

Edited by
Shoshana Dreyfus, Susan Hood and
Maree Stenglin

Semiotic Margins

Meaning in Multimodalities

Semiotic Margins

Also available from Continuum

Mathematical Discourse

Kay O'Halloran

Multimodal Semiotics

Edited by Len Unsworth

Semiotic Landscapes

Edited by Crispin Thurlow and Adam Jaworski

Semiotic Margins

Meaning in Multimodalities

Edited by
Shoshana Dreyfus,
Susan Hood
and
Maree Stenglin



Continuum International Publishing Group

The Tower Building
11 York Road
London SE1 7NX

80 Maiden Lane
Suite 704
New York, NY 10038

www.continuumbooks.com

© Shoshana Dreyfus, Susan Hood, Maree Stenglin and contributors 2011

All rights reserved. No part of this publication may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying, recording, or any information storage or retrieval system, without prior permission in writing from the publishers.

British Library Cataloguing-in-Publication Data

A catalogue record for this book is available from the British Library.

ISBN: 978-1-4411-7322-5 (hardcover)

Library of Congress Cataloging-in-Publication Data

A catalog record for this book is available from the Library of Congress.

Typeset by Newgen Imaging Systems Pvt Ltd, Chennai, India
Printed and bound in Great Britain

Contents

Introduction	1
<i>Shoshana Dreyfus, Susan Hood and Maree Stenglin</i>	
Part One: Beyond Paralinguistics	
Chapter 1 The Interpersonal Semiotics of Having a Laugh	7
<i>Naomi K. Knight</i>	
Chapter 2 Body Language in Face-to-face Teaching: A Focus on Textual and Interpersonal Meaning	31
<i>Susan Hood</i>	
Chapter 3 Grappling with a Non-speech Language: Describing and Theorizing the Non-verbal Multimodal Communication of a Child with an Intellectual Disability	53
<i>Shoshana Dreyfus</i>	
Part Two: Evolving Accounts of Space and Music	
Chapter 4 Spaced Out: An Evolving Cartography of a Visceral Semiotic	73
<i>Maree Stenglin</i>	
Chapter 5 Dealing with Musical Meaning: Towards an Embodied Model of Music	101
<i>Edward McDonald</i>	
Part Three: Intermodality between the Visual, Verbal and Aural	
Chapter 6 Organizing Visual Meaning: FRAMING and BALANCE in Picture-Book Images	125
<i>Clare Painter, J.R. Martin and Len Unsworth</i>	
Chapter 7 Integrating Visual and Verbal Meaning in Multimodal Text Comprehension: Towards a Model of Intermodal Relations	144
<i>Eveline Chan</i>	

Chapter 8	Rhythm and Multimodal Semiosis <i>Theo van Leeuwen</i>	168
Chapter 9	Meaning beyond the Margins: Learning to Interact with Books <i>David Rose</i>	177
Part Four: Imaging Representations of Meaning		
Chapter 10	Visualizing Logogenesis: Preserving the Dynamics of Meaning <i>Michele Zappavigna</i>	211
Chapter 11	Visualizing Multimodal Patterning <i>David Caldwell and Michele Zappavigna</i>	229
Chapter 12	Multimodal Semiotics: Theoretical Challenges <i>J.R. Martin</i>	243
	<i>Index</i>	271

Introduction

Shoshana Dreyfus
University of Sydney

Susan Hood
University of Technology, Sydney

Maree Stenglin
University of Sydney

The initial inspiration for this book was a conference held at the University of Sydney in December 2007. The conference, entitled 'Semiotic Margins: Reclaiming Meaning', brought together scholars interested in multimodal discourse studies from a social-semiotic perspective. As the terms 'semiotic' and 'margins' suggest, it was motivated by a strong desire to explore meaning-making resources other than language, especially those modes that are often considered to lie on the borders, fringes and peripheries of semiosis, and which have tended to receive less attention from the field of semiotics. Unifying the contributions to that conference were connections to social semiotics within the Systemic Functional tradition, acknowledging a theoretical heritage in the work of Michael Halliday and his colleagues in language as a social-semiotic system. Even though Halliday (e.g. as in Halliday & Hasan 1985) has long emphasized that language is only one semiotic system among many, the work on language as a resource for meaning-making has to date dominated the semiotic landscape.

This volume of current work in the field reflects similar interests and concerns with pushing at the margins of our understandings of semiosis. The contributions present analyses of the meaning-making potential of a wide range of modalities including body language, colour and ambience, laughter, architectural spaces, music, diagramming and image-verbiage relations. Contributions also engage with a second interpretation of the title Semiotic Margins, one to do with the relationship of other modalities to language, the question of what mode is marginal to what, and the ways in which different modes co-articulate, or co-pattern to create meaning.

The contributions to the volume have been subdivided into 4 key themes:

1. Beyond paralinguistics;
2. Evolving accounts of space and music;
3. Intermodality between the visual, verbal and aural;
4. Imaging representations of meaning.

We introduce the contributions within each theme briefly here.

Beyond Paralinguistics

The first theme explores modalities of meaning that have been categorized as paralinguistic, here including body language and laughter.

In Chapter 1, Naomi Knight explores the meaning-making potential of laughter, articulating its semiotic functions. Of particular interest, are the ways **laughter construes affiliation and bonding**. The data in her study are drawn from conversational humour.

In Chapter 2, Susan Hood investigates the ways in which teachers use body language in face-to-face teaching in tertiary classrooms, with a particular focus on the ways in which body language functions to **give salience to particular information** and to manage interaction and engagement. From analyses of the data, Hood begins to **construct system networks of choices in interpersonal and textual meaning**.

In Chapter 3, Shoshana Dreyfus analyses the non-verbal communication of a boy with a **severe intellectual disability** and shows how this requires that a range of modes of semiotic behaviour be analysed. In particular, behaviours that have traditionally been regarded as 'paralinguistic' become central rather than peripheral to communication.

Evolving Accounts of Space and Music

The second theme covers theoretical cartographies of modalities for which **systematic theorization is still evolving**, in particular those of space and **music**.

In Chapter 4, Maree Stenglin highlights some of the theoretical challenges involved in developing a metafunctionally diversified theory of three-dimensional space. Stenglin's examples are drawn from domestic architecture and she explores the genesis of Australian housing using ideational, interpersonal **and** textual lenses.

In Chapter 5, Edward McDonald is concerned with discourses about music from a social-semiotic perspective. He critiques several key texts exemplifying these discourses, concluding that embodiment is the crucial element in any theorization of music.

Intermodality: Visual, Verbal and Aural

The third theme is concerned with theorizing the co-articulation of visual and verbal meaning in children's picture books, and the implications of that for student literacy.

In Chapter 6, Clare Painter, J.R. Martin and Len Unsworth present two resources for analysing aspects of visual semiosis that fall within the **textual metafunction. They are 'framing' and 'balance'**. The chapter illustrates these systems using images from well-known picture books.

In Chapter 7, Eveline Chan presents findings from a research project investigating how visual and verbal meanings co-pattern in reading comprehension tests. Chan introduces a new model that begins with the interplay between representation and composition, and applies it to school literacy test materials. The study shows that successful comprehension requires the recovery of meanings across semiotic modes.

In Chapter 8, Theo van Leeuwen addresses the question of intermodal relationships, challenging the notion that we can communicate at all in a single modality. Central to his explanation of relationships across modalities is the element of 'rhythm', which is seen as an essential framework for coordinating and aligning different modalities in meaning making.

In Chapter 9, David Rose explores how children learn to engage with books as a mode of communication, and how this engagement is pivotal to how they learn from reading in school. Rose's account of student literacy includes a methodology for supporting those children relegated to the socioeconomic margins of meaning making, and facilitating their apprenticeship into mainstream literacy practices.

Imaging Representations of Meaning

The final theme engages with new approaches to transcribing and mapping representations of multimodal semiosis in screen-based technologies.

In Chapter 10, Michele Zappavigna presents three new tools for visualizing text patterns: text arcs, stream graphs and animated networks. Zappavigna demonstrates that the advantages of these tools are two-fold: they visualize linguistic patterns that the eye is unable to identify and provide a synoptic perspective on text without sacrificing logogenesis.

In the Chapter 11, David Caldwell and Michele Zappavigna explore how one visualization method, the arc diagram, is able to represent the way meanings build up as a text unfolds. The text chosen for analysis is a rap song by Kanye West and two of his collaborators and their analysis focuses on one system of meaning: graduation, a system from appraisal theory (Martin & White 2005). The key focus of the analysis is how graduation is materialized in the end-rhymes of a rap song: a feature that distinguishes the rhyming capacity of a rap artist.

Theory and Challenges

Chapter 12, by J.R. Martin, concludes this volume with an overview of what it means to say that language is a semiotic system, and a summary account of what the dimensions of that semiosis are. The discussion generates a set of questions for the multimodal analyst in terms of how they theorize the modalities they are exploring. Martin then addresses a number of significant challenges in multimodal research, some of which are the focus of ongoing work, some of which are newly emerging and some of which we have barely begun to consider.

Conclusion

The contributions in this book are stunning in terms of both their depth and breath. They are also at the **cutting edges** of multimodal discourse analysis in three main ways. First, they extend the field of semiotic analysis; second, they expand the theory; and finally, they contribute to the ongoing dialogue between linguistics and other disciplines, including those of psychology, architecture, music, education, language **disorder**, advertising and information technology.

References

- Halliday, M.A.K. & Hasan, R. (1985). *Language, context and text: Aspects of language in a social-semiotic perspective*. Geelong: Deakin University Press.
- Martin, J.R.M. & White, P.R.R. (2005). *The language of evaluation: Appraisal in English*. Hampshire/NY: Palgrave Macmillan.

Part One

Beyond Paralinguistics

This page intentionally left blank

Chapter 1

The Interpersonal Semiotics of Having a Laugh

Naomi K. Knight
University of Sydney

How much lies in Laughter: the cipher-key, wherewith we decipher the whole man.
(Thomas Carlyle)

Introduction

Studies of laughter have a long history in the literature, with references as far back as Plato (Glenn 2003). Some studies have focused on laughter as a discrete phenomenon with particular qualities (Sroufe & Wunsch 1972, Provine 1992, 2000, 2004, Bachorowski & Owren 2001). However, the apparent pervasiveness of laughter, particularly in conversational speech has also given rise to a growing body of research into its relation to speech (Jefferson 1979, 1984, Norrick 1993, Stewart 1995, Bonaiuto et al. 2003) and particularly to humour in this context (Koestler 1964, Morreall 1983, Zijderfeld 1983, Apte 1985, Martin 2001, Archakis & Tsakona 2005, Warren et al. 2006). Studies in conversation have found that laughter has a specific and highly functional meaning potential with significant implications for social interaction (Provine 2000, Glenn 2003, Partington 2006). In this chapter, we explore the social potential of laughter in conversation from a grounding in Systemic Functional Linguistics (SFL) as a social semiotic theory of language (Halliday 1978a). In particular, the analyses apply discourse semantic systems related to interpersonal meaning, namely Appraisal (Martin & White 2005) and Negotiation (Martin 1992, Eggins & Slade 1997). The aim is to begin to **develop, through such tools as system networks, a semiotic account, of how we can mean interpersonally in laughter in conversation.** The data set from which examples are drawn for this chapter comprises conversations among groups of Canadian friends (drawn from Knight, in preparation).

Literature on Laughter

We begin with a review of just some of the broad scope of literature that provides a backdrop to the current study of laughter and interpersonal meaning, a scope that ranges from an interest in the phylogenetic origins of laughter in the body language of apes, to the ontogenetic development of laughter in babies and its importance in the development of mother tongue, and to the growing body of studies from a number of theoretical orientations that concern the role of laughter in human conversation.

A Phylogenetic – Ontogenetic Take on Laughter

From a physiological perspective, laughter is first a sounding phenomenon, expressed with the same organs as speech including the vocal apparatus. As air passes through, it is pulsed out through the vibration of the vocal folds when voiced (Owren & Bachorowski 2003:189). Provine (2000:57–62) describes the ‘acoustic signature’ of laughter as a series of short, vowel-like syllables that he calls *laugh-notes* (e.g. *ha, ho, he*) each with a duration of about 75 milliseconds, spaced at regular intervals of about 210 milliseconds. Laughing is thus periodic and includes a recognizable sound character. The initial burst of air pressure that characterizes laughter is often the point of departure for laughter theorists, especially those who attend to it as a physiological or psychological reaction to stimuli (e.g. Spencer 1911, Freud 1976 [1905], Morreall 1983, see also Raskin 1985, Chafe 2007). However, a look to our ape ancestry suggests that laughter is not only a sounding phenomenon. It is also connected with ‘the adjustment of facial muscles that we call smiling’ (Chafe 2001:38, Darwin 1965 [1872], Koestler 1964, Ruch & Ekman 2001, Owren & Bachorowski 2003), and laughter and smiling can often be mutually interpretable. As physical expressions and in the evolution of their functions laughing and smiling are apparently intertwined.

What of the phylogenetic origins of this phenomenon of laughter? van Hooff (1972, see also 1967) suggests that a combination of two types of ape display converged in humans. He distinguishes two proximal facial expressions of apes that combine in humans: a ‘relaxed open-mouth display’ (accompanied by quick breathing with vocalizations in chimps), which is related to human laughter, and a ‘silent bared-teeth display’, which relates to the human smile (see Image 1.1). These combine an appeasement function with a more modern friendly or affiliative function (van Hooff 1972:225), and human laughter and smiling are placed on an evolutionary track showing that they have also come together through time to shade into each other.

Similarly, Morris (1967) suggests that the ‘play-face’ and soft play-grunt in chimpanzees, stimulated by a mixture of fright and safety or friendliness,

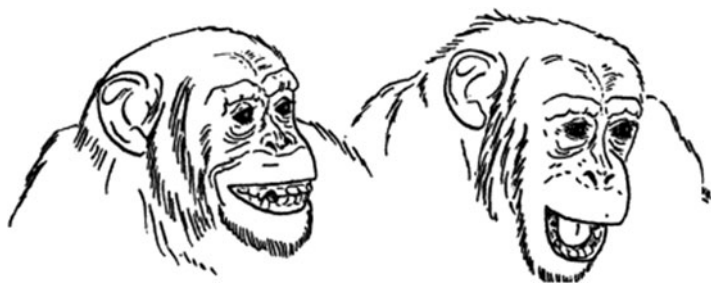


IMAGE 1.1 A silent or 'horizontal' bared-teeth display and a relaxed open-mouth display in two chimpanzees (*pan troglodytes*) (van Hooff 1972:219)

combine two otherwise distinct expressions to signal appeasement or 'non-threat', evolving into the human smile (see Morris 1967:137–139, 141–144).

