Chapter 1

Introduction and Theoretical Foundations

Outline

1.1 Introduction to the thesis
   1.1.1 Motivations, aims and data
   1.1.2 Research questions
1.2 Summary of theoretical foundations
   1.2.1 The SFL model
   1.2.2 Interpersonal meaning
   1.2.3 Visual semiosis
   1.2.4 Intersemiosis
1.3 Organisation of thesis

1.1 Introduction to the thesis

1.1.1 Motivations, aims and data
This thesis aims to develop further theoretical tools for analysis of visual meanings (visual semiosis) to use alongside the sophisticated tools already available for language (verbal semiosis). More broadly it is concerned with expanding present understandings of knowledge and literacy, in which the visual has come to play an increasingly important role. The more specific theoretical aims of the thesis are to describe and explain evaluative meanings in one kind of visual semiosis, that of naturalistic news photos, and in one kind of verbal-visual intersemiosis, that involving the news photo(s), captions and headlines that introduce long news review feature articles. Put simply, the thesis asks: how are values, beliefs, attitudes or feelings expressed in news photos and how do these interact with surrounding words?
Consider the verbal evaluative meanings in the introduction to two news feature stories about the asylum seeker issue, as shown in Example 1.1 below, from an Australian newspaper and Example 1.2 from a Greek newspaper. In the first story, *Wait in fright*, a strong negative feeling in the attitudinal word *fright* is explicitly ascribed in the headline to someone not referred to in that headline. In the second story, “*Slightly legal*” *immigrants* (my translation), a negative opinion of this way of classifying immigrants is less explicitly expressed by the odd pairing of the first two words and ‘scare’ quotes. Identifying and describing these different types of evaluative meaning are important in explaining how readers understand and are positioned by these words in respect to the knowledge also represented here. But in both cases, the headlines are accompanied by striking images, and it is obvious that, besides providing visual information not in the headlines, evaluation is also an important element in these images. It is important both to the positioning of readers in respect to what is depicted in them, and in respect to what is written in words around them.

Clearly, various evaluative meanings are made in the photos in these two examples and they have a significant effect on meanings in the nearby headlines, and due to their
prominence, also to meanings made in the story. Looking at each text, we can see that, though the prominent photo in each has similar ‘content’ (each seems to depict ‘foreign’ people in detention), quite different visual choices have been made. In one, two adults are depicted clearly in bright daylight but from behind and cropped so that they have no heads. In the other, two children are depicted in a less focused photo, with unusual light and colour, but in an extreme frontal close up giving us direct eye contact with them. Theoretically, analytically and practically, we cannot ignore these differences and their evaluative effect on readers, yet accounts of such choices in analysis of media discourse are limited. We need to be able to give much fuller accounts if we are to understand visual meaning as we have come to understand verbal. So far there is no comparable means of describing and analysing evaluation in visual semiosis. This thesis addresses this analytic and descriptive gap.

At present a detailed and principled analysis of a wide range of types of evaluation and the forms it takes in language is available. The systemic-functional linguistics (SFL) Appraisal system offers a comprehensive and systematic tool for doing this. However, in visual analysis we have not gone far beyond commonsense observations. Visual meanings in general, and in naturalistic photos in particular, are often still taken for granted and not given the same analytic scrutiny as verbal, even while all around us their use is rapidly increasing in all kinds of discourse. Further, while many tools have been developed for semiotic analysis of verbal discourse and specifically news discourse, far fewer have been directed at news photos or developed for visual meaning. For example, SFL grammar and register descriptions have been used for critical discourse analysis since the 1970s (Trew, 1979; Fowler, Hodge, Kress and Trew, 1979) to reveal implicit evaluation or ideology in seemingly ‘objective’ news stories, but few, even today, examine news photos. In journalism scholarship, calls for more attention to be paid to the evaluative function of news photos have only recently been made (Zelizer, 1992, 1995, 1998, 2004a, b, 2005, 2007; Zelizer, et al 2002). In descriptions of visual meaning, even the widely used ‘visual grammar’ generated by applying certain SFL notions and categories (Kress and van Leeuwen, 2006) explores evaluative meaning the least. The challenge is clear. Unless we can describe evaluative meaning in news photos in a theoretically informed way, we cannot do satisfactory analysis of news media texts that use photos. The ever-increasing use of visuals in the news and many other contemporary contexts makes it a matter of some urgency to incorporate, better and
more fully, visual evaluative meaning into our presently verbally-based descriptions of discourse.

In addition to developing analysis of visual evaluative meaning, the thesis addresses its overarching aims concerning knowledge and literacy by focussing on the informal pedagogic context of the broadsheet mass news media. In doing so it draws on sociologist Bernstein’s definition of pedagogic as any discourse “that shapes people’s consciousness” and the pedagogic process as one in which “power relations are transformed ... into discourse and discourse into power relations” (1996: 12). He specifically refers to the public media as a context, beyond family and school, in which the pedagogic process occurs, albeit in “more generalised and diffuse forms” (1996: 12). Recognising this, the thesis investigates news photos; and rather than exploring how they are used in front-page breaking news or news stories, it concerns itself with how they are used to introduce long articles in that part of the paper we trust the most – the weekly news review. Though referred to as news feature or investigative stories, these articles are not ‘stories’ in any strict sense. They deal with recent events and related issues judged (by the newspaper) as especially significant for our community and so are considered civic journalism texts. We trust these articles to be more considered and better researched than daily news stories, more balanced and objective than op-ed pieces or news commentaries. Their epistemic status in the community is high. We go to them for fresh knowledge on serious issues, knowledge not settled enough to be in formal pedagogic texts such as academic journals and books. They are also brought into academic contexts for that reason, often without any distinctions drawn between knowledge in these news texts and in formal academic texts. Perhaps better labelled ‘news analysis’ (Iedema, Feez and White, 1994; Martin and White, 2005) these long news texts have received the least attention in descriptions of news discourse, where the focus has typically been on daily news reports or news commentaries.

