). Interacting with the multimodal text: Reflections on image and verbiage in *artexpress*. *Visual Communication, 3*(1), 5-26.

Mandela, N. (1996). *The illustrated long walk to freedom: The autobiography of Nelson Mandela*. London: Little, Brown and Company.

Martin, J. (1992). *English text: System and structure*. Amsterdam: Benjamins.

Martin, J. (2000). Beyond exchange: Appraisal systems in English. In S. Hunston & G. Thompson (Eds.), *Evaluation in text: Authorial stance and the construction of discourse* (pp. 142-175). Oxford: Oxford University Press.

Martin, J. (2002). Fair trade: Negotiating meaning in multimodal texts. In P. Coppock (Ed.), *The semiotics of writing: Transdisciplinary perspectives on the technology of writing* (pp. 311-338). Begijnhof, Belgium: Brepols & Indiana University Press.

Martin, J. (2003). Voicing the "other": Reading and writing indigenous Australians. In G. Weiss & R. Wodak (Eds.), *Critical discourse analysis: Theory and interdisciplinarity* (pp. 199-219). London: Palgrave.

Martin, J., & Rose, D. (2003). *Working with discourse: Meaning beyond the clause* (1st ed. Vol. 1). London/New York: Continuum.

Martinec, R. (1999). Cohesion in action. *Semiotica, 1/2*, 161-180.

Martinec, R. (2000a). Rhythm in multimodal texts. *Leonardo, 33*(4), 289-297.

Martinec, R. (2000b). Types of process in action. *Semiotica, 130*(3/4), 243-268.

Martinec, R., & Salway, A. (2005). A system for image-text relations in new (and old) media. *Visual Communication, 4*(3), 337-371.

McCloud, S. (1994). *Understanding comics: The invisible art*. New York: Harper Collins.

New South Wales Board of Studies. (1998). English K-6 syllabus and support documents. Retrieved 7th September, 2005, from http://k6.boardofstudies.nsw.edu.au/english/english_index.html

O'Halloran, K. (Ed.). (2004). *Multimodal discourse analysis: Systemic functional perspectives*. London and New York: Continuum.

O'Toole, M. (1994). *The language of displayed art*. London: Leicester University Press.

Queensland Studies Authority. (2005). *Years 1-10 English syllabus.* Retrieved 7th September, 2005, from http://www.qsa.qld.edu.au/yrs1to10/kla/english/syllabus.html

Quinn, M. (2004). Talking with Jess: Looking at how metalanguage assisted explanation writing in the middle years. *Australian Journal of Language and Literacy, 27*(3), 245-261.

Richards, C. (2001). Hypermedia, internet communication, and the challenge of redefining literacy in the electronic age. *Language Learning and Technology, 4*(2), 59-77.

Royce, T. (1998). Synergy on the page: Exploring intersemiotic complementarity in page-based multimodal text. *Japan Association Systemic Functional Linguistics Occasional Papers, 1*(1), 25-50.

Royce, T. (2002). Multimodality in the TESOL classroom: Exploring visual-verbal synergy. *TESOL Quarterly, 36*(2), 191-205.

Russell, G. (2000). Print-based and visual discourses in schools: Implications for pedagogy. *Discourse: Studies in the Cultural Politics of Education, 21*(2), 205217-.

Schleppegrell, M. (2004). *The language of schooling: A functional linguistic perspective*. Mahwah, New Jersey and London: Erlbaum.

Schleppegrell, M., Achugar, M., & Oteíza, T. (2004). The grammar of history: Enhancing content-based instruction through a functional focus on language. *TESOL Quarterly, 38*(1), 67-93.

Sendak, M. (1962). *Where the wild things are*. London: The Bodley Head.

SteinbeckSeries. (1996). Of mice and men. New York: Penguin Electronics.

Torr, J., & Harman, J. (1997). Literacy and the language of science in year one classrooms: Implications for children's learning. *Australian Journal of Language and Literacy, 20*(3), 222-237.

Unsworth, L. (2001). *Teaching multiliteracies across the curriculum: Changing contexts of text and image in classroom practice*. Buckingham, United Kingdom: Open University Press.

Unsworth, L. (2004). Comparing school science explanations in books and computer-based formats: The role of images, image/text relations and hyperlinks. *International Journal of Instructional Media, 31*(3), 283-301.

Unsworth, L. (2006a). *Describing meaning-making at the intersection of language and image: Towards a metalanguage for multi-modal literacy pedagogy.* Paper presented at the Future Directions in Literacy, University of Sydney.

Unsworth, L. (2006b). *E-literature for children: Enhancing digital literacy learning*. London and New York: Routledge/Falmer.

Unsworth, L. (in press). Explicating inter-modal meaning-making in media and literary texts: Towards a metalanguage of image/language relations. In A. Burn & C. Durrant (Eds.), *Media teaching: Language, audience, production*. London: AATE-NATE/Wakefield Press.

Unsworth, L., Thomas, A., Simpson, A., & Asha, J. (2005). *Children's literature and computer based teaching*. London: McGraw-Hill/Open University Press.

van Leeuwen, T. (1999). *Speech, music, sound*. London: Macmillan.

Veel, R. (1998). The greening of school science: Ecogenesis in secondary classrooms. In J. Martin & R. Veel (Eds.), *Reading science: Functional and critical perspectives on the discourses of science* (pp. 114-151). London: Routledge.

Williams, G. (1999). Children becoming readers: Reading and literacy. In P. Hunt (Ed.), *Understanding children's literature* (pp. 151-162). London: Routledge.

Williams, G. (2000). Children's literature, children and uses of language description. In L. Unsworth (Ed.), *Researching language in schools and communities: A functional linguistic perspective* (pp. 111-129). London: Cassell.