The phylogenetic evolution of laughter (from apes to human) can be complemented with notions of ontogenesis (from protolanguage to language). From the perspective of ontogenesis in human infants, Morris (1967) links the responses to fright and friendliness in chimpanzees to the combined gurgle and cry that results when infants are presented with 'shock stimuli' (e.g. tickling) by a parent they recognize as a 'safe protector' (Morris 1967:103). This combination of cry and gurgle is identified as the ontogenesis of laughter. The expression occurs when bonding between the mother (or other caregivers) and child allows the child to identify the trusted protector. The functions of laughter as designating safety or 'non-threat' (Partington 2006), friendliness and play in reaction to non-threatening activities and stimuli signal the interpersonal potential of laughter.

In the earliest phase of ontogenesis, laughter is part of multimodal protolanguage and is treated by caregivers as expressing meaning along with other protolinguistic sounds and features (Halliday 1975, Painter 1984, 2003, Matthiessen 2006). Laughter is a mode of expression with microfunctional meaning potential, but it seems to express a particular orientation to meaning that is a precursor to the interpersonal orientation of the developed mother tongue. According to Matthiessen (2006:5) different modes of expression 'may be brought together within one microfunctional meaning potential or dispersed across more than one such potential'. At the same time there is 'a strong tendency for a particular mode of expression to go with a particular mode of meaning'. Laughter as a mode of expression in protolanguage can be seen to express meaning particularly within the *reflective* mode of consciousness. That is, it makes meanings in the personal and interactional microfunctions that are components of the reflective sphere, but not in the regulatory and instrumental microfunctions of the active sphere (Halliday 1978b). This can be

observed both in the infant's laugh and in the origins of ape displays as discussed above. An infant's laugh is a signal to the trusted mother of 'non-threat' (removal of fear) and function interactionally to bring infant and caregiver together in bonding. It is a signal of togetherness with the meaning of 'you and me' (Halliday 1975).

Painter (2003) suggests that in the ontogenetic transition from protolanguage to language, laughter again plays an important role, and is in fact a driving force behind the move into the language proper. As the child transitions from protolanguage into the mother tongue language and incorporates the notion of metafunctions, the functionality of laughter develops as well. Following Matthiessen (2006), laughter becomes a semiotic system that fuses with language in the semantic stratum where 'different semiotic systems are integrated as complementary contributions to the making of meaning in context' (2006:2). In other words, different expressive resources are distributed across semiotic systems, including that of laughter, 'paralanguage' and 'body language'. Each has different semiotic affordances on the content plane, but each combines with language to achieve a unified 'performance' (2006:3) of meaning in context. While language is distinguished from other semiotic systems because it has a level of grammar, or a 'higher degree of systematisation of its meaning potential' (2006:7), laughter is a semiotic system that construes meanings of language by its own expression system. Laughter is represented as a semiotic system that coordinates with language to make meaning in Figure 1.1.

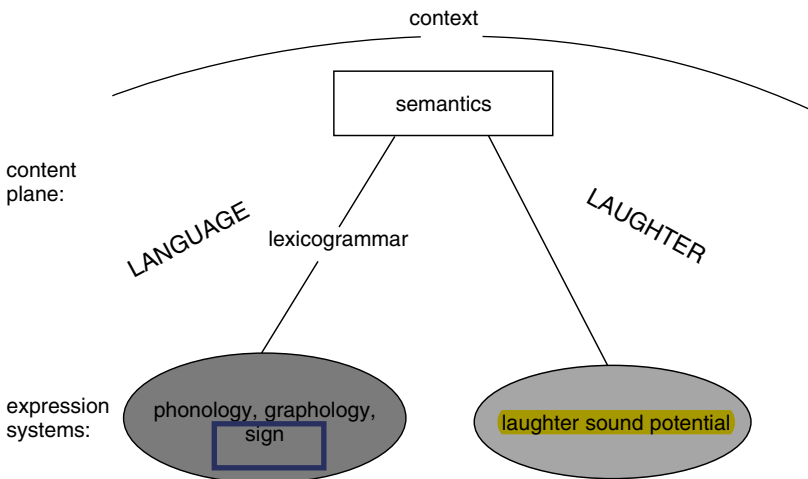


FIGURE 1.1 Laughter as a semiotic system alongside language, fusing in the semantic stratum but differing in expression systems

According to Matthiessen (2006:7), '[b]ody language and paralanguage emerge as sets of distinct semiotic systems when protolanguage is transformed into language', and laughter takes on expressive resources including some from both paralanguage (e.g. loudness) and body language (e.g. facial expression) to make meaning in the content plane as a semiotic system in its own right. This classification of laughter as both paralanguage and body language suggests an increased meaning potential for this social semiotic. This perspective is hence a divergence from traditional perspectives on laughter as 'paralanguage' in the sense of 'non-linguistic elements in conversation' (Abercrombie 1968:56) or as 'contextualization cues' (Gumperz 1982) that work to modify or qualify speech (Poyatos 1993). In these perspectives, laughter is seen to have the potential to communicate context (such as humorous play) but is thought not to have any semantic meaning potential of its own. From a social semiotic perspective laughter does not just provide contextual clues for the interpretation of speech. Rather it has semantic meaning potential, which coordinates with speech to create meaning in context. As a semiotic system in this sense, laughter is meaningful and not incidental to, or 'parasitic' on language (Halliday & Matthiessen 1999:606). Both 'complement one another in the creation of meaning' (Matthiessen 2006:1).

At the same time, the semiotic affordances of laughter are limited in relation to language in a number of ways: the expression systems of laughter make meaning in discourse semantics but not through a level of lexicogrammar; its discourse semantic meaning potential is limited to interpersonal systems, as it makes meaning only in the interpersonal metafunction (following from its origins in the reflective mode in protolanguage); and it is dependent on the linguistic co-text and the context to be interpreted. Due to these limitations, laughter may not realize meaning in the same way that language does, but it suggests interpersonal semantic meanings, which are interpreted only in relation to the context and the co-text.¹ Therefore, we propose that laughter indicates interpersonal meaning through choices in the discourse semantic systems.

Notwithstanding its limitations, laughter makes meanings in context through its own system of expression, and may even substitute for speech in conversation as an interactional move (discussed further below), indicating that it coordinates with language in varying degrees in discourse. The interpersonal potential of laughter includes the discourse semantic systems of *move* in Negotiation and *attitude* in Appraisal. Through the particular choices that laughers make in these systems in combination with speech, an instance of laughter also impacts upon the social relations between interactants, as the co-articulation of meanings construes their cultural relations as members of social networks with particular sets of values. Hence, in terms of semiosis, laughter can be shown to function as a semiotic system in its own right, making meaning in the discourse semantic systems within the interpersonal metafunction, and combining with speech in ways that affect affiliation.

why
sem
antic
and
not
prag
mati
c?

what
syst
agai
proc

what
sema

so it i
aspe

The Interpersonal Function of Laughter

Laughter is understood to be inherently social. Glenn (2003) notes that 'its occurrence, form and meaning are shaped deeply by the presence of others, roles, relationships, activities and other contextual features' (2003:32). People most often laugh together rather than alone (Provine 2004:215). Laughter has been explored in terms of its role in aligning interactants towards the talk in progress (Goodwin 1986), in expressing attitudes (Chafe 2007) and in influencing the behaviour of others (Owren & Bachorowski 2003), in constructing affiliation and disaffiliation (Ellis 1997, Glenn 2003) and creating intimacy (Jefferson et al. 1987) between interlocutors. Coates (2007) has described laughter as having multiple functions in playful conversational talk, including signalling the 'presence of a collaborative floor' (2007:44) and constructing solidarity.

Research has also focused on the association of laughter to humour. Sacks (1974), for instance, found that laughter or 'laughings' are the prime minimal response sequence in the telling of dirty jokes. He also revealed that a laugh could grade the joke and its teller (affiliating with a preceding utterance or not), exhibiting that evaluative meaning may be communicated through a laugh not only towards the verbiage but also towards the speaker. In naturalistic social interactions such as casual conversation, Provine (2004:215) shows laughter to be affected by the dynamics of the face-to-face mode and to be stimulated by the interaction between people rather than jokes per se. Nonetheless, its occurrence supplies an overall 'meta-message' of play (Bateson 1987) or funniness to an utterance. What is said may be thought of as 'funny' because of its relationship to the immediate micro-context (and co-text) and/or to the macro cultural context (Eggins & Slade 1997:157). In fact, laughter has even been shown to be necessary for the interpretation of a conversational text as humorous (Archakis & Tsakona 2005).

Analysing Laughter in Relation to Attitude and Affiliation

The inherent sociability of laughter orients us to the interpersonal meta-function as the most salient in our semiotic perspective. Building on studies such as those cited above, we explore laughter's potential to convey interpersonal meanings and to promote affiliation between persons in social life in the context of convivial conversational humour.²

Laughter and Attitude

From a social semiotic perspective we draw on a theorization of *attitude* within the system of Appraisal (Martin & White 2005). *Attitude* is modelled as having

three dimensions: *affect* as emotion and feelings, *appreciation* as the evaluation of phenomena or events, and *judgement* as the evaluation of people and behaviour. These attitudinal meanings may carry either a positive or negative polarity. Attitudinal meanings may be expressed explicitly or *inscribed*, or they may be implied or *invoked*. An attitudinal meaning may also be graded up or down with resources of *graduation*. A number of these options are evident in Examples 1 and 2. (Transcription conventions are modified from Eggins & Slade 1997.)

In Example 1, three female Canadian university students at first share positive attitudes about the holidays they have just spent away with family. They inscribe positive appreciation of the food they consumed in ‘ate *well*; ate *well*; good pie’. The prosody of positive appreciation of eating is then disrupted with U’s mention of diet. In this context ‘on a diet now’ might be interpreted as an instance of invoked negative appreciation of the impact of the eating.

- U = = Yeah I saw like my family and friends . . . I ate well (*laughs*)
 N We all ate well.
 (*all laugh*)
 N Dude we all (*laughing*) ate good pie!
 (*continuous laughing*)
 U Yes I agree. (*continuous laughing*) On a diet now.
 (*all laugh*)

Example 1 Conversational participants share attitudinal meanings

In Example 2, the three university students are discussing a previous event in which a close friend, Marissa³, has reacted unfavourably as a director in auditioning potential actors for her play.

- C And Yana somehow it’s like she only got in because she was
 N ()
 C When she was auditioning, = =
 N = = Oh yea::h () laughing
 C Marissa laughed so she felt bad (*laughs*) so she let her in
 N, C (*laugh*) = =
 F = = Oh:::!
 N (*laughs*) = =
 C = = It was like she just started laughing in the middle of our audition so she was like felt so bad (*N laughs*) and she was like ‘Let her in’ (*laughs*)

Example 2 Conversational participants share evaluative meanings of laughter

In Example 2, Marissa is reported as having negatively judged herself [negative judgement as propriety] for laughing at Yana in the audition. C reports this

with the use of inscribed attitude in 'she felt *bad*', and later as graded up in force in 'she felt *so bad*'. C also negatively evaluates Yana in terms of judgement as capacity in the opening move but this time the attitudinal meaning is invoked rather than inscribed in *Yana somehow . . . got in*.

In Examples 1 and 2 the speakers share laughter in the interaction (which we will attend to shortly). However, in Example 2, the speakers also discuss instances of the shared experience of laughter and its meaning potential, providing insight into the potential for laughter to convey attitudes on its own without the complement of speech. In telling her funny story to her friends, C is not suggesting that Marissa herself inscribed in speech her negative judgement of Yana's audition. Rather, C interprets the attitudinal meaning potential of Marissa's laughter as suggesting this negative judgement. Neither does C suggest that Marissa inscribed in speech a negative self-judgement. Rather this is intuited from the fact that having laughed at Yana's performance she subsequently gave her a part in the play to minimize the negative judgement she had seemingly conveyed towards Yana. In this conversation, the speakers' references to laughter in a previous interaction are seen to convey evaluative meanings.

In their re-telling, the participants also share in laughter, and this gives rise to another level of analysis in the relationship of laughter to attitude to do with the sharing or otherwise of attitudes in processes of affiliation.

Laughter and Affiliation

Affiliation (Knight 2008, forthcoming) is, in brief, a theory of how linguistic choices function to construct communities and our identities as members of those communities. It builds on Martin's (2000) concept of 'coupling', that is the association of values and entities or activities, and Stenglin's (2004) notion of 'bonding'.⁴ In Affiliation theory, a bond is a higher-order social semiotic unit by which we negotiate affiliation (see Knight, forthcoming, for further discussion). Affiliation refers to the ways in which people come together in a community around particular values 'invested' in activities (Martin & White 2005:211). In other words, as individuals, we value experience differently, and we present valued experiences through linguistic couplings of attitude and ideational meaning. As we talk in casual conversation, we present these couplings of attitude with ideational meaning (e.g. 'good + holidays'⁵ in Example 2) to share them as recognizable bound meanings in communities (e.g. the family community).

However, couplings are not only presented to commune within a single community, but may be differentially negotiated. Any presented coupling (e.g. *ate + well*) may create tension in terms of what we can share as members of another particular community. Such tensions must then be laughed off by the interlocutors in order to continue affiliating.⁶ Affiliation thus concerns the negotiation of various communities of values. Couplings in the linguistic

text are the points around which we negotiate our alignments in degrees of 'otherness' and 'in-ness' and laughter is a key resource in this process of negotiation. In this way, laughter serves as our *way in* to the study of affiliation, as it offers an explicit signal that this social process is going on.

In humour, laughter therefore does not only indicate attitude but attitudes coupled with experience that are laughable to the interactants. These couplings are only found laughable in relation to how the interactants have been affiliating together. In Example 3, the interactants laugh off the coupling of the ideational meaning of 'eating' with the positive appreciation *well* (in N's utterance *We all ate well*) because it creates a tension with underlying values around eating too much that they share together.

- U = = Yeah I saw like my family and friends . . . I ate well (*laughs*)
 N We all ate well.
 (*all laugh*)

Example 3 Laughing off the coupling ate + well

In Example 4, U adds *On a diet now*, indicating that this young female community expects a negative appreciation to be coupled with eating a lot. This contrast is then responded to as humorous (all laugh). Together they laugh off the notion of dieting as necessary. They indicate that they share values that do not take eating too much, or dieting, too seriously.