The primary data set for this inquiry comprises eight broadsheet news review stories on the issue of asylum seekers, four from an Australian and four from a Greek broadsheet. The specific focus of inquiry is the large verbal-visual “display” introducing the news

---

1 Though anecdotal, my own observations of broadsheet news review stories used in the Australian tertiary education context include a sociology lecture in which a front page story was used as visual aid as well as quoted and summarised; a pharmacy writing assessment task requiring a response to a quote from a news feature story; and student reading lists across a number of disciplines that included such stories.
feature story. These displays are called ‘standouts’ in this thesis. The standouts on the front page of the review are the largest in the news review and introduce the longest article, written by the most respected journalists. If readers trust in the news review, its stories and writers, then they typically extend this trust also to the standouts, particularly where they use ‘factual’ news photos. However, not only has the news photo itself and its interaction with surrounding verbal text been paid scant attention in news discourse analysis, but no attention at all has been given to news review standouts or to any news texts with a large visual component.

Although SFL analysis, particularly appraisal analysis of news texts, has been productive in deconstructing evaluation and stance in verbal news text (Martin and White, 2005; Martin, 2004a, 2005; White, 2003a,b, 2006) little investigation using these tools has been carried out on visual or verbal-visual news. This is a significant limitation given that most news now has a visual component and in some texts, like standouts, it is the most prominent component. Evaluative analysis is especially important in these news review standouts, which have wide reach and are associated with civic journalism that has influence in our community. How such standouts create a stance towards knowledge, and align readers with this stance through image and language warrants our close attention. We need to describe and explain evaluation not only in the language, but also in the photos used in news review standouts.

The first step taken by this thesis therefore is to explore evaluative meaning in news photos, using a data set of 450 news photos, most from the Australian broadsheet The Sydney Morning Herald. Deploying SFL descriptions as a model and heuristic, in particular the SFL appraisal system, different visual categories are proposed, tested and developed for describing evaluation in news photos in Chapters 3-6. These categories are then used to describe news photos and their interaction with language in a set of verbal-visual standout texts in Chapter 7. The research questions that directed these two phases of the thesis project are presented below, followed by a summary of its theoretical foundations.

---

2 The name ‘standout’ was suggested by Helen Caple, personal communication, 2004
1.1.2 Research questions

The overarching question:
How do verbal-visual displays introducing civic journalism texts in prestige broadsheet sites create an evaluative stance in respect to critical social issues and align a mass readership with this stance?

Theoretical/Analytical questions

About evaluation:
1. How are evaluative meanings realised by visual resources in the naturalistic news photo?

2. How are evaluative meanings co-articulated across visual and verbal semiotic modalities in a verbal-visual print text, and how are they distributed within and across all verbal-visual components in the text to construct the overall evaluative stance?

About news discourse:
1. Can a schematic structure be identified for the verbal-visual ‘standout’ as a distinctive news text type, and how does it relate to the written story it introduces?

2. How is evaluative stance constructed in a series of standouts on the same issue in a similar news context in two different countries, and how does the standout stance compare to the stance in the ensuing written story?

3. How do differences between the two contexts impact on stance in each set of standouts?

1.2 Summary of Theoretical Foundations

1.2.1 The SFL Model

The Systemic Functional model used in this thesis has language as its primary object of study. In order to discuss how this model is relevant for the exploration of visual and
verbal-visual texts, it is first necessary to outline key parameters of the model and then its potential adaptations to visual and visual-verbal meaning.

One fundamental principle of the SFL model is its view of language as a resource, with every human language seen as a vast ‘meaning potential’ from which users choose when they engage in language behaviour. At the same time, the model views language in use as text, with every text seen as an instance of the meaning potential available to users. These two perspectives on language – as underlying potential and as actualised instance of text – are argued to be complementary views of the same phenomenon. The instantiation relation between them is likened to that between climate and weather: the former, the seemingly more stable general underlying patterns and the latter, the ever-changing daily local conditions (Halliday & Matthiessen, 1999: 328; 2004: 26-29). What distinguishes SFL from other semiotic and social semiotic models is the attention it gives to both these perspectives. It is equally concerned with building up descriptions of the meaning potential of a language, and with closely analysing individual textual instances.

As an instance of language in use, a text is always situated in a specific social context and SFL models this as context of situation (Halliday and Hasan, 1975), identifying three variables, or contextual dimensions that impact on language use: those of field, tenor and mode. Presented more fully in Table 1.1 below, these can be broadly glossed in terms of field being to do with what is going on in the situation, tenor with who is taking part, and mode with what part the language is playing.

Table 1.1: Variables of the situational context

<table>
<thead>
<tr>
<th>Context of Situation</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Field</strong></td>
</tr>
<tr>
<td>the social institution ; the social activity including subject matter</td>
</tr>
<tr>
<td><strong>Tenor</strong></td>
</tr>
<tr>
<td>temporary and permanent social roles and role relationships of participants including power and solidarity</td>
</tr>
<tr>
<td><strong>Mode</strong></td>
</tr>
<tr>
<td>medium and channel the language uses; how much language is part of action</td>
</tr>
</tbody>
</table>
SFL theory argues that contextual choices of field, tenor and mode can be explicitly related to linguistic choices in the text. One way the model makes explicit this context-text hook up is by relating each contextual dimension to a different kind of meaning available in the potential. Referring to these three broad areas of meaning as metafunctions, field is shown to be associated with the ideational metafunction, tenor with the interpersonal and mode with the textual. Each metafunction is described in Table 1.2:

### Table 1.2 The three metafunctions of language

<table>
<thead>
<tr>
<th>Context of Situation</th>
<th>Metafunctions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Field</td>
<td>Ideational: representing the material world and our experience of it</td>
</tr>
<tr>
<td>Tenor</td>
<td>Interpersonal: acting upon each other to create our social world</td>
</tr>
<tr>
<td>Mode</td>
<td>Textual: enabling the other two metafunctions by creating text</td>
</tr>
</tbody>
</table>

Though the relationship between context and meaning is commonly referred to as realisation, SFL emphasises its mutually determining nature whereby meaning both construes and realises context. This mutual relationship is indicated by a double-headed arrow in Table 1.2. Most socio-semiotic approaches have adopted this model of the text-context relationship, referred to as SFL register theory, to explain how and why language used in different contexts varies in systematic ways (Eggins and Martin, 1997). The particular configuration of field, tenor and mode values in a context constrains what can be meant and said in the context and at the same time, what is meant and said can also constitute that context.