- U Yes I agree. (*continuous laughing*) On a diet now.
 (*all laugh*)

Example 4 Laughing off dieting as necessary

The laughter shows that attitudinal couplings are presented that create laughable tensions for the participants in the negotiations of affiliation. Laughter provides a window into how interactants negotiate their communal values, identities and alignments, indicating degrees of 'otherness' and 'in-ness' (Eggins & Slade 1997:155). In phases of humour in conversation, laughter is a reaction to (and marker of) value-infused meanings that need to be negotiated in processes of affiliation. It is through laughter that interactants manage tensions that may arise as they construe themselves as members of communities.

There is more to be considered here in terms of the particular expressive features of the laughter in the talk and how this relates to attitudinal meaning, but first we will look more closely at the placement of the laughter in terms of the move structure of the interaction.

The meanings of the laughter are dependent upon its placement with speech as a conversational move in the exchange. That is, the meanings are affected by the speech function the laughing constitutes. Laughter can mark humorous tension, such as in U's own laughter (in Example 1) following her

utterance *I ate well*, speakers can laugh off a tension they do following N's utterance: *We all ate well*. Move choices that are made in the articulation of the laugh are described in the following section.

Laughter as a Conversational Move

Laughter's role in turn-taking has been studied in conversation analysis, most prominently by Jefferson (cf. 1979, 1984, 1985, Jefferson et al. 1987), who places it in an adjacency pair with the previous turn at talk in an invitation-acceptance sequence.

In SFL, move choices are analysed not just as formal choices as in CA but as meaning-making options within the discourse semantic system of Negotiation (for a detailed explanation of the system of Negotiation see Eggins & Slade 1997 and Martin & Rose 2007). Such distinctions in move are important in an analysis of the meaning potential of laughter.

In terms of moves, participants can open an exchange or sustain an opening by continuing or reacting. Following Eggins and Slade's (1997) speech function network, laughter in conversation may constitute both opening and continuing moves in tandem with speech. For instance, in Example 1, U's laughter combines with speech to fulfil a continuing extension move, in:

SUSTAIN: CONTINUE: APPEND: EXTEND	U: Yeah I saw like my family and friends
SUSTAIN: CONTINUE: PROLONG: EXTEND	... I ate well (<u>laughs</u>)

Reacting moves are determined in regard to previous moves, and laughter may constitute a reacting move with or without accompanying speech. If alone, it is coded as an acknowledging move, as a response to a statement of fact. Such moves indicate a willingness to accept the speaker's proposition and are often realized by minimal expressions (see e.g. Eggins & Slade 1997:206–207). Examples are taken from Knight (in preparation).

SUSTAIN: CONTINUE: APPEND: EXTEND	SH: And so she waddles up to my snack (<u>laughing</u>)
	tray = =
REACT: SUPP: ANSWER: ACKNOWLEDGE	N: (<u>laughs</u>)

When in combination with speech, a laugh co-articulates the move that the speech is construing. In the following example, the reaction move functions to 'counter' the speaker's claim and is expressed in verbiage and laughter:

SUSTAIN: CONTINUE: PROLONG: EXTEND	T: They-they wanna just have fun an-an I don't know
------------------------------------	---

pick up girls that's the
idea of the thing. Well
that's how they () = =
REACT: CONF: CHALLENGE: COUNTER K = = °Dressed like a girl°
(laughs) = =

The meaning of the laughter therefore depends on the kind of move function it realizes in the interaction, whether it is part of an initiation or a reaction, and where it occurs in relation to verbiage.

Before we bring together the partial analyses presented so far to demonstrate how laughter makes meaning as a conversational move in relation to attitude and affiliation, there is one further aspect of the meaning potential of laughter that we need to attend to, that of the choices in expressive features. As a social semiotic, speakers vary the meaning indicated in their laughter by changing the characteristics of its expression. By considering laughter within a particular context, it is possible to represent these choices of sound features systematically and paradigmatically in a system network. The following section will present a system network that models the sound potential of laughter in convivial conversational humour, relating its systems of meaning with particular uses in the social context.

Laughter, Sound and Meaning

In their study of conversational laughter, Vettin and Todt (2004) found that acoustic features are systematically linked to conversational context. They differ between situations, convey information about and depend on the conversational role of the laugher, and are sensitive to communicative norms. It is proposed here that articulatory features can also be systematically linked to meaning differences in laughter in convivial conversational humour, and are sensitive to contextual constraints.

A System Network of Sound Potential for Laughter

Specific expression options for laughter indicating Appraisal and Negotiation and construing Affiliation in convivial conversational humour can be identified and systematized, and are presented in a system network in Figure 1.2. This system indicates the valeur of possible choices that may be combined into distinguishable laughter expressions. Meanings of laughter in convivial conversational humour depend on what speech role the laugher takes on, and on its relation with speech, so that combinations of sounding options in the laugh expression indicate discourse semantic systems within the frame of these considerations.

how
How
varia

what
role?

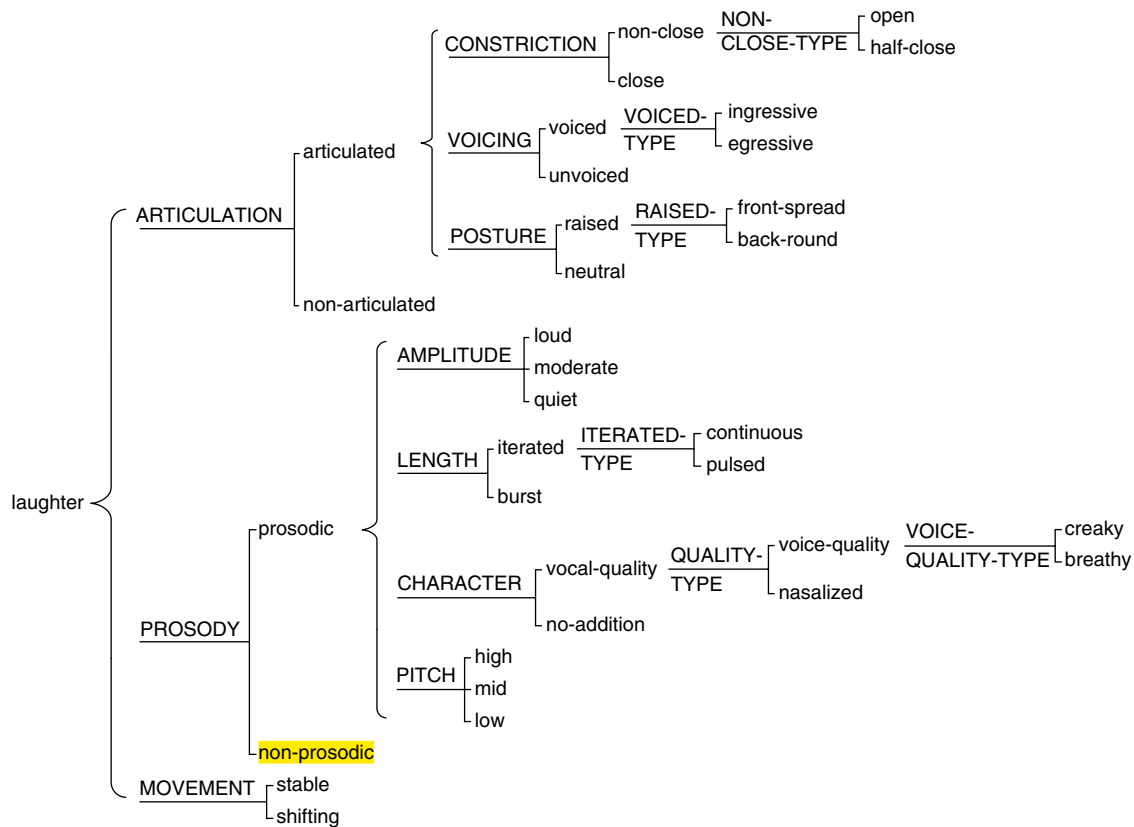


FIGURE 1.2 A system network of laughter sound potential in convivial conversational humour

As laughter indicates rather than realizes meaning, its expressions are dependent on the context and surrounding co-text in the interaction; and in making choices from the specified options, it is important to interpret these in their situational environment. That is to say that *who* is laughing (e.g. speaker or hearer(s)) must be considered in relation to the utterances in the text. Whether the verbal co-text specifically precedes or follows or is overlapped by the laugh conveys information about meanings produced as well. In combination with language the semiotic potential of the laugh is more fully revealed.

The paradigmatic options in sound for laughter are given with respect to its articulatory and its prosodic features,⁷ and these can change depending on its movement, providing three simultaneous subsystems in the system network. The 'stable' versus 'shifting' distinction in the subsystem of *movement* follows Halliday's (1992) classification for Chinese syllabic phonology,⁸ and captures the possibility of a laugh changing its course, from which the laugher re-enters the system; or if stable, only one entry into the system is necessary. This may impact upon the meaning made as the speaker may alter his or her attitude by changing to a different combination of sound features.

Choices in the *articulation* of the laugh are presented in a close-up version in Figure 1.3.

As a laugh is articulated, the closure of the mouth is captured in the system of *constriction*. While it is represented as distinct choices, it may be seen to follow Stewart (1995) in that the vocalic sound that combines with the initial /h/ aspiration (e.g. 'ha') can be measured on a continuum based on the opening of the mouth.⁹ A laugh can also be articulated as voiced or unvoiced in the system of *voicing*, and this combined with the closure of the mouth affects whether the laugh is more nasal or oral. Chafe (2007:28) also identifies ingressive voicing (recovery inhalations with enough laryngeal friction to make audible) as a feature of laughter that is not found in ordinary speech, and that has a highly distinctive sound. This has been incorporated into the

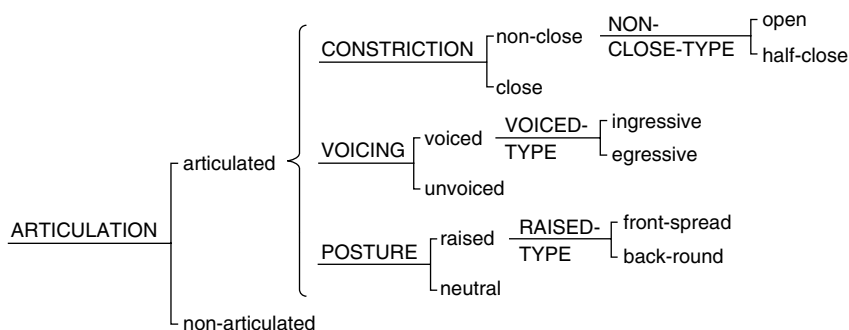


FIGURE 1.3 ARTICULATION system of laughter sound potential

why
refer
wha

network as an option in voicing – voiced between ‘ingressive’ and ‘egressive’ (with ‘ingressive voicing’ as the marked choice). The system of *posture* follows Halliday’s (1992) classification for Chinese syllabic phonology, to capture the placement of the tongue and lips as neutral (ə) or raised, and if raised, the tongue can be fronted and lips spread (y), or the tongue can be back with the lips rounded (w).

These choices in the articulation of laughter affect the meanings it makes in the discourse semantic system of attitude and the social context of affiliative meanings, and must be considered in analysis of convivial conversational humour. For instance, a difference between voiced and unvoiced laughter has been shown to make an impact on attitudinal meanings. Through acoustic and experimental analyses, scholars have found that there are differences between positive and negative emotional correlates with voiced and unvoiced laughter. Unvoiced laughs are perceived as related to ‘negative’ emotions and attitudes, while voiced laughs are more often perceived as related to ‘positive’ emotions (Devillers & Vidrascu 2007) and these cause similar emotional responses in listeners (Bachorowski & Owren 2001). This suggests that there is an impact on attitudinal polarity in relation to the presence of phonation in a laugh.

Articulation is complemented by simultaneous choices in the system of prosody, which accounts for the non-segmental features of a laugh (Figure 1.4).

Amplitude involves the intensity or loudness of the laughter utterance, and can be interpreted perceptually by its volume levels as either moderately intense or especially loud or quiet. Duration and repetition of the laughter is included under *length*, as continuous or pulsed laughter has a longer duration with distinct features from a laughter burst, which often involves a more forceful expulsion of air in the initial instigation. *Pitch* differences range from low to mid to high pitch, and additional sound quality can affect the *character* of the laugh through voice quality (creaky or breathy) (Clark & Yallop 1990:60–61) or nasalization. Voice quality has been shown to have its own meaning potential (van Leeuwen 1999), but as a combined variable in a laughter expression, it gives the laugh a particular ‘character’, indicating attitudinal and social interactional meanings associated with it as well.

The initial choices in the articulation network between ‘articulated’ and ‘non-articulated’ and in the prosody network between ‘prosodic’ and ‘non-prosodic’ allow for laughter that may be classified through a smile only or integrated into speech, as in ‘speech laughs’ (Nwokah et al. 1999, Trouvain 2001) or ‘laughspeak’ (Provine 2000). That is to say that laughs may occur on their own, in combination with separate speech, or they may punctuate the speech itself: when laughs occur within speech their articulation and prosodic features cannot be separated from the sound features of the verbiage. It is important that the co-textual environment of speech is considered in interpreting a laugh, especially since a laugh punctuating a speaker’s speech will orient towards the meaning conveyed within that speech.

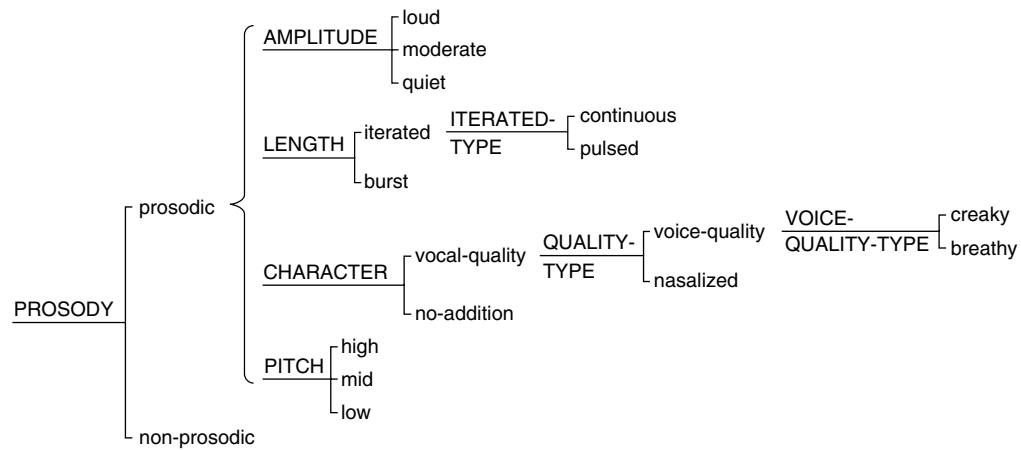


FIGURE 1.4 PROSODY system of laughter sound potential

Examples of Laughter in Convivial Conversational Humour

In Example 5, a combination of options in laughter demonstrates its use in conveying affiliation while contributing additional attitude in relation to speech. The two female interactants have just discussed how the owners of a high-end restaurant in the city of Toronto, which has just passed a by-law banning smoking in restaurants, continue to allow indoor smoking, and they now jokingly imagine why this is possible:

- P They probably **paid off somebody**
 G *Yea::h*
 P From the government or something
 G (*laughs*)₍₁₎
 P To make sure that uh (*laughing*) *inspectors* don't come
 G (*laughs*)₍₂₎

Example 5 Laughter around smoking indoors humour; implicit coupling in bold

This example shows speaker P joking that both the restaurant owners and the government have broken the law and engaged in bribery. By doing so, P implies a coupling of positive judgement for bribery as sanctioned behaviour on the part of both the restaurant owners and the government, creating an affiliative tension with the values shared with her interactants as law-abiding citizens. In the system of move in Negotiation, the laughs by G indicate reactions, which work to **laugh off the tension created by the positive attitudinal coupling**. At the same time, the laugh shows that an additional attitude is being shared towards the restaurant and the government, as their 'breaking of the law' is laughable (rather than actually sanctionable), and so in reality the supposed 'law-breakers' are positively judged as aligning with these interactants as law-abiding citizens (and administrators of the law).

This additional attitude and their affiliation with the restaurant are varied, however, by the particular choices in sounding of G's reacting laughter. In Example 5, the first laugh (1) is a mid-pitch, neutral, single burst that is moderate in volume, and is noticeably short and stable. G's second laugh (2) is similar and only slightly higher in pitch, with a short, quick pulse. These follow from her first reaction to the utterance *They probably paid off somebody* with *Yea::h*, in which she indicates a possible agreement for negatively judging the restaurant for its sanctionable behaviour (because she does not at first laugh it off). With her short, quick laughter pulses that follow, G indicates that she cannot fully share the underlying restaurant community by which these utterances are taken to be funny. In fact, G makes clearer in later talk that she is not a regular visitor to the restaurant (*I haven't been ever!*), while P is ('I've been two times! . . . just this year'). The laugh expressions thus show that to laugh off

the tension of those couplings as 'funny', the participants must be able to construe both of the communities involved in the contrast (while it is easy to laugh off the government as law-breakers, it is not so easy to laugh off this behaviour by the restaurant unless one is a member of its community). The choices made in the laughter expressions in this example alter the affiliative negotiation as the laugh indicates attitude and fulfils a move in the exchange, and this exhibits the significance of the work that laughter is doing in conversational humour.

The laughter in Example 6 not only marks the humorous tension that is detected, but in its expression it also conveys negative judgement towards the other conversational speaker who is not 'in on the joke'. The speakers K and T are a married couple discussing an aspect of Brazilian culture with their Canadian dinner guests. T is a Brazilian man and K a Canadian woman, and they live together in Canada. A tension between their differential cultural (and perhaps gender) memberships is acknowledged by K. She attempts to laugh off the couplings made by Brazilian men, while T tries to share them:

- K: Yeah but you see a lot of guys in Brazil who aren't necessarily gay who **like to dress like women** an . . . Because I remember being at = =
- T = = Oh you're talking about (festival) right
- K the Carnival and like a whole group of guys they were all **dressed like women** = =
- T = = Yeah but they're not men dressed like women; they're like in a *costume* like a little costume like you know whaddamean? You can li- they're not reading into this about women's feelings you know what I mean? They-they don't wanna know what it's about to be a woman. They-they **wanna just have fun** an-an I don't know **pick up girls** that's the idea of the thing. Well that's how they () = =
- K = = °**Dressed like a girl**° (*laughs*) = =
- T = = Well they don't really *dress* like a girl! Alright?

Example 6 Laughter indicating negative judgement and marking tension; couplings in bold

K reacts with laughter to the couplings presented by T, which include for instance a positive appreciation for 'picking up' girls. This is laughable to her because they contrast with the positive appreciation coupling with cross-dressing that she has presented as an activity loved by Brazilian heterosexual men. K punctuates her own speech with laughter, marking her implied coupling as laughable in relation to what T has presented.

In terms of sound, her laugh is quiet, half-close, and pulsed through her own speech, but as the laugh continues past her own speech and through T's following utterance, the pitch moves from low to high, and the constriction

moves to nearly close. The quiet, near close and high-pitched quality of her laughter indicates negative judgement (and may also indicate self-consciousness on her part for doing so, cf. Edmondson 1987) and further suggests that T is the target as she continues laughing through his following speech. While T continues to construe himself as a serious member of the Brazilian male community, K's laughter expression not only acknowledges the tension his values create but also conveys her own judgement of him, affecting their affiliation.

It is also informative to consider the prosody of laughter in a text, or specifically, a humorous phase of discourse, as the changes in participant laughter as they construe a humorous sequence indicate the affiliation process that is occurring. We expand on the laughter description for Example 1 to exhibit the way that the changing laughter expression affects the meanings it conveys (see Example 7). Recall that the three interactants construe a family community in which eating heavy foods was a value they shared with that community over the holidays, but something that they need to laugh off to share a young female community in their conversation:

- U = = Yeah I saw like my family and friends . . . I **ate well** (*laughs*) ₍₁₎
 N We all **ate well**.
 (*all laugh*) ₍₂₎
 N Dude we all (*laughing*) ate **good pie!**
 (*continuous laughing*)
 U Yes I agree. (*continuous laughing*) **On a diet** now.
 (*all laugh*) ₍₃₎

Example 7 Laughter indicating meaning through prosody of conversation; couplings in bold

While the speakers present the couplings as creating tension, the laughter exhibits a rising solidarity as they all share belonging to both of the communities construed.¹⁰ The first speaker marks her coupling ('I ate well') as laughable by expressing a single quiet breathy burst with a front-spread posture following her own speech in a continuing move in ₍₁₎ (see Example 7). She has coupled positive appreciation with heavy eating, and her somewhat nervous (or self-conscious) laughter indicates a negative self-judgement for having done so and creating an affiliative tension that needs to be laughed off with the others.

This is made more explicit when the following speaker shifts the underlying judgement towards all those in the conversation, and reiterates the laughable coupling ('We all ate well'). The reacting laughter ₍₂₎ (see Example 7) is marked by an increase in amplitude and a decrease in pitch with continuous iteration, and is shared by speaker and all hearers. Together the participants laugh off the tension that their 'ate + well' coupling causes together, and they exhibit their shared memberships to both the family and the young female communities being construed (as they all participated in this 'bad behaviour'). Towards the

end of the phase, they begin to negotiate even within the young female community by laughing off dieting and their own negative self-judgements, and as a response, the laughter in ⁽³⁾ (see Example 7) is even louder, with a more open constriction in its iteration. The negative self-judgement is now jubilantly laughed off, and this roar is shared as the interactants achieve solidarity in affiliation by identifying as close members of similar contrasting communities, laughing off those tensions that their respective couplings cause for all of them as family members. The **prosodic unfolding of the laughing** from a single quiet burst marking the humorous tension to a **shared roar** indicates the progression of affiliation, and the **achievement in the moment-to-moment negotiation of community through convivial conversational humour**.

These examples display various combinations of sounding choices in laughter that speakers can make in convivial conversational humour, and suggest that in relation to the context and verbal co-text, particular forms of laughter indicate distinguishable attitudinal and affiliative meanings. Their placement in the text also shows how moves in Negotiation are indicated and impact upon the affiliative relations of the participants. The meanings indicated by laughter in this context show a development from the social functions of laughter as a signal of 'non-threat' to its role as an indicator of a humorous (or non-threatening) tension between the social values of communities in convivial conversational humour. Laughter is not only a semiotic system that combines facial expression and vocalization to construe various interpersonal meanings, but in combination with speech as well, its meaning potential grows as an essential component of the social negotiation of affiliation.

Conclusion

While laughter has been variously linked to differing origins and social functions, its development as a semiotic system functioning interpersonally, and complementing speech in the negotiation of affiliation, exhibits its role as a meaningful mechanism for the maintenance of **cohesive relations** between interactants. The meaning potential of laughter has developed from the micro-functions of the reflective mode into the array of **interpersonal discourse semantic systems** that are shared in language. **Systematized choices** in sound may be combined to make particular meanings within a **specified context**, and this has been exhibited through convivial conversational humour. Interactants combine these **variables** to indicate particular attitudes and to negotiate different degrees of affiliation in relation to their complex identities, and in this way, laughter functions as a powerful tool in casual conversation for the management of social values that bring people together in communities of the culture. As a complementary semiotic to language, their co-articulation demonstrates not only the intrinsic functionality and expanding meaning potential that these combined semiotic systems make possible, but it also displays the development

this
ana
mak
'sys
are
dete
Aga
crite

is co
analy

of laughter as a social semiotic in its own right, which enables the constant negotiation of similarity and difference that characterizes casual conversation. Beyond conversation, laughter has also been shown to convey a variety of meanings from play to a display of superiority, indicating that laughter may be the cipher-key for unlocking a world of semiotic potential beyond speech in systemic functional linguistic research. This study has provided an initial attempt to open the door.

Acknowledgements

This chapter is based on a presentation with Chris Cléirigh at the Semiotic Margins conference at the University of Sydney in December 2007. I extend my utmost appreciation to him for his help and contributions to earlier drafts of this chapter.

Notes

- ¹ This can be explained in relation to the systemic functional classification of intensive identifying processes in lexicogrammar, in which classes of signifying processes are separated according to the relationship between Token and Value (what is identified) (cf. Halliday & Matthiessen 2004:238). Meanings that are *realized* or *denoted* within semiotic systems (in lexical items such as ‘signify’ and ‘realize’) are distinguished from those that are *suggested* rather than denoted (in lexical items such as ‘indicate’ and ‘suggest’) (and these are also distinguished from relationships between non-semiotic manifestations and their meanings) (Martin 1992:280–282), and this reflects the difference between language and semiotic systems like laughter.
- ² This is similar to Provine’s (2000) term ‘convivial humor’, but is specific to conversation between friends and intimates, characterized by shared laughter and the negotiation of community values. This is also a reformulation of the earlier term ‘cooperative conversational humour’ used by Knight (2008), and was recommended by Salvatore Attardo (personal communication, 2008) to remove its association with the pragmatics notion of ‘cooperation’.
- ³ Names have been changed for privacy.
- ⁴ Because they can be variously negotiated, bonds here differs from Stenglin’s (2004) bonding and the notion of ‘bonding icons’ in that bonding icons bring interactants together into communities around quite strong and serious values such as nationhood (see Stenglin 2004:410) and peace (see Martin 2008:131) that cannot be laughed off.
- ⁵ The ‘+’ symbol will hereafter denote the coupling of attitude with ideational meaning.
- ⁶ We may also reject couplings altogether in the ‘condemning’ strategy of affiliation (see Knight 2008), such as in discourses of gossip.
- ⁷ These may correspond with ‘calls’ in laughter for the former and ‘bouts’ of laughter for the latter (see Owren 2007). Chafe (2007) also refers to these as

'pulses' and 'laugh clusters'. Each laugh should be considered as a whole, but its pulses can be distinguished by the constriction, posture and voicing characteristics, while the whole cluster makes differences in amplitude, length, character and pitch in relation to all of its pulses.

⁸ This classification is, however, incorporated here as an overall option rather than in relation only to what Halliday has classified as 'aperture'.

⁹ This is adapted from Halliday's (1992) system for aperture, but in constriction, it is the opening and closure of the vocal chamber with the lips rather than its narrowing or opening by the placement of the tongue that is chosen from in laughter.

¹⁰ Extended thanks to John Knox for his feedback in regard to the prosody of laughter in this clip.

References

- Abercrombie, D. (1968). Paralanguage. *British Journal of Disorders of Communication*, 3, 55–59.
- Apte, M.L. (1985). *Humor and laughter: An anthropological approach*. Ithaca, NY: Cornell University Press.
- Archakis, A. & Tsakona, V. (2005). Analyzing conversational data in GTVH terms: A new approach to the issue of identity construction via humor. *Humor*, 18(1), 41–68.
- Bachorowski, J.A. & Owren, M.J. (2001). Not all laughs are alike: Voiced but not unvoiced laughter readily elicits positive affect. *Psychological Science*, 12(3), 252–257.
- Bateson, G. (1987). *Steps to an ecology of mind*. New Jersey, London: Jason Aronson.
- Bonaiuto, M., Castellana, E. & Pierro, A. (2003). Arguing and laughing: The use of humor to negotiate in group discussions. *Humor*, 16(2), 183–223.
- Chafe, W. (2001). Laughing while talking. In D. Tannen & J.E. Alatis (Eds), *Linguistics, language, and the real world: Discourse and beyond* (pp. 36–49). Washington, D.C.: Georgetown University Press.
- Chafe, W. (2007). The importance of not being earnest: The feeling behind laughter and humor. Amsterdam/ Philadelphia, PA: John Benjamins.
- Clark, J. & Yallop, C. (1990). *An introduction to phonetics & phonology*. Oxford: Basil Blackwell.
- Coates, J. (2007). Talk in a play frame: More on laughter and intimacy. *Journal of Pragmatics*, 39, 29–49.
- Darwin, C. (1965 [1872]). *The expression of the emotions in man and animals*. Chicago, IL: University of Chicago Press.
- Devillers, L. & Vidrascu, P. (2007). Positive and negative emotional states behind the laughs in spontaneous spoken dialogs. Paper presented at the Interdisciplinary Workshop on The Phonetics of Laughter, Saarbrücken, 4–5 August.
- Edmondson, M.S. (1987). Notes on laughter. *Anthropological Linguistics*, 29, 23–34.
- Eggins, S. & Slade, D. (1997). *Analysing casual conversation*. London, New York: Cassell.
- Ellis, Y. (1997). Laughing together: Laughter as a feature of affiliation in French conversation. *Journal of French Language Studies*, 7 (2), 147–161.