The SFL model also uses sociologist Bernstein’s notion (1971, 1975, 1996: 157) of code or coding orientation to explain another kind of variation to do with the language user him/herself. Bernstein shows that an individual’s social positioning, predominantly (but not solely) in terms of socio-economic class, determines what meanings are seen by the
individual as at risk in a specific context. The meanings available to, and seen as appropriate by one individual in a situation may thus be different from those available to, and seen as appropriate by another individual in the same situation. Work on coding orientation identifies how differential access is provided to the meaning potential of a culture according to an individual’s social position. The orientation to meaning acquired in early family socialisation has been argued in many studies to affect children’s success in the school context, depending on their alignment with the code assumed or required at school\(^3\) (Cloran, 1999; Hasan, 1992, 1996, 2005; Williams, 1999, 2001).

Work on social factors influencing coding orientation in SFL relates to work on subjectivity and ideology developed in other frameworks such as Critical Discourse Analysis (CDA), as discussed in Chouliaraki and Fairclough (1999: 139-155) and Martin and Wodak (2003). Recent SFL work uses the term \textit{individuation} (Martin, 2008c) to refer to the process by which a user develops over a lifetime – from family, through school to later work and leisure affiliations – their own individual ‘repertoire’ of resources associated with contexts, from the large ‘reservoir’ of meaning resources available to the whole culture.

\textbf{Genre} is posited by SFL linguist, Martin, as an additional layer of contextual potential (at a more abstract level than register) to account for rhetorical aspects of a text not accounted for by the register variables of field, tenor and mode (Martin, 1992: Chapter 7; Martin 1999a, Eggins and Martin, 1997). Martin defines genre as the relatively finite set of social purposes and ways of achieving them available in any culture, though the term genre is also used for the relatively predictable set of staged text structures that realise these social goals and processes. Genres and their associated text stages (e.g. narrative, comprising orientation\(^\ast\) complication\(^\ast\) resolution; and exposition, comprising thesis\(^\ast\) argument\(^\ast\) thesis reiteration\(^4\)) are thus comparable even across social contexts where register variables differ. The two genres of narrative and exposition have been described in the formal education context and in the news media context examined in this thesis. Both register and genre are also seen to work together in text, in that each

---

\(^3\) What Bernstein refers to as the ‘elaborated’ code/coding orientation and found in formal schooling contexts was not found in all home contexts. It was strongly associated only with families of ‘higher autonomy’ professions.

\(^4\) The symbol \(^\ast\) is used to represent ‘followed by’.
functional stage of a genre is realised in text by the changing patterns of register variables within a single social context.

SFL accounts of genre take into account relationships between genres, such as the combining of basic or ‘elemental’ genres within a text. SFL genre theory describes the different ways in which elemental genres (like the recount) can be combined through layering or embedding and/or sequencing within larger *macrogenres* (Martin and Rose 2003: 209-19; Martin, 1999b). For example, the longer news feature story examined in this thesis is described as a macrogenre in Iedema, Feez and White (1994:166-187) since it incorporates a number of elemental genres. The mixing of genres is revealed by discourse analysis studies as common in contexts where rapid social change exerts pressures on conventional social processes, such as the contemporary media context explored in this thesis.

SFL models the realisation of different genres within a larger text by referring to theories of dialogism (Bakhtin, 1986) and intertextuality (Fairclough, 1992a, 1992b: Chapter 6, 1992c: 110-135). Every text is seen as dialogic in nature; intertextuality refers to the incorporation of external voices into a text, in relation to which the internal textual voice is positioned. Bakhtin (2000: 85) also refers to secondary more complex genres which “absorb and digest various ... primary (simple) genres”. More recently, intertextual sourcing has been identified in SFL as the instantiation of external meanings in a text from potentially any semantic stratum, including the higher level of context. The way one text is said to be ‘inspired’ (Martin, to appear) by another can thus be explained in terms of external sourcing at the level of external genre or register. Also, the author’s selection of which external context or text to bring into a text has been explained in terms of individuation. That is, it is likely to be chosen to more easily align putative readers around shared values and in this way, to serve the writer’s interests (Martin, to appear).

Moving away from SFL’s concern with the relationship between language and society, or text and context, we return to its concern with language as a meaning resource. The meaning potential of language is described in SFL in terms of interrelating sets of options, with each set of options organised as a *system*. Using the formalism of the *system network*, as illustrated in Figure 1.1, possible meanings are shown as options...
available at different levels of ‘delicacy’. That is, the selection of an option at the entry level of the system (on the left) leads to further, more delicate sub-options (towards the right)\(^5\). For example in Figure 1, representing the grammatical Mood system, the choice of ‘indicative’ leads to a further choice of ‘declarative’ or ‘interrogative’, as well as the choice of ‘modalised’ or not. This kind of description of language also makes clear the relationship between systems of meaning (paradigms) and the linguistic structures (syntags) that are the output of choices from available systems. That is, a choice from a grammatical system (e.g. Mood) may have a specific structural realisation, such as the presence of a grammatical element or the sequential ordering of elements of grammatical structure. Thus, while SFL is known for bringing to language description a paradigmatic perspective and typological descriptions through system networks, the syntagmatic aspect is equally important in its description of language (Halliday and Matthiessen, 2004: 19-24).