- Freud, S. (1976 [1905]). *Jokes and their relation to the unconscious* (J. Strachey, trans. and A. Richard, ed.). Harmondsworth: Penguin Books.
- Glenn, P. (2003). *Laughter in interaction*. Cambridge: Cambridge University Press.
- Goodwin, C. (1986). Audience diversity, participation and interpretation. *Text*, 6 (3), 283–316.
- Gumperz, J. (1982). *Discourse strategies*. Cambridge: Cambridge University Press.
- Halliday, M.A.K. (1975). *Learning how to mean: Explorations in the development of language*. London: Edward Arnold.
- Halliday, M.A.K. (1978a). *Language as social semiotic: The social interpretation of language and meaning*. London: Edward Arnold.
- Halliday, M.A.K. (1978b). Meaning and the construction of reality in early childhood. In H.L. Pick & E. Saltzman (Eds), *Modes of perceiving and processing of information* (pp. 67–96). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Halliday, M.A.K. (1992). A systemic interpretation of Peking syllable finals. In P. Tench (Ed.), *Studies in systemic phonology* (pp. 98–121). London, New York: Pinter.
- Halliday, M.A.K. & Matthiessen, C.M.I.M. (1999). *Construing experience through meaning: A language-based approach to cognition*. London: Cassell.
- Halliday M.A.K. & Matthiessen, C.M.I.M. (2004). *An introduction to functional grammar* (3rd edn). London: Edward Arnold.
- Jefferson, G. (1979). A technique for inviting laughter and its subsequent acceptance/declination. In G. Psathas (Ed.), *Everyday language: Studies in ethnomethodology* (pp. 79–96). New York: Irvington Publishers.
- Jefferson, G. (1984). On the organization of laughter in talk about troubles. In J.M. Atkinson & J. Heritage (Eds), *Everyday language: Studies in ethnomethodology* (pp. 346–369). New York: Irvington Publishers.
- Jefferson, G. (1985). An exercise in the transcription and analysis of laughter. In T.A. van Dijk (Ed.), *Handbook of discourse analysis: Volume 3* (pp. 25–34). London: Academic Press.
- Jefferson, G., Sacks, H. & Schegloff, E. (1987). Notes on laughter in the pursuit of intimacy. In G. Button & J.R.E. Lee (Eds), *Talk and social organisation* (pp. 152–205). Clevedon: Multilingual Matters.
- Knight, N.K. (2010). Wrinkling complexity: Concepts of identity and affiliation in humour. In M. Bednarek & J.R. Martin (Eds), *New discourse on language: Functional perspectives on multimodality, identity, and affiliation* (pp. 59–98). London: Continuum.
- Knight, N.K. (2008). 'Still cool . . . and american too!': An SFL analysis of deferred bonds in internet messaging humour. In N. Norgaard (Ed.), *Systemic functional linguistics in use* (pp. 481–502). Odense: Odense Working Papers in Language and Communication, vol. 29.
- Knight, N.K. (in preparation). Laughing our bonds off: Conversational humour in relation to affiliation. PhD Thesis in progress, Department of Linguistics, University of Sydney.
- Koestler, A. (1964). *The act of creation*. London: Hutchinson.
- Martin, J.R. (1992). *English text: System and structure*. Philadelphia, PA: John Benjamins.
- Martin, J.R. (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.R. (2008). Intermodal reconciliation: Mates in arms. In L. Unsworth (Ed.), *New literacies and the English curriculum: Multimodal perspectives* (pp. 112–148). London: Continuum.
- Martin, J.R. & Rose, D. (2007). *Working with discourse: Meaning beyond the clause* (2nd edn). London: Continuum.
- Martin J.R. & White, P.R.R. (2005). *The language of evaluation: Appraisal in English*. New York: Palgrave Macmillan.
- Martin, R.A. (2001). Humor, laughter and, physical health: Methodological issues and research findings. *Psychological Bulletin*, 127(4), 504–519.
- Matthiessen, C.M.I.M. (2006). The multimodal page: A systemic functional exploration. In T. Royce & W. Bowcher (Eds), *New directions in the analysis of multimodal discourse* (pp. 1–62). Mahwah, NJ: Lawrence Erlbaum Associates.
- Morreall, J. (1983). *Taking laughter seriously*. Albany, NY: State University of New York.
- Morris, D. (1967). *The naked ape*. New York: Dell.
- Norrick, N.R. (1993). *Conversational joking: Humor in everyday talk*. Indianapolis, IN: Indiana University Press.
- Nwokah, E., Hsu, Hui-Chin & Davies P. (1999). The integration of laughter and speech in vocal communication: A dynamic systems perspective. *Journal of Speech, Language and Hearing Research*, 42, 880–894.
- Owren, M.J. (2007). Understanding acoustics and function in spontaneous human laughter. Proceedings of the Interdisciplinary Workshop on the Phonetics of Laughter, Saarbrücken, Germany, 4–5 August.
- Owren, M.J. & Bachorowski, J.A. (2003). Reconsidering the evolution of non-linguistic communication: The case of laughter. *Journal of Nonverbal Behaviour*, 27 (3), 183–200.
- Painter, C. (1984). *Into the mother tongue: A case study in early language development*. London: Frances Pinter.
- Painter, C. (2003). Developing attitude: An ontogenetic perspective on appraisal. In M. Macken-Horarik & J.R. Martin (Eds), *Text, special issue – Negotiating heteroglossia: Social perspectives on evaluation* (pp. 183–209). Berlin, New York: Mouton de Gruyter.
- Partington, A. (2006). *The linguistics of laughter: A corpus-assisted study of laughter-talk*. London: Routledge.
- Poyatos, F. (1993). *Paralanguage: A linguistic and interdisciplinary approach to interactive speech and sound*. Amsterdam/Philadelphia, PA: John Benjamins.
- Provine, R. (1992). Contagious laughter: Laughter is a sufficient stimulus for laughs and smiles. *Bulletin of the Psychonomic Society*, 30, 1–4.
- Provine, R. (2000). *Laughter: A scientific investigation*. London: Faber.
- Provine, R. (2004). Laughing, tickling, and the evolution of speech and self. *Current Directions in Psychological Science*, 13(6), 215–218.
- Raskin, V. (1985). *Semantic mechanisms of humor*. Dordrecht: D. Reidel Publishing Company.
- Ruch, W. & Ekman, P. (2001). The expressive pattern of laughter. In A.W. Kaszniak (Ed.), *Emotion, qualia, and consciousness* (pp. 426–443). Tokyo: Word Scientific Publisher.
- Sacks, H. (1974). An analysis of the course of a joke's telling in conversation. In R. Bauman & J. Sherzer (Eds), *Explorations in the ethnography of speaking* (pp. 337–353). Cambridge: Cambridge University Press.

- Sacks, H., Schegloff, E.A. & Jefferson, G. (1974). A simplest systematics for the organization of turn-taking for conversation. *Language*, 50, 696–735.
- Spencer, H. (2007 [1911]). The physiology of laughter. In H. Spencer (Ed.), *Essays on education and kindred subjects* (pp. 298–309). London: Dent.
- Sroufe, L.A. & Wunsch, J.P. (1972). The development of laughter in the first year of life. *Child Development*, 43, 1326–1344.
- Stenglin, M.K. (2004). Packaging curiosities: Towards a grammar of three-dimensional space. PhD Thesis, University of Sydney.
- Stewart, S. (1995). The multiple functions of laughter in a Dominican Spanish conversation. Paper presented at the Language South of the Rio Bravo Conference, Tulane University.
- Trouvain, Jürgen (2001). Phonetic aspects of ‘Speech-laugh’. Proceedings of the conference on orality & gestuality (ORAGE) (pp. 634–639). Aix-en-Provence (France).
- van Hooff, J.A. (1967). The facial displays of the catarrhine monkeys and apes. In D. Morris (Ed.), *Primate ethology* (pp. 7–68). Chicago, IL: Aldine.
- van Hooff, J.A. (1972). A comparative approach to the phylogeny of laughter and smiling. In R. Hinde (Ed.), *Non-verbal communication* (pp. 209–241). Cambridge: Cambridge University Press.
- van Leeuwen, T. (1999). *Speech, music, sound*. Basingstoke: Macmillan.
- Vettin, J. & Todt, D. (2004). Laughter in conversation: Features of occurrence and acoustic structure. *Journal of Nonverbal Behaviour*, 28(2), 93–115.
- Warren, J.E., Sauter, D.A., Eisner, F., Wiland, J., Dresner, M.A., Wise, R.J., Rosen, S. & Scott, S.K., (2006). Positive emotions preferentially engage an auditory-motor ‘mirror’ system. *The Journal of Neuroscience*, 26(50), 13067–13075.
- Zijderveld, A.C. (1983). The sociology of humour and laughter. *Current Sociology*, 31, 1–100.

Chapter 2

Body Language in Face-to-face Teaching: A Focus on Textual and Interpersonal Meaning

Susan Hood

University of Technology, Sydney

Introduction

The rapid expansion of computer-mediated interaction in pedagogic contexts has focused much critical attention on the modalities of e-teaching and e-learning, suggesting an array of ‘new’ modes of interaction. At the same time, however, there has also been much renewed interest in the multimodality of what is sometimes dismissively referred to as the ‘traditional’ face-to-face classroom. Face-to-face classrooms are now recognized as most complex pedagogic sites involving simultaneous engagements with at least the modalities of speech, written texts, visuals, space and body language, including facial expression and gaze (Kress et al. 2001, Jewitt 2008, Bourne 2003, Lund 2007). The analyses in this chapter focus in particular on the modalities of speech and body language. The intention is to make visible ways in which body language functions in collaboration with spoken language in teachers’ discourse to shift student attention to particular kinds of information, to manage processes of student interaction and engagement, and interpretation of meanings. The aim is to contribute to social semiotic theorizing of the meaning potential of body language, and ultimately to identify ways in which teachers’ embodied meaning-making can contribute to effective pedagogic practice.

Theory

The study of body language presented in this chapter builds on foundational studies in gesture from a number of fields, including the seminal work in cognition of McNeill (1992, 1998, 2000), Kendon (1980, 2004) and more recently, Enfield (2009). More directly, it draws on a growing field of social semiotics which has over recent years extended beyond language to include modelling of the semiotic modes of image (Kress & van Leeuwen 2006, Painter 2007), space (Stenglin 2004, 2007, Martin & Stenglin 2006), typography (van Leeuwen 2006),

sound, music and voice quality (van Leeuwen 1999, McDonald, Chapter 5), facial expression, gesture and position (Martinec 2000, 2001, 2002, 2004, Munitgl 2004), and importantly to theorizing the relationships within and across different semiotic systems (Bednarek & Martin 2010, Painter & Martin, in press, Martinec & Salway 2005, Royce & Bowcher 2007, Ventola, Charles & Kaltenbacher 2004).

While referencing the influences of studies in cognition and in social semiotics, it is important to note that each discipline approaches research on body language from different premises and different theories, or interpretations of theory. Studies in cognition, as articulated, for example, in Enfield (2009), are primarily interested in cognitive processes of intention and interpretation. Enfield explains the quest as understanding 'how it is that interpreters may derive meaning from composite utterances, or how we recognise "others" communicative and informative intentions' (2009:1). From a grounding in cognition, Enfield critiques what he describes as a (neo-)Saussurean view of meaning – 'that a sign has meaning because it specifies a standing-for relation between a signifier and a signified'. This interpretation is then negatively evaluated as a view of sign as 'static, arbitrary and abstract' (2009:2). Enfield argues the need to explain meaning as processes of interpretation of signs that are 'dynamic, motivated and concrete'. He suggests that the only alternative to 'a static view of meaning' is available through Peircean semiotics (e.g. Peirce 1955) or through pragmatics (e.g. Grice 1975, Levinson 1983).

In this chapter, I take a different perspective, based on a different conceptualization of meaning and a different interpretation of Saussure than is considered in Enfield's argument. Revisiting the notion of the sign in Saussurean linguistics, an alternative interpretation to that in Enfield (2009) is provided by Martin (1992, 2007) who argues with reference to Hjelmslev (1961) that the domain of social semiotics is not a theorization of the relation between signifier and signified, but is in fact the theorization of the delineating line – the space between the two dimensions. The Saussurean contribution is to bind the signified and signifier into sign and then to theorize language as system of signs. As a system of signs, the potential to mean is in the relationship of signs to other signs in the system. We mean in relation to what we could have meant but did not (Martin 1992). Hjelmslev expands the meaning potential of this space (of systems of signs) as a stratified system of signs, that is, as expression form and content form. In Systemic Functional Linguistics (SFL) (Halliday 1978, 1992, Martin 1992, Martin & Rose 2007), the content form of language as a system of signs is then further stratified as *discourse semantics* and *lexicogrammar*. The relationship across these strata is one of abstraction. Martin (1992) extends the system of signs further to a stratified context plane (context form) of *genre* and *register*. This already rich theorization of sign systems acquires greater explanatory power when the *hierarchy of realization* (briefly articulated above as stratification) is complemented with the *hierarchy of instantiation* (Halliday 1991, 1992). Instantiation has to do with constraints on the generalized meaning potential of the system through genres and registers to specific instantiations in texts. The resultant theorization of the system of signs, of how we mean in

process

language and beyond language in other semioses in Systemic Functional Semiotics (SFS) (e.g. Painter & Martin, in press, Martinec & Salway 2005, Royce & Bowcher 2007, Ventola, Charles & Kaltenbacher 2004) is far from Enfield's description of a 'standing-for relation between a signifier and a signified'.

As the discussion above makes evident, the way in which language itself is theorized is a significant variable in the analysis and interpretation of body language and its relationship with spoken language. In this light, there are some concepts in Systemic Functional Semiotic theory that require additional explanation. The first is the conceptualization of meaning as *metafunctional*, incorporating notions of ideational meaning, interpersonal meaning and textual meaning (Halliday 1994). Ideational meaning refers to the way we represent 'reality' as configurations of kinds of processes, participants and circumstances. Interpersonal meaning refers to the ways in which we exchange values with each other and construct relationships of power and of solidarity. Textual meaning refers to the ways in which we make sense in the context within which we interact; how we organize and package ideational and interpersonal meanings to make ourselves understood. In language, all three metafunctions mean simultaneously in discourse such that the same wordings can be re-analysed for the ways in which they function in relation to each kind of meaning. In a social semiotic analysis of body language too, we can ask how postures and gestures offer the potential to mean metafunctionally. We can also consider the extent to which particular metafunctional meanings might be fused or co-instantiated in a 'single' gesture.

A further aspect of SFL theory, the **hierarchy of instantiation** models the relationship between systems of meaning potential and any instance of meaning in a text (Halliday 1992, Halliday & Matthiessen 1999, Martin & Rose 2007, Martin 2010). A critical concept in relation to instantiation is that of *commitment* which refers to the degree of meaning potential instantiated at any point in the discourse (Martin 2006, Hood 2008). **To the extent that meanings are instantiated in more than one semiotic system, the theorization of instantiation provides us with a framework for interpreting and explaining the relationship of meanings committed in instances of speech and in instances of accompanying body language. We can consider the kinds of meanings and the degree of meaning potential committed in each.**

Finally, we need to explore more closely the relationship of body language to language. Cléirigh (in Martin, Chapter 12) differentiates kinds of body language in terms of their relation to language. He distinguishes in body language 'three types of semiotic systems: protolinguistic, linguistic and epilinguistic'. Protolinguistic body language 'is a development from infant protolanguage', that is, it is the systems 'left behind in the transition to the mother tongue'. It consists of expression and meaning only, does not need accompanying speech to mean and is exemplified, for example, in a postural orientation realizing involvement, or fidgeting realizing discomfort. **Linguistic body language on the other hand 'only occurs during speech'.** These movements synchronize with the rhythm and intonation of prosodic phonology in language and so express

salience and tone, co-instantiating textual and interpersonal meanings. The third system, that of epilinguistic body language, is 'made possible by transition [from protolanguage] into language, but [is] not systematically related to the lexicogrammar of language (. . .) realis[ing] meanings rather than wordings'. When accompanying speech, epilinguistic body language makes visible the semantics of speech. Without speech it constitutes mime. Epilinguistic body language can instantiate all three metafunctions: ideational, interpersonal and textual. It is the system of epilinguistic body language that is explored in this study, that is, body language that relates to meanings rather than wordings. In order to provide adequate depth of analysis in this chapter I restrict the focus to embodiments of textual and interpersonal meaning. Within textual meaning I explore meanings of identification and of the phasal shifts in the discourse. Within interpersonal meaning, the focus is on the invocation of attitude and in the management of space for other voices. Finally, I attend briefly to the ways in which a gesture may fuse both textual and interpersonal meaning.

Locating the Study in Face-to-face Classrooms

There are several reasons why classrooms provide rich sites for analysis of body language accompanying spoken discourse. In the first place face-to-face classrooms are sites for a complex interaction between different semiotic modes, technologies and artefacts (Kress et al. 2001, Flewitt 2006). Teachers' roles in managing and integrating these resources in interaction with students can be expected to involve considerable bodily movement. Face-to-face classrooms are also sites for the enactment of a range of different spoken genres, from procedures and protocols, to explanations, discussions and arguments, 'story' genres and casual conversation. The range of social purposes provides a range of social functions for gestural expression (Gullberg 1999). Some genres of the classroom are likely to provide increased density in gestural expression. If teachers are engaged in tasks of detailed explanation, particularly where it is anticipated, tasks will challenge students because they involve complex and abstract concepts and ideas, and/or because students are learning a second language, then we might expect a greater redundancy in meaning-making in teacher discourse. A need for redundancy can be expected to impact on both verbal and gestural expression of meaning.

For this study, digital video recordings were made of six advanced level adult classes in English for academic purposes, with the written consent of teachers and students. In total, three teachers were filmed, all experienced in this context. The filming took place in the first hour of a 4-hour class after which there was a short break, signalling a shift in activity or task focus. The data from each lesson therefore represents a stage in the longer curriculum macrogenre (Christie 1997) marked by its own initiation and closure discourse (and actions). Within the initial stage of approximately 60 minutes there are identifiable phases of interaction, once again marked in the discourse and representing

sub-stages of activities or tasks. And at a more micro level again, within such phases a series of episodes of interaction can be identified by shifts in the pattern of interaction in which the teacher and students are engaged. The data were viewed multiple times as whole lessons enabling the researchers to track shifts in patterns of body language as phases of lessons. Detailed transcriptions were made of the verbiage and descriptions of gestures for selected phases and sequences of phases.

In this chapter, I focus on phases of lessons in which the teacher is fronting the class and engaged in episodes of instruction and explanation with some teacher-coordinated discussion. The language is dominantly monologic and the analyses focus on the teacher's embodied meanings co-expressed with spoken language. In all instances the teachers are intent on engaging students and guiding them to a greater understanding of aspects of content (academic writing).

Gesture and the Textual Patterning of Meanings

The bodily enactment of textual meaning can be considered from a number of perspectives. Here I focus initially on the way in which body language functions in the service of identification, in the integration of entities and multiple semiotic systems into the discourse.

Identification: Getting People, Places and Things into the Text

We readily recognize the role of the canonical 'pointing' gesture of the extended index finger in the service of identification. The vector constructed with the body can express directionality **towards another entity (human, material, semiotic)** that is thus referenced in the discourse, regardless of whether there is co-articulation of a verbal expression such as 'this', 'that', 'you', 'her' etc. Such gestures are representative of Cléirigh's category of epilinguistic body language (Martin, Chapter 12). Such body language need not be co-expressive with language and expresses meanings rather than wordings. Other parts of the body can also serve to construct vectors and hence directionality to a referent, including, for example, movements of the chin or head or directionality of gaze (Kendon 2004, Enfield 2009). The vector need not only direct away from the 'meaner' but can also self-reference, although there are then apparent restrictions on the parts of the body that can come into play. It is also noted here that pointing gestures can be subject to strong social and cultural taboos meaning that some options can be highly marked in one context or another (Efron 1972). It was noted that in the adult classrooms analysed in this study there were very few instances of teachers using finger-pointing gestures to identify students. Where this did occur it was a fleeting movement, often with supine hand (see later discussion). Other data suggest that this is a more commonly used gesture in primary school classrooms.



(a)



(b)

IMAGE 2.1a Identifying actual wordings**IMAGE 2.1b** Identifying potential wordings

While identification gestures can make reference to an entity that is co-visible in the shared material space of the meaner and interactant, they can also reference an entity that is not recoverable in that material space, one that is assumed to exist elsewhere or is hypothetical, imagined or potential. This meaning distinction is frequently enacted by one of the teachers in this study as she engaged students' attention with a model of a text projected on a whiteboard at the front of the room (Images 2.1a & 2.1b). While referring to the actual wording in the text the teacher points with her left hand to that wording. At other times she offers alternative meanings, to suggest what was not written and could have been written. At such points in the talk she would raise her hand off the text and point away and upward to her left (see Brady et al. 1995 on contact and distal gestures). In spoken language and gesture there is a corresponding shift from referencing actualized meanings to referencing potential meanings. In these episodes of interaction the teacher's body language plays a significant role in signalling to students a shift in orientation from what is there to what is not yet there, and could or perhaps should be there.

These analyses inform the development of a tentative system network for the potential of the body to mean textually in terms of identification, as in Figure 2.1a and 2.1b.

The hand and fingers are well suited to the construction of vectors and hence to the expression of direction and identification. The teachers in the study make use of the whole hand, side of the hand, index finger, index and middle finger and little finger, and in some cases other instruments are incorporated into the gesture such as a pen, or a whiteboard marker. The variations illustrated in Images 2.2a, 2.2b and 2.2c occur in one short episode

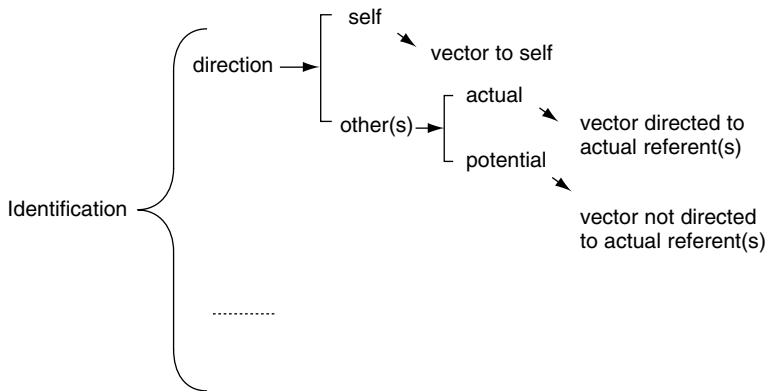


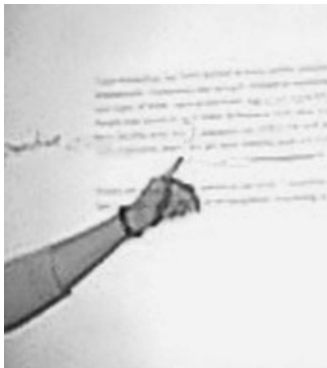
FIGURE 2.1a Partial system network for the body language of identification



(a)



(b)



(c)

IMAGES 2.2 (a), (b), (c) Identifying with different degrees of specificity

in the data. The teacher points with her hand, index finger then little finger, as she guides the students to attend to more general or more specific parts of the text or wordings. The teacher's body language functions here in relation to another dimension of identification, that of specificity.

We can interpret the variation in the bodily resources evident in Images 2.2a, 2.2b and 2.2c as varying along a cline of specificity. The smallest body part that enables the highest degree of specificity is the little finger. The system network for Identification can therefore be extended as in Figure 2.1b.

There is yet a third dimension of identification enacted in body language, that of specification as delineation. In this case a gesture is formed in such a way as to indicate boundaries. It may, for example, include two hands extended with palms vertical and facing inwards, as in instances where the teacher indicates three students sitting side by side as the ones she wants to form one group. But the delineation can also be enacted with boundaries formed by the bent finger and thumb as in Image 2.3 where the teacher specifies the boundary of what she wants students to attend to on a projected text.

The more complete system network of identification, represented in Figure 2.1c can be interpreted as meaning that where identification is enacted in body language, the gesture encodes direction to a referent (actual or potential) *and* a degree of particularization, *and* +/- delineation of boundaries for the referent.

In the data in this study, teachers typically rely heavily on resources of body language to construe meanings of identification. **We could say that considerably more meaning of identification is committed in the teachers' body language than is committed in their spoken language.** A general verbal reference to students as 'you', for example, might be committed with additional meaning of specificity in particularization and delineation in gesture. Similarly, verbal reference to a segment of text as 'this' could be further committed in body

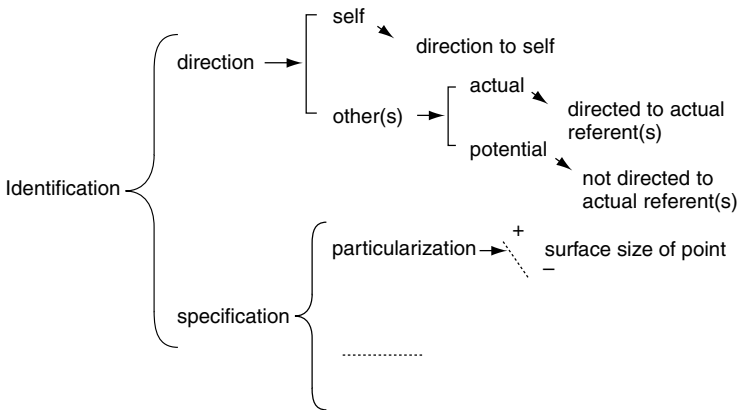


FIGURE 2.1b Partial system network for the body language of identification

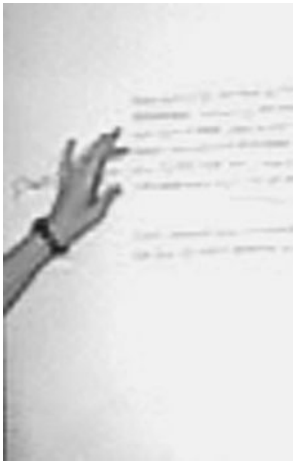


IMAGE 2.3 Identifying and delineating

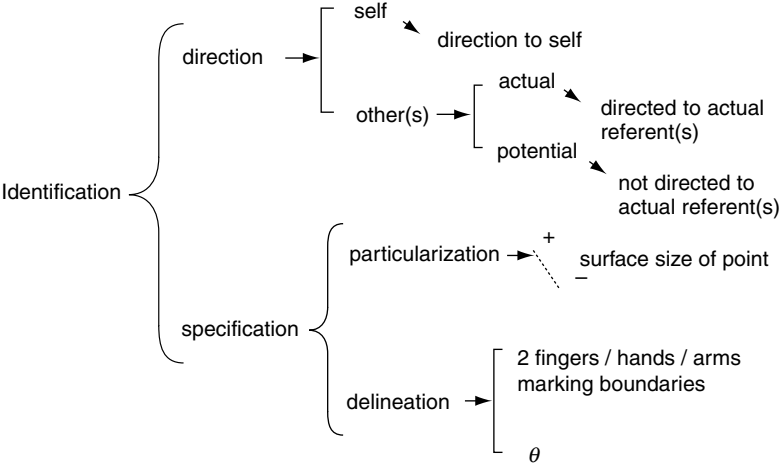


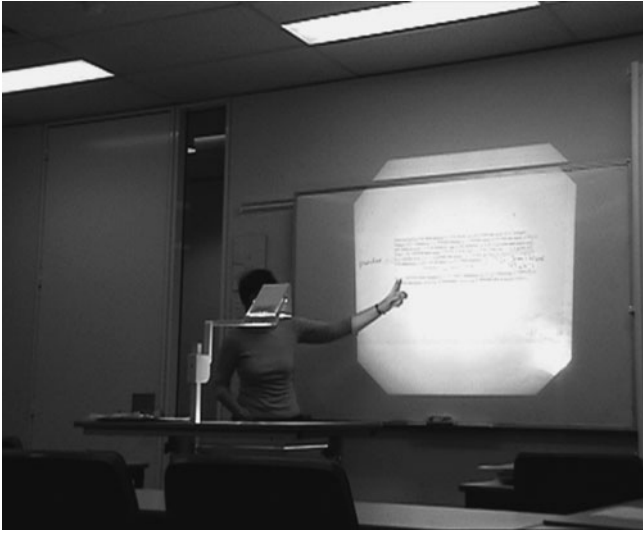
FIGURE 2.1c System network for the body language of identification

language as meaning generally ‘this’ or specifically ‘this’. The interpretation of teacher instructions and explanations typically requires students to interpret both body language and spoken language in interaction.