**Figure 1.1 Simplified (partial) Mood system**

\(^5\) Names of options in a network, also termed ‘features,’ are conventionally presented in square brackets when referred to in the body of a text where necessary for clarity.
Syntagmatic structures involve part-whole constituency relations, whereby each constituent of a structure is made up of constituents from a lower level or **rank**. For example, a grammatical clause is made up of groups or phrases, which in turn are made up of words, which in turn are made up of morphemes – on a four-tier grammatical rank scale (Halliday, 1976a; Halliday and Mathiesson, 2004: 31-33). It should be noted that while description in terms of paradigms of meaning can be readily applied to other semiotics, the application of notions of structure and rank may turn out to be problematic in relation to visual meaning. Our present notions of structure, constituency and rank are based on language, which is for the most part expressed in linear and typological fashion, whereas visual meaning is for the most part expressed spatially and topologically. While acknowledging this, this thesis applies the notion of constituent structure heuristically as a useful way in to the examination of visual texts.

The systems of options in SFL descriptions of language are organised into three relatively independent networks, each relating to one of the three metafunctions – the interpersonal, the ideational and the textual – with those of Mood and Modality illustrated in Figure 1.1 belonging to the interpersonal metafunction. This notion of the metafunctional diversity of meaning systems has proved a fruitful one in the transferring of the SF linguistic framework to analyses of the visual. The theory proposes that while the systems **within** each metafunction are highly interdependent, choices from different metafunctions are relatively independent of one another. It is essential to keep in mind, however, that choices from all three metafunctions interplay in the production of text.

The SFL metafunctional principle not only concerns the grouping of paradigmatic choices according to three regions of meaning (interpersonal, ideational and textual) but argues that each metafunction is associated with its own structuring principles (Halliday, 1979, 1981b: 29-41). (See Figure 1.2) The ideational metafunction is associated with **particulate structuring** originally interpreted by Halliday as part-whole (as in constituency) structures for experiential meaning or part-part segmentation (as in clause dependency) for logical meaning. The interpersonal metafunction is characterised by **prosodic structuring** allowing for the ‘spreading’ of interpersonal meaning across ideational structural elements. For example, polarity can be realised throughout a clause,
as in underlined items in *No, I don’t want* any, ever or modality in *Perhaps, it could happen like that in some cases... just possibly, sometimes*. Finally, textual meaning is associated with **periodic structuring**, which organises meanings according to their significance in wave-like structures with peaks of prominence across ideational and interpersonal elements. For example, the two informational peaks of a clause are Theme realised by initial position and News by final position, as in *The story* (Theme) *is about many things, but mostly about greed* (New).

![Figure 1.2 Types of meaning and types of structure (Martin, 1996: 45)](image)

This thesis follows Martin’s (1995a and 1996) reinterpretation and elaboration of the metafunctionally diverse structuring principles identified by Halliday, summarised in Figure 1.2. In Martin’s description of particulate structuring, experiential meaning is arranged orbitally (one part of the whole being an obligatory nucleus and other parts optional satellites) and logical meaning is arranged serially (by interdependent chaining of a potentially infinite series of wholes). Martin has also demonstrated that the different structuring principles identified in the clause also operate at the level of text and across semiotic systems and more importantly, interact with each other. (See his analysis of a rock song in Martin, 1996: 37-44).
Chapter 1: Introduction and Theoretical Foundations

One question addressed in this thesis relates to the extent to which the structuring principles described above may prove applicable to other semiotics, in particular to the visual and also to intersemiosis in verbal-visual text. Though further structuring principles clearly need to be proposed for other semiotics, for example, the visual or spatial ‘array’ perhaps in the visual semiotic, the metafunctionally diverse structuring principles shown to operate at the levels of clause and text have not yet been systematically applied to visual and verbal visual text (Martin, 1995a, 1996).

Another important feature of language described by SFL is its organisation in terms of different levels or *strata*. Below the two most abstract contextual levels of genre and register (already discussed) there are three strata for language – one stratum of ‘expression’ - either *phonology* or *graphology* – and two strata of ‘content’ - the *lexicogrammar* (with lexis seen as a more delicate level of grammar) and the higher level of *discourse semantics*. Discourse semantics maps meaning systems available at the level of text (the basic unit of analysis for discourse) while the lexicogrammar maps meaning as systems at the level of clause (the basic unit of analysis for grammar). In the SFL literature theoretical descriptions have been provided for each of these levels (see
Chapter 1: Introduction and Theoretical Foundations


Table 1.3 presents the basic elements and systems in each of these strata according to metafunction. (The systems most relevant to this thesis enquiry, the interpersonal discourse semantic systems, are elaborated below in Section 1.2.2). The discourse-semantic systems in Table 1.3 and as used in this thesis are as modelled by Martin (1992) and elaborated by Martin and Rose (2003). Though the systems are presented here according to the metafunction they are most aligned with, Martin (1992: 26) emphasises that because discourse semantic systems operate at the level of text (the highest ranking unit of discourse), they are all concerned with ‘texturing’ meaning. This principle is relevant to Martin’s distribution of cohesion systems (asterisked in Table 1.3), which in the earlier SFL model (Halliday and Hasan, 1976) are included under the textual metafunction only.