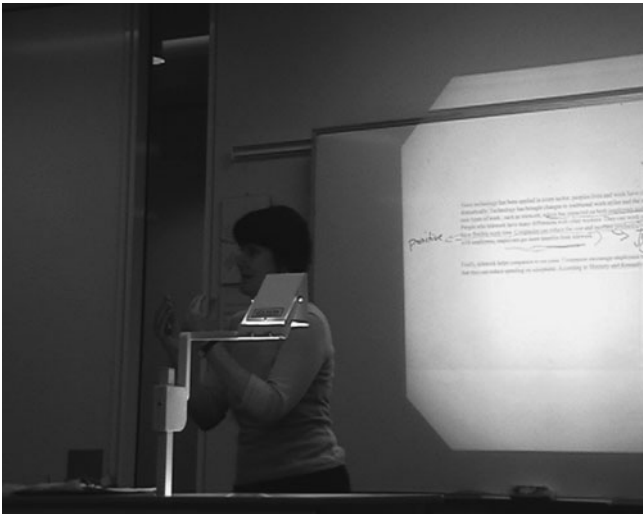
Body Language and Texturing the Flow of Discourse

A second way in which body language functions in relation to textual meaning is in periodic movements of the whole body. Such movements contribute to the

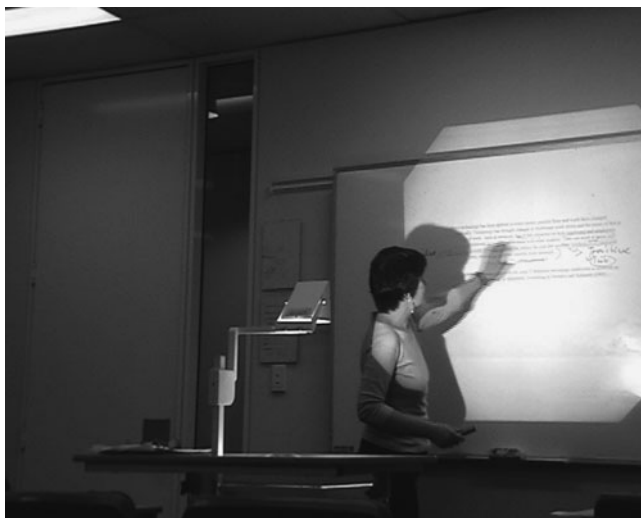
construal of phasal shifts or shifts in the patterning of metafunctional meanings (Gregory & Malcolm 1995) that are also evident in the spoken discourse. Over a segment of a lesson the teacher moves position away from the students to a whiteboard attached to the wall at the front of the room, then back closer to the students, then back to the board, back to the class and so on in several cycles of movement. Examples of the differing positions are illustrated in Images 2.4a, 2.4b and 2.4c.



(a)



(b)



(c)

IMAGES 2.4 (a), (b), (c) Cyclic movements towards board and back to class

The spoken language associated with one or other position (i.e. close to students or close to board) was transcribed, identifying tone groups, using Halliday's analytical framework of 5 tones: 1 = falling ('certain'); 2 = rising ('uncertain'); 3 = level ('unfinished'); 4 = fall-rise ('but'); 5 = rise-fall ('surprise') (Halliday 1963 (2005)), as well as shifts in ideational and interpersonal meaning choices. The epilinguistic body language characterizing each position is also described. The extract of the analyses across one wavelength in Table 2.1 (from close to board to close to class to close to board) represents the kinds of shifts in language and body language that are repeated across subsequent shifts of position.

An analysis of the teacher's spoken language and body language associated with each stage shows evidence of a shift in the level of actualization of key ideational meanings associated with the content of the text that the students are attending to. When the teacher is positioned at the board these meanings are construed as actual rather than potential. She refers to meanings realized in the discourse and verbalizes and gestures specific locations in the text. The referents for her identifying gestures are dominantly parts of the text. Her tone is predominantly one of certainty (tone 1: this is what is). Her role is dominantly to inform. When the teacher is close to the class the ideational meanings are construed as potential (to be elicited/negotiated) rather than actualized and they are de-specified through resources of focus (*some kind of*). Her tone is predominantly 'uncertain' (tone 2: what is it?). Her body language functions dominantly to identify her students and herself, and also to potential meanings

Table 2.1 Patterns of spoken language and body language construing phases of interaction

Spoken language in tone groups	Teacher's position	Body language patterns	Multi-semiotic phases of interaction
// 1 <u>We've got a 'but' in here.</u>	close to board	thumb and forefinger delineation of segment of text; touching board	informing + actual meaning tonally: 1 ('certain') verbally: generalized meaning of concession ('a "but"'); actual location (<u>here</u>) bodily: identify specific referent on text
// 2 <u>Yeah?</u>		(gesture held)	checking tonally: 2 ('uncertain')
// 1 <u>There's more 'benefits for the employers'.</u>		underlines wordings on text '...' with a pen	informing + actual wording tonally: 1 ('certain') verbally: specific wordings from text ('benefits ...') bodily: identify specific location/wordings on text
// 2 <u>So what would</u>		points to specific location in the text; touching text	transition into . . .
<u>we need to say before we say that?</u>	close to class	forearms to class and supine hands curled to self (see Image 2.4b)	eliciting + potential meaning / wording tonally: 2 ('uncertain') verbally: interrogative ('what . . . ?')
[St answer]		fleeting point with pen to student responding	bodily: identify sts and self.
// 2 <u>We'd need to say some kind of disadvantage or problem for the</u>		forearms to class and supine hands curled to self	eliciting + potential (defocused) meaning tonally: 2 ('uncertain') verbally: defocused meaning (<u>some kind of</u>) bodily: identify sts and self.
Ss + T: // 1 . ^ . . . <u>employees .</u>		2-hand delineated identification in space to right side of body	confirming + potential meaning tonally: 1 ('certain') verbally: specific meaning (<u>employees</u>) bodily: identify potential meaning (realized in space but not yet in text)
// 1 <u>Okay.</u>		backs off slightly from class	consolidating tonally: 1 ('certain') transition into . . .
// 1 <u>So we need to . . . bring in here</u>	moving to board close to board	sweeps left hand under wordings on the text (see Image 2.4c) finishing with fingers horizontal to board	informing + specific meanings tonally: 1 ('certain') verbally: high modulation; circumstance of location (' <u>here</u> ') bodily: identify specific location

(in space, not on the text). Her role is dominantly to elicit and engage. The teacher's shifts in position in the room correspond to shifts in the meanings she is orienting students to, from actual to potential, and from the written text to the students as potential writers. The teacher's shifts in position and accompanying shifts in patterns of body language function to texture the discourse and hence the teaching-learning activity into phases of interaction. Each multimodally constructed phase makes salient different kinds of information to be attended to by the students, and expresses different expectations in terms of student engagement and participation. These cyclical shifts can be interpreted as an aspect of the teacher's scaffolding of students' academic writing as she opens up space for new possibilities and guides students towards new instantiations of meaning.

While not analysed in this chapter, there are also movements of the body at much smaller wavelengths mainly involving the fingers and hands, and small movements of the head, which are synchronous with phonological rhythms of stress and intonation, movements that constitute linguistic body language according to Cléirigh (in Martin, Chapter 12) in contrast to the epilinguistic body language analysed here. While Eisentein (2008) does not differentiate kinds of movements in functional terms, the analyses of body movement presented here do correspond to his description, suggesting that

the small linguistic units (e.g., phrases) are synchronized with fast moving body parts (e.g., hands and fingers) and large discourse units (e.g., topic segments) are synchronised with slower moving body parts (e.g., the torso). (. . .) posture shifts occur much more frequently at segment boundaries. (Eisentein 2008:29)

The synchrony of body language and verbalized meanings functions, as Martinec (2000:293) suggests, 'to create order out of what may otherwise appear chaotic', and contributes significantly to a sense of ordered and organized flow of interaction in face-to-face teaching.

Body Language and the Expression of Interpersonal Meaning

Before attending to analyses of interpersonal meaning in epilinguistic body language a very brief review of interpersonal meaning in SFL is necessary. At the level of discourse semantics, interpersonal meaning is theorized as appraisal (Martin & White 2005), identifying domains of *attitude* (or the expression of values and feelings as affect, appreciation and judgement), *graduation* (where meanings are graded as degrees of force and sharpness of focus) and *engagement* (or the management of other voices in the discourse in terms of whether and how space is opened up or closed down to other voices).

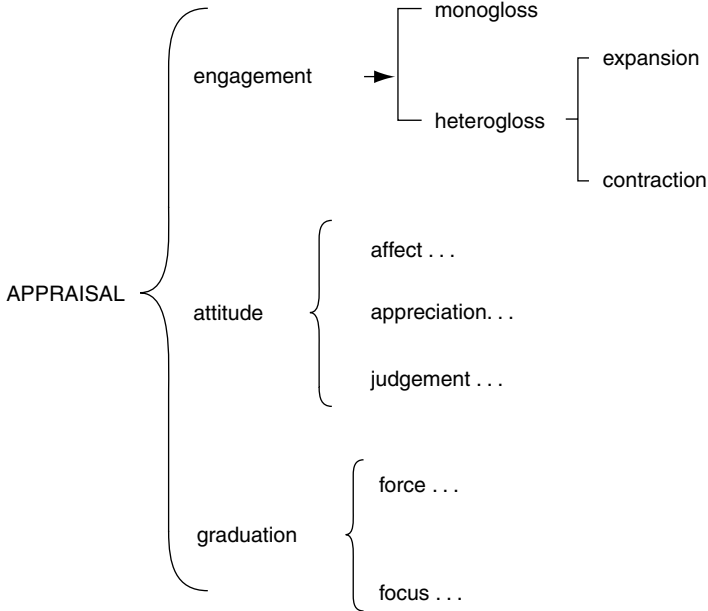


FIGURE 2.2 Dimensions of appraisal

(See Figure 2.2 for a skeletal model of appraisal and Martin & White 2005 for a comprehensive explanation).

In analysing attitude in verbal discourse a distinction can be drawn between attitude that is explicitly expressed or *inscribed* and attitude that is implied or *invoked* (Martin & White 2005). Graduation provides one important means by which attitudinal meanings can be invoked (Hood 2004, 2006, 2010). By grading an objective (ideational) meaning the speaker gives a subjective slant to that meaning, signalling for the meaning to be interpreted evaluatively. So, for example, when a teacher says ‘you all need to listen to this’, both *all* and *need* are instances of grading the force of what is said, implying though not explicitly encoding the meaning of ‘this is *very important*’.

In analysing body language co-expressed with spoken language we can consider these same dimensions of meaning (see Macken-Horarik 2004 on appraisal analysis of images). Resources of facial expressions are not analysed here but can, for example, function to express affect as happiness, sadness etc. But the body can also play a role in invoking attitude through the grading of meanings along a number of clines. Meanings can be graded in intensity in the muscle tension employed in gestures accompanying the verbiage. The intensification may or may not be co-expressed in the verbiage. Tension realizing intensification can be expressed in various parts of the body and is illustrated as tensed and relaxed hand muscles in Image 2.5a and 2.5b.



(a)

IMAGE 2.5a + muscle tension expressing intensification (. . . *the grammar rules*)



(b)

IMAGE 2.5b – muscle tension expressing lack of intensification (*how did you . . . ?*)

In Image 2.5a the teacher's gesture enacts a meaning of graduation as force, as 'mustness'. In Image 2.5b such a meaning is absent.

Another option for grading meanings is in terms of quantification. Here gestures offer resources in terms of size or scope (see Chafai et al. 2007 on expressivity). Grading up gesturally in this respect (greater size and/or scope) can co-instantiate with verbal amplification (e.g. 'this is *a big problem*' or 'this is *really significant*'), or the graduation can occur in gesture but not in wording, as illustrated in Image 2.6 where the verbiage in the caption does not of itself instantiate graduation. To the extent that interpersonal meaning is co-instantiated in body language and verbiage there can be mutual reinforcement of amplified meaning. **Where an interpersonal meaning is instantiated only in body language there is a distribution of the metafunctional load across speech and gesture.** The verbiage can carry the ideational load while the body appraises. An interpretation of the interpersonal meaning in the message relies on students interpreting the evaluative meaning in the gesture.

Bodily expressions of intensity or quantity offer a potential to be interpreted interpersonally. The extent to which they are so interpreted will depend on the context and co-text, necessarily implicating the co-instantiated spoken language. Body language in the service of interpersonal meaning is represented in Figure 2.3 in a tentative and partial system network. The dimension of focus has not been explored in this study.



IMAGE 2.6 Amplifying size: Invoking value (*That's what we're talking about!*)

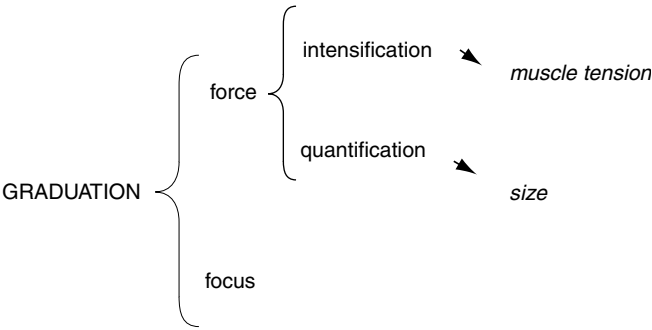
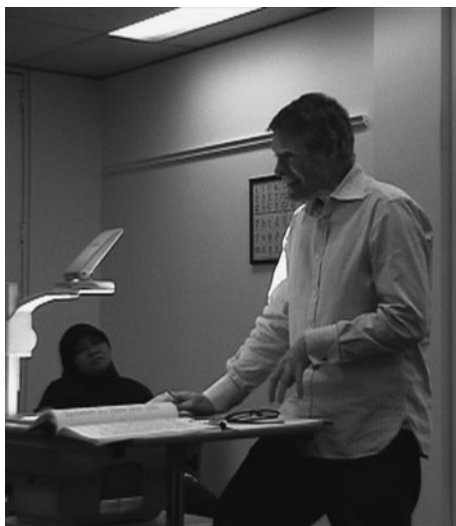


FIGURE 2.3 A partial system network of graduation in body language

Another set of findings that relate to interpersonal meaning concern the positioning of hands as a means for expressing what is referred to in appraisal theory as *engagement*. This has to do with the extent to which a speaker (or writer) expands or contracts space for other voices in their discourse. A basic distinction can be made in terms of the positioning of hands, contrasting a supine (palms up) position to a prone (palms down) position. The palms-up positioning embodies an elicitation move on the part of the teacher, enacting an expansion of heteroglossic space and so inviting student voices into the discourse (such gestures are evident in the elicitation moves by the teacher in Table 2.1). This may correspond to an interrogative structure in the verbiage. In Image 2.7a, for example, the teacher is asking students: *How did you work*



(a)



(b)

IMAGE 2.7a Expanding space for negotiation (*How did you work out the answer?*)

IMAGE 2.7b Contracting space for negotiation (*a draft isn't a complete rewrite*)

out the answer? A prone-hand gesture, in contrast, functions to close down space for other voices, and typically accompanies verbal discourse that functions in a corresponding way. In Image 2.7b the teacher is negating the possibility of other positions, with phonological stress on the negation (isn't) as he says *a draft isn't a complete rewrite*. While a supine and prone distinction is most often enacted with the hand, it may also be evident in the positioning of the index finger in pointing gestures. So pointing to a student to *invite* them to contribute to the discussion can be made with the inside of the index finger facing up, while a direction to do something can be made with the inside of the finger facing down.