Table 1.3 Strata, metafunction and systems (summarised from Martin, 1992; Martin and Rose, 2003)

<table>
<thead>
<tr>
<th>Metafunction</th>
<th>Ideational</th>
<th>Interpersonal</th>
<th>Textual</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Stratum</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Discourse Semantics</td>
<td>ideation - patterns of lexis in transitivity choices - relations of taxonomy &amp; synonymy in lexical strings*</td>
<td>negotiation exchange structure patterns across text *</td>
<td>identification - introducing &amp; tracking participants in reference chains *</td>
</tr>
<tr>
<td></td>
<td>external conjunction - logical connection between events (addition comparison, sequence, cause)</td>
<td>appraisal - types of attitude values, levels of intensification, types of intertextual positioning</td>
<td><strong>internal conjunction</strong> - logical connection between parts of discourse *</td>
</tr>
<tr>
<td></td>
<td><strong>involvement</strong> - solidarity vs exclusion markers</td>
<td><strong>periodicity</strong> - information flow organised across text either by scaffolding of smaller units into a larger or by chaining or serial expansion of units</td>
<td></td>
</tr>
<tr>
<td>Lexico-grammar</td>
<td>Transitivity - Process, Participant Circumstance [nominal group classification description, enumeration]</td>
<td>Mood, Modality, Polarity, Comment, Vocation; [nominal group attitude, person]</td>
<td>Theme and Information; tense and deixis; ellipsis and substitution</td>
</tr>
<tr>
<td>Graphology/Phonology</td>
<td>tone sequence sentence</td>
<td>tone, voice quality, formatting, colour emoticons</td>
<td>punctuation, layout tonality, tonicity</td>
</tr>
</tbody>
</table>
Chapter 1: Introduction and Theoretical Foundations

The fact that language is organised on both discourse-semantic and lexicogrammatical strata allows for the possibility of additional meanings created by tension between the two levels. For example, the discourse semantic option of ‘demanding goods and services’ (i.e. a command) is congruently realised by the grammatical option of imperative, as in *open the window*. However it can also be realised by other grammatical moods, such as declarative, *it’s stuffy in here* or modalised interrogative, *could you open the window?* Such forms create ‘stratal tension’ (Martin 1997: 33) because they function as commands but are ‘dressed up’ as if informative or questioning remarks. Halliday and Mathiesson (2004: 592) refer to such non-congruent grammatical realisations of a semantic option as ‘grammatical metaphors’, whether they are within the interpersonal metafunction, as in the above examples or within the ideational metafunction. An example of the latter would be when an action, normally construed as a clause configuration of participant and process (e.g. *the audience clapped*) is dressed up as a ‘thing’ when realised as a noun group (*the applause*). This kind of ideational grammatical metaphor is referred to as nominalisation and its relevance for analysis of the visual semiotic is an area still to be explored.

1.2.2 Interpersonal meaning

Interpersonal meanings cannot easily be expressed as configurations of discrete elements... The essence of the meaning potential of this part of the semantic system is that most of the options are associated with the act of meaning as a whole... this interpersonal meaning ...is strung throughout the clause as a continuous motif or colouring... the effect is cumulative... we shall refer to this type of realisation as 'prosodic', since the meaning is distributed like a prosody throughout a continuous stretch of discourse. (Halliday, 1979: 66-67)

Because this thesis is concerned primarily with evaluative meaning at the level of discourse, it is worth elaborating on Martin’s account of the most relevant discourse semantic systems, the three systems of interpersonal meaning.

**Negotiation** is concerned with conversational exchange structure across a text and tracks this in the patterning and sequencing of speech function choices (such as Command or Statement), via Mood resources, such as the choice of imperative or
declarative (most fully presented in Martin, 1992: 31-91). As it is associated with interactional meanings and spoken discourse, this semantic area is not seen as of great relevance to texts examined in this thesis.

**Involvement** is concerned with evaluative (not interactional) meanings that establish relations of solidarity or otherwise between writer and reader by means of non-gradable lexical resources such as technicality, slang and vocatives (Martin and White, 2005: 33-35). Involvement is relevant to data in this thesis and may be referred to, but as this semantic area has so far been sparsely described (one example is in Halliday, 1976b), it is not used to frame the thesis exploration.

**Appraisal** is also concerned with evaluative meanings that establish solidarity between writer and reader. Within Appraisal this is done by means of gradable (rather than non-gradable) lexical resources across a text. As developed by Martin (2000a) and Martin and White (2005), Appraisal extends SFL descriptions (e.g. Halliday, 1985a) but also incorporates other related theoretical perspectives – on intertextuality (Bahktin, 1986; Fairclough, 1992c; Lemke, 1995a, b.) and on evaluation (Biber and Finegan, 1988; Chafe and Nichols, 1986; Lemke, 1998c; Hunston and Thompson, 2000; Lyons, 1977).

### Table 1.4 Interpersonal resources (across strata) (adapted from Martin, 1997b: 21 & Martin and White, 2005: 35)

<table>
<thead>
<tr>
<th>genre</th>
<th>register (tenor)</th>
<th>discourse semantics</th>
<th>lexicogrammar</th>
<th>phononology graphology</th>
</tr>
</thead>
<tbody>
<tr>
<td>prosodic structuring</td>
<td>power (status)</td>
<td>NEGOTIATION</td>
<td>- mood - tagging - polarity</td>
<td>- tone (and ‘key’)</td>
</tr>
<tr>
<td></td>
<td>solidarity (contact)</td>
<td>APPRAISAL</td>
<td>- reporting/quoting - modal verbs, - modal adjuncts - evaluative lexis - intensifiers</td>
<td>- loudness. - pitch movement - voice quality - phoneaesthesia - [formatting] emoticon, colour</td>
</tr>
<tr>
<td></td>
<td></td>
<td>INVOLVEMENT</td>
<td>- vocatives/ names - technical lexis - specialised lexis - slang &amp; taboo lexis</td>
<td>- ‘accent’ - whisper - acronyms - pig latin - secret scripts</td>
</tr>
</tbody>
</table>
The overall set of interpersonal systems of which appraisal forms a part is presented across contextual and linguistic strata in Table 1.4 above. As shown there, appraisal brings together in its three subsystems diverse linguistic resources from each level (lexis, grammar and discourse) that construct and contribute to evaluative meaning – resources that are often described separately in other linguistic approaches (e.g. Hunston and Thompson, 2000). The three sub-systems in Appraisal, presented in the system network in Figure 1.4 below, are Engagement, to do with intertextual positioning; Attitude, kinds of attitudinal values; and Graduation, adjustment of intensity.

**Figure 1.5 The Appraisal System**

The **Engagement** system identifies ways in which attitude, values or pieces of information are sourced to the writer or to external sources in order to contribute to evaluation. Where no voice or alternative position is acknowledged [monogloss], information is asserted as ‘fact’. Where external voices are introduced [heterogloss], this may be done more explicitly by using the grammatical system of projection, as in quotes or reported speech, or more implicitly through acknowledgement of alternative positions via modal expressions like *perhaps*, which signal the subjectivity of the writer’s position.

The **Attitude** system identifies and describes three kinds of values, usually expressed lexically and in different parts of the grammar. The most basic kind is **affect**, to do with
people’s emotions or feelings, as in afraid/fear or happy/laugh. The two other kinds of values, judgement and appreciation, are described by Martin and White (2005: 45) as two ways in which feelings have been institutionalised in society. **Judgement** has to do with ethical assessments of people’s behaviour or character, as in kind or cruel and **appreciation** concerns valuation of things (material, abstract or semiotic) as in pleasing or chaotic.

The expression of attitude values is also described in terms of a cline, from explicit to implicit (Martin and White, 2005: 61-68). A value may be explicitly ‘inscribed’ in attitudinal lexis like kind and selfless, ‘provoked’ by a stable lexical metaphor such as in a real samaritan (widely associated with meanings of ‘caring’, ‘kind’ or ‘selfless’) or at the other extreme, an attitude may merely be ‘afforded’ by ideational meanings. An example of this would be he carried the old lady’s bags, which also allows the evaluative meaning of ‘kind’ or ‘caring’ to be inferred. Such ideational ‘tokens’ of attitude may be commonly associated with a certain attitude among one social group but not another. Attitude values may thus interact in complex ways across a text, not only in terms of where they are realised in the grammar and how they accumulate across sections but also in terms of how attitude inscriptions interact with evocations.

The **Graduation** system describes different ways of grading meanings – most obviously by adjusting the strength of attitude values, as for example by intensifiers raising the force of a value as in very, very scared or lowering it as in slightly chaotic. Non-attitudinal meanings can also be adjusted for force to create an evaluative effect, as in very, very big. And non-gradable ideational meanings like straight or red can be adjusted for focus to create an evaluative effect, by sharpening as in absolutely straight, or by softening in sort of red. A common means of evoking or ‘flagging’ attitude is by such graduation of ideational meaning.

Martin and White (2005: 211) have interpreted appraisal choices and their interaction in patterns across a text in terms of the communal alignments or affiliations negotiated with readers. People align around shared attitudes and values, and writers strategically imbue their text’s experiential content with different types of attitude (referred to as ‘bonding’ by Stenglin, 2004) to construct potential alignment with readers. Thus, there may be a clustering of affect values aligning readers around empathy (common in
entertainment texts), or a clustering of judgement values aligning readers in terms of principles or morals (as in news editorials), or, in an absence of affect and judgement, appreciation values may align readers in a more detached way around assessments of phenomena (as required in academic texts).

A number of ways in which evaluation may be prosodically structured in text has been identified in Martin and White (2005). The clustering of multiple attitude values, referred to as ‘saturation’ (pp.19-23), is one way. Another is ‘domination’, associated with the ‘propagation’ (Lemke, 1998c) or the spreading of highly salient values, and a third is ‘accumulation’ of values leading to their intensification. Propagation of evaluation can refer to the interaction between an attitude inscription and ideational meanings. For example, if an explicit attitude is in a prominent thematic position (domination) it can initiate a spread (propagation) of that attitude over an ensuing stretch of text with no overt attitudinal meanings. Accumulation of attitude may occur through repetition of lexis or even of grammatical structures across text in a way that creates or intensifies attitude and thus reader alignment. The co-occurrence of different meanings, referred to as ‘coupling’ (Martin, 2000), is also involved in prosodic structuring. For example, an appraisal value may be consistently coupled with an ideational meaning or a graduation value with an attitude value. Attending to all these prosodies enables a systematic description of the dynamic unfolding rhetoric of a text – aligning and realigning readers around different values. These ideas relating patterns of appraisal choices in a text to how readers are aligned around certain values are taken up in this thesis in relation to verbal and visual text.

Appraisal has also been used to explore rhetorical positioning in discourse in terms of the ‘voice’ of the writer and in terms of the writer’s ‘stance’. Martin and White (2005: Ch. 4) propose that particular configurations of evaluative meanings operating in specific settings can be linked with rhetorical objectives and construct authorial identities. These configurations, as found across many texts in a particular domain, are recognised as conventionalised ‘styles’ associated with a specific social role in a social context. Referring to these styles of evaluative positioning as ‘evaluative keys’ or ‘voices’, Martin and White (2005: 164-184) describe three recognisable journalistic roles in the news media context based on a study of texts in this domain (See Figure 1.5). ‘Writer Key’ is distinguished from ‘Reporter Key’ and then further distinguished
into ‘Commentator’ and ‘Correspondent Key’. The most constrained in terms of appraisal choices is Reporter Key, associated with ‘neutral’ or ‘objective’ hard news stories. The least constrained is Commentator Key, with the full range of appraisal choices available as in ‘subjective’ op-ed pieces. Corresponding evaluative keys have been proposed for different contexts – the writing of history (Coffin, 2000); narrative and literary response texts (Macken-Horarik, 2003); academic research texts (Hood, 2004) and, from work related to this thesis, in news photos (Economou, 2008). These ideas relating appraisal patterning in a text to a particular evaluative style will be taken up in the thesis in relation to both visual and verbal-visual news text.

<table>
<thead>
<tr>
<th>Journalistic Voices</th>
<th>No (extremely low probability of) unmediated inscribed judgement</th>
<th>No authorial affect</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reporter Voice</td>
<td>Lower probability (relative to writer voice) unmediated inscribed appreciation</td>
<td>+ observed affect (e.g. the angry nation)</td>
</tr>
<tr>
<td>Writer Voice</td>
<td>Inscribed authorial judgement</td>
<td></td>
</tr>
<tr>
<td>Correspondent Voice</td>
<td>No (low probability of) unmediated inscribed social sanction</td>
<td>No authorial affect</td>
</tr>
<tr>
<td>Commentator Voice</td>
<td>Higher probability (relative to reporter voice) unmediated inscribed appreciation</td>
<td>+ observed affect (e.g. the angry nation)</td>
</tr>
<tr>
<td></td>
<td>No co-textual constraints on judgement (free occurrence of unmediated social sanction and social esteem)</td>
<td>+ authorial affect</td>
</tr>
</tbody>
</table>

Figure 1.6 Journalistic keys – attitudinal profile (Martin and White, 2005: 178)

At a less general level, Martin and Rose (2005: 186-206) propose that a further, more localised co-patterning or ‘syndrome’ of evaluation choices might be taken up within a particular evaluative key and identify this as ‘evaluative stance’. They illustrate this by identifying three different evaluative stances taken within Commentator Key (damning, conciliatory and sceptical) in their analysis and comparison of three news commentaries on the same issue (2005: Ch 4). And they point to further research needed to identify what the range of available stances might be for each of the different media keys. These ideas will also be taken up in the thesis in relation to verbal and visual text.
Appraisal is used as an organising framework in this thesis in two ways. Firstly, it has been used to explore visual evaluative meaning in news photos, providing ways of systematically describing and identifying evoked evaluation, on which visual meaning relies heavily. Because news photos are predominantly ideational and ‘factual’, it is especially valuable to use appraisal to explore ways in which evaluation may be evoked by visual ideational meanings, and how these may interact with any explicit evaluation to accumulate or spread a value to verbal text nearby. Secondly, appraisal is used in this thesis to complement genre analysis by an analysis and comparison of both evaluative key and evaluative stance in a set of verbal-visual news texts from a similar news context in Chapter 7. The intention is to use the comprehensive range of appraisal tools to successfully deconstruct overall evaluative stance in these verbal-visual texts in a way that is analogous to the use of appraisal in discourse analysis studies of verbal text, reviewed in Chapter 2 (Martin, 2004a; Martin and White, 2005). To achieve this, the thesis will build on the initial attempts to apply appraisal to verbal-visual texts, also reviewed in Chapter 2 (Martin, 2001, 2004b; Macken-Horarick, 2003a, b, 2004).

1.2.3 Visual Semiosis

Language is one kind of semiosis. The architecture of SFL, and in particular its concept of the metafunctions, has provided a model that has been adopted for the analysis of other kinds of semiosis. The model has been applied to describe a range of non-linguistic (non-verbal) meaning systems such as music (van Leeuwen, 1999), action and gesture (Hood, 2007; Martinec, 1998, 2000, 2001, 2004) mathematical symbolism (O’Halloran, 2005, 2007a), architecture (O’Toole, 1994; Stenglin, 2004) and images (Kress and van Leeuwen, 2006; O’Toole, 1994). Most relevant here are the last two studies of the visual semiotic as they have both shown that a metafunctional framework can help systematically describe features of visual images. Kress and van Leeuwen’s (2006) work, now widely used for visual analysis (Fulton, 2005; papers in McCabe, O’Donnell and Whittaker, 2007 and Schirato and Webb, 2004; Stenglin and Iedema, 2001), is of most interest to this thesis as it uses system networks to describe visual meaning potential and also includes realisation statements, describing how visual meanings are realised by range of different visual forms. Although they do not apply the rank scale to describe image structure, as O’Toole does, their focus is broader.
O’Toole’s concern is with the description of examples of fine art, whereas Kress and van Leeuwen have explored a wider range of image types including news photos.

Kress and van Leeuwen (2006) incorporate ideas that derive from fine art description and theory but for the most part use the SFL metafunctional model with some changes in terminology. Their experiential metafunction is called ‘Representation’, their interpersonal metafunction comprises ‘Interaction’ and ‘Modality’ and their textual metafunction is called ‘Composition’. Many meaning resources identified in their systems parallel those in language. One of their Compositional systems, for example, positing a Given-New structure realised by Left-Right spatial organization of the image is modelled on the linguistic system of Information Structure similarly realised by left-right positioning in written text. Their Compositional system of Ideal and Real realised by Top-Bottom positioning, however, reveals the influence of fine art description.

Despite being open to criticism for being too tied to linguistic meaning, Kress and van Leeuwen’s systems could also be criticised for making insufficient use of SFL discourse semantic descriptions. Martin (2001) questions the validity of their focus on interactional meanings within the interpersonal, arguing that evaluative meaning is more relevant when considering the visual, and recommending the use of appraisal. In his study reviewed in Chapter 2, Martin (2001) demonstrates how a prominent image can function as an evaluative macro-Theme in a verbal-visual text by means of interaction between appraisal and periodicity. He shows how an image provoking attitude, which is made prominent in size and by initial position, can both thematise and spread the attitude to ensuing verbal text.

In acknowledgement of the wide currency of Kress and van Leeuwen’s (2006) descriptions, their categories will be incorporated in visual analysis in the thesis to some extent (with adjustments and extensions as noted in Chapter 3). The thesis will also apply, sometimes heuristically, SFL dimensions not used by Kress and van Leeuwen, including genre, stratification and rank, to explore visual and verbal-visual meaning. However, the primary focus of the thesis will be on using the metafunctional model as presented in Table 1.3 and, in particular, the appraisal system, as the way into visual and verbal-visual description. Most importantly for this thesis, these descriptions take into
account how interaction between systems across metafunctions and across text contribute to visual evaluative meaning and evaluative stance.

1.2.4 Intersemiosis

When it comes to describing the interaction between linguistic and non-linguistic meaning, referred to here as intersemiosis, researchers have attempted to find different ways to deal with the phenomenon. Most commonly, SFL resources for describing different ways of linking meanings within linguistic text have been used as a model for linking image with text. Although this work has not yet produced a comprehensive framework, progress is being made by researchers working with different kinds of texts – verbal-visual print texts (Baldry and Thibault, 2006; Bateman, 2008; Bateman, et al, 2004; Caple, 2008; Kress and van Leeuwen, 2006; Martin, 2001; Martinec, 2004, 2005; O’Halloran, 1999, 2005, 2007b; Royce, 1998; van Leeuwen, 2005; Ventola et al, 2004); electronic and film texts (Baldry and Thibault, 2006; Djonov, 2006; Iedema, 2001a; Lemke, 2002); and 3D sites (Pang, 2004; Ravelli, 2000; Stenglin, 2004). Many of these descriptions of intersemiosis have used SFL textual categories of lexical cohesion, information structure, conjunction, and logical relations. More recent work has suggested modelling intersemiosis on ideational categories of relational transitivity (Unsworth and Cleirigh, in press). Most recently, O’Halloran (2008) has brought together the various attempts to model intersemiosis and proposes six intersemiotic mechanisms.

These descriptions of intersemiosis provide useful models and some will be taken up in the analysis of verbal-visual texts in Chapter 7. One drawback of most of this work is that it does not foreground evaluative meaning, and few descriptions focus on the news media context. The most useful work in terms of this thesis are studies of individual multimodal or verbal-visual texts concerned with evaluative stance (e.g. Iedema et al, 1994; Macken, 2004; Martin, 2001, 2004b; Thibault, 2000), most of which are reviewed in more detail in Chapter 2. Though all these take a metafunctional approach, only two (Macken-Horarick, 2004 and Martin, 2001) attempt to incorporate evaluative meaning within their metafunctional analysis.
An approach that allows more attention to the interpersonal is Martin’s (2008) focus on instantiation in verbal-visual texts, offering a view of intersemiosis that is not confined to the textual or any one metafunction. He refers to the co-instantiation of meanings from different semiotic modes in the one text. He adds to this the notion of ‘commitment’ of meaning (2008c) to describe how meanings from different semiotic modalities are distributed in text. He points out that different kinds and ‘amounts’ of meanings may be instantiated by each of the semiotic systems in play in the text (and it may also be the case that particular semiotic systems afford particular kinds of meaning more readily). As an example of how two semiotic modes may ‘commit’ kinds of meanings differently, one can imagine a book cover where colours, decorative borders and images commit more evaluative meaning and less ideational meaning than accompanying titles and author information.

However, even Martin’s notions of co-instantiation and commitment do not entirely capture the potential for the unique ‘multiplication’ of meaning (Lemke, 1998b) made possible by intersemiosis. As another step towards this, the exploration of verbal-visual evaluative stance in Chapter 7 of this thesis will use a range of different tools to describe intersemiosis from the perspective of evaluation. This thesis, in particular, sees the appraisal system and the ways it has been used in analysis of verbal text as offering a useful way of bringing all dimensions of the SFL model to bear on describing evaluative meaning in verbal-visual text. It allows for further consideration of the extent to which the visual affords appraisal and how this is realised (explored in Chapters 3-6), and thus, a better reflection on the affordances of (and potential for different levels of commitment in) the two semiotic systems in relation to evaluation. It also allows for comparison of the degrees to which evaluative meaning is committed by each semiotic in any text, and for observation of any multiplicative effects of patterned couplings, both within one semiotic modality and across two semiotic modalities (see Chapter 7).

1.3 Organisation of the thesis

This thesis is organised in eight chapters.
Chapter 1: Introduction and Theoretical Foundations

This chapter has introduced the thesis project and presents its theoretical foundations in SFL, including a review of SFL literature related to appraisal and visual meaning.

Chapter 2. Literature Review and Research Design

Chapter 2 surveys relevant literature, first, in 2.1 from the overlapping fields of journalism practice, media/communication studies and journalism studies, and then, in 2.2, from social semiotics critical discourse analysis (CDA), particularly SFL CDA. In 2.3, all phases of the thesis project are described, including an initial pilot study that suggested the directions the thesis ultimately followed.

Chapter 3. Introduction to visual meaning

This chapter establishes a basic analytical framework to be applied in the thesis to the news photo as a semantic unit. It does so by adapting and interrogating available SFL, and SFL-based visual descriptions in the literature, particularly Kress and van Leeuwen’s widely used visual ‘grammar’ (2006).

Chapter 4. Visual appraisal: attitude in news photo

Chapter 5. Visual appraisal: graduation in the news photo

Chapter 6. Visual appraisal: engagement and interdiscursivity in the news photo

In these three chapters, each of the subsystems in the present appraisal system – attitude, graduation and engagement – is used to model and investigate evaluative meaning in the news photo. In the process, the present appraisal system is interrogated and an additional category based on ‘interdiscursivity’ (following Fairclough, 1992c), is proposed for the engagement system. Ultimately, the three chapters provide a visual appraisal system – a new set of tools for analysing three types of visual evaluation in the news photo.

Chapter 7. Verbal-visual evaluative stance in news review standouts

In this chapter, both verbal and visual appraisal tools are systematically applied in a comparative critical discourse analysis study of eight verbal-visual news texts (standouts) from two cultural contexts (Australian and Greek). This demonstrates the potential of the new tools developed for analysis of news photos and their productive use alongside present appraisal tools for analysis of evaluative stance in verbal-visual news discourse.
Chapter 8. Conclusion

In this chapter, the results from all analyses are summarised. Implications for theory, methodology and practice are outlined, as are the limitations of the study and directions for further research in this area.