The data also reveals a gesture constructed as a movement back and forth between that of supine and prone positions in an oscillating gesture. This is interpreted as expressing modality of possibility, and in terms of engagement, as expanding heteroglossic space by entertaining other possible positions. In these data it was always co-instantiated with a verbal expression of modality (congruent or metaphoric), and the extent to which a possibility is entertained as relatively likely or unlikely seems to depend on additional resources such as facial expression or voice quality. In these data the oscillation is typically enacted with the hands, but other parts of the body such as the head or even the upper torso can also be used in the expression of this meaning. The representation of these options as a system network is shown in Figure 2.4.

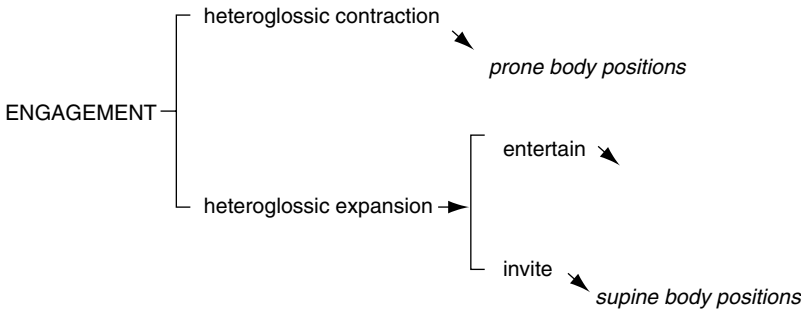


FIGURE 2.4 A system network for expanding and contracting space for negotiation

The frequency with which the teachers use these supine, prone and oscillating gestures varies from one stage of a lesson or pedagogic activity to another. The more frequent use of elicitation gestures with supine hand position characterizes phases of lesson in which teachers coordinate discussion. They function in this context to open up space for students to contribute. The extent to which individual teachers engage in dialogue with students is also no doubt a reflection of a more general pedagogic model (Bourne 2003). There is an urgent need for more research into the ways in which interpersonal epilinguistic body language functions in relation to teaching and learning in face-to-face classrooms, and in turn into the impact a lack of access to embodied meanings might have in computer-mediated online learning.

Metafunctional Fusions in Body Language

The instances of body language described above highlight the ways in which metafunctional meanings can be co-instantiated in both speech and body language, albeit in ways that commit meaning potential to a greater or lesser degree. It is also noted that **the metafunctional load can be distributed differently across modes**, so that meaning in relation to one metafunction may be instantiated in gesture but not the verbiage, and vice versa, and in any one instance of body language there may be fused different kinds of metafunctional meanings. Pointing gestures, for example, doing the work of identification, readily fuse with other gestural expressions of interpersonal meaning. A pointing gesture identifying a participant in the discourse can do so in the context of an elicitation with a supine hand position or in the context of a command with a prone-hand position. In Image 2.5b, for example, the gesture integrates a meaning of elicitation together with a meaning of identification of the intended interactant in the directionality of the fingers of the hand. Muscle

tension can function simultaneously with meanings of identification or even quantification adding a dimension of intensity of attitude, and so on.

Conclusion

The intention in this chapter has been to contribute to research on body language from a social semiotic perspective through an analysis of the ways in which teachers exploit the meaning potential of bodily postures and movements as they construct meaning in face-to-face classrooms. Building on work of others in this field (especially Martinec 2000, 2001, 2002, 2004, Munitgl 2004, and Cléirigh, in Martin, Chapter 12), the aim has been to explore further of the meaning potential of body language from a metafunctional perspective, focusing here on textual and interpersonal meaning. The data reveal ways in which these metafunctional meanings may be co-instantiated in both spoken language and body language or distributed across the different semiotic systems. It also reveals ways in which more than one metafunctional meaning may be infused in a single gesture.

The research has resulted in the development of some tentative and partial system networks to represent meaningful options in bodily expression in identification, graduation and engagement. Highlighted too is the role of the body in construing phasal shifts in the flow of discourse. From a pedagogic perspective, these embodied movements and syndromes of gestures function to guide students' attention, signalling shifts in what is salient for them in the teacher's talk. Evident too is the extent to which body language can cue students into the values attached to certain information, and can expand or contract perceived space for their participation in the discourse. While some teachers are more or less gestural in the enactment of their pedagogic practice, in each of the classrooms studied body language was intrinsic to the teacher's interaction with the students. **It contributes to building redundancy in meaning-making potential and to expanding the meaning potential available in the spoken discourse alone.** The teachers' body language is also a resource in mediating between potential and actual meanings and as such is an intrinsic part of the process of scaffolding students' learning.

Acknowledgements

I would like to thank Linda, Matt and Juliana who together with their students generously allowed me to film their classrooms, and Insearch Language Centre for its ongoing support for and cooperation in research at a time when more and more barriers are being constructed for research in classrooms in Australia. I also thank my research assistant, Catherine Baird, for her technical know-how

and insight. This research was undertaken with the assistance of a grant from the University of Technology, Sydney.

References

- Bednarek, M. & Martin J.R. (Eds) (2010). *New discourse on language: Functional perspectives on multimodality, identity, and affiliation*. London: Continuum.
- Bourne, J. (2003). Vertical discourse: The role of the teacher in the transmission and acquisition of decontextualised knowledge. *European Educational Research Journal*, 2(4), 496–521.
- Brady, N.C., McLean, J.E., McLean, L.K. & Johnston, S. (1995). Initiation and repair of intentional communication acts by adults with severe to profound cognitive disabilities. *Journal of Speech and Hearing Research*, 38, 1334–1348.
- Chafai, N.E., Pelachaud, C. & Pele, D. (2007). A case study of gesture expressivity. *Language resources and evaluation*, 41, 341–365.
- Christie, F. (1997). Curriculum macrogenres as forms of initiation into a culture. In F. Christie & J.R. Martin (Eds), *Genres and institutions: Social processes in the workplace and school* (pp. 134–160). London: Cassell.
- Efron, D. (1972). *Gesture, race, and culture*. The Hague, Netherlands: Mouton de Gruyter.
- Eisenstein, J. (2008). Gesture in automatic discourse processing. PhD Thesis, Massachusetts Institute of Technology.
- Enfield, N. 2009. *The anatomy of meaning: Speech, gesture, and composite utterances*. Cambridge: Cambridge University Press.
- Flewitt, R. (2006). Using video to investigate pre-school classroom interaction: Education research assumptions and methodological practices. *Visual Communication*, 5 (1), 25–50.
- Gregory, M. & Malcolm, K. (1995). Generic situation and discourse phase: An approach to the analysis of children's talk. In Jin Soon Cha (Ed.), *Before and towards communication linguistics: Essays by Michael Gregory and Associates* (pp. 154–195). Seoul, Korea: Sookmyung Women's University.
- Grice, H.P. (1975). Logic and conversation. In P. Cole & J.L. Morgan (Eds), *Syntax and semantics, Vol III, Speech acts* (pp. 41–58). New York: Seminar Press.
- Gullberg, M. (1999). Gestures in spatial descriptions. *Lund University Department of Linguistics Working Papers*, 47, 87–97.
- Halliday, M.A.K. (1963). The tone of English. *Archivum Linguisticum*, 15(1), 1–28. Republished in 2005 as Chapter 8: The tone of English. In J. Webster (ed.), *Collected Works of M.A.K. Halliday*, Volume 7. London: Continuum.
- Halliday, M.A.K. (1978). *Language as social semiotic: The social interpretation of language and meaning*. London: Edward Arnold.
- Halliday, M.A.K. (1991). The notion of context in language education. In T. Le & M. McCausland (Eds), *Language education: Interaction and development: Proceedings of the international conference* (pp. 1–26). Ho Chi Minh City, Vietnam.
- Halliday, M.A.K. (1992). The act of meaning. In J.E. Alatis (Ed.), *Georgetown Round Table on languages and linguistics: Language, communication and social meaning* (pp. 7–21). Washington, D.C.: Georgetown University Press. [Republished in

- Volume 3 in the *Collected Works of M.A.K. Halliday: On language and Linguistics*, Chapter 17, 375–389].
- Halliday, M.A.K. (1994). *Introduction to functional grammar*. London: Edward Arnold.
- Halliday, M.A.K. & Matthiessen, C.M.I.M. (1999). *Construing experience through meaning: A language-based approach to cognition*. London: Cassell.
- Hjelmslev, L. (1961). *Prolegomena to a theory of language*. Madison, WI: University of Wisconsin Press.
- Hood, S. (2004). Managing attitude in undergraduate academic writing: A focus on the introductions to research reports. In L. Ravelli & R. Ellis (Eds), *Analysing academic writing: Contextualised frameworks* (pp. 24–44). London: Continuum.
- Hood, S. (2006). The persuasive power of prosodies: Radiating values in academic writing. *Journal of English for Academic Purposes*, 15, 37–49.
- Hood, S. (2008). Summary writing in academic contexts: Implicating meaning in processes of change. *Linguistics and Education*, 19, 351–365.
- Hood, S. (2010). *Appraising research: Evaluation in academic writing*. London: Palgrave Macmillan.
- Jewitt, C. (2008). Multimodality and Literacy in School Classrooms. *Review of research in education*, 32, 241–267.
- Kendon, A. (1980). Gesticulation and speech: Two aspects of the process of utterance. In M.R. Key (Ed.), *The relationship of verbal and nonverbal communication* (pp. 207–227). The Hague, Netherlands: Mouton de Gruyter.
- Kendon, A. (2004). *Visible action as utterance*. Cambridge: Cambridge University Press.
- Kress, G., Jewitt, C., Ogborn, J. & Tsatsarelis, C. (2001). *Multimodal teaching and learning: The rhetorics of the science classroom*. London: Continuum.
- Kress, G. & van Leeuwen, T. (2006). *Reading images: The grammar of visual design* (2nd edn). London: Routledge.
- Levinson, S.C. (1983). *Pragmatics*. Cambridge: Cambridge University Press.
- Lund, K. (2007). The importance of gaze and gesture in interactive multimodal explanation. *Language resources and evaluation*, 41, 289–303.
- Macken-Horarik, M. (2004). Interacting with the multimodal text: Reflections on image and verbiage in ArtExpress. *Visual Communication*, 3 (1), 5–26.
- Martin, J.R. (1992). *English text: System and structure*. Amsterdam: John Benjamins.
- Martin, J.R. (2006). Genre, ideology and intertextuality: A systemic functional linguistic perspective. *Linguistics and the Human Sciences*, 2(2), 275–298, (Special Issue on Genre, J. Bateman (Ed.)).
- Martin, J.R. (2007). Multimodality – Some issues. Paper presented at Semiotic Margins: Reclaiming Meaning Conference, Dec 10–12, University of Sydney.
- Martin, J.R. (2010). Semantic variation: modelling system, text and affiliation in social semiosis. In M. Bednarek & J.R. Martin (Eds), *New discourse on language: Functional perspectives on multimodality, identity and affiliation* (pp. 1–34). London: Continuum.
- Martin, J.R. & Rose, D. (2007). *Working with discourse: Meaning beyond the clause* (2nd edn). London: Continuum.
- Martin, J.R. & Stenglin M. (2006). Materialising reconciliation: Negotiating difference in a trans-colonial exhibition. In T. Royce & W. Bowcher (Eds), *New directions in the analysis of multimodal discourse* (pp. 215–238). Mahwah, NJ: Lawrence Erlbaum Associates.

- Martin, J.R. & White, P.R.R. (2005). *The language of evaluation: Appraisal in English*. London: Palgrave Macmillan.
- Martinec, R. (2000). Rhythm in multimodal texts. *Leonardo*, 33(4), 289–297.
- Martinec, R. (2001). Interpersonal resources in action. *Semiotica* 135, 1(4), 117–145.
- Martinec, R. (2002). Rhythmic hierarchy in monologue and dialogue. *Functions of language*, 9(1), 39–59.
- Martinec, R. (2004). Gestures that co-occur with speech as a systematic resource: The realization of experiential meaning in indexes. *Social semiotics*, 14(2), 193–213.
- Martinec, R. & Salway, A. (2005). A system for image–text relations in new (and old) media. *Visual Communication*, 4(3), 337–371.
- McNeill, D. (1992). *Hand and mind: What gestures reveal about thought*. Chicago, IL and London: University of Chicago Press.
- McNeill, D. (1998). Speech and gesture integration. In J.M. Iveron & S. Goldin-Meadow (Eds), *The nature and functions of gesture in children's communication*, 79 (pp. 11–27). San Francisco, CA: Jossey-Bass publishers.
- McNeill, D. (Ed.) (2000). *Language and gesture: Window into thought and action*. Cambridge: Cambridge University Press.
- Muntigl, P. (2004). Modelling multiple semiotic systems: The case of gesture and speech. In E. Ventola, C. Charles & M. Kaltenbacher (Eds), *Perspectives on multimodality* (pp. 31–50). Amsterdam: John Benjamins.
- Painter, C. (2007). Childrens' picture book narratives: Reading sequences of images. In A. McCabe, M. O'Donnell & R. Whitakker (Eds), *Advances in language and communication* (pp. 40–59). London: Continuum.
- Painter, C. & Martin, J.R. (In press). Intermodal complementarity: Modelling affordances across verbiage and image in children's picture books. *Ilha do Desterro: A Journal of English Language, Literatures in English and Cultural Studies* (Special Issue on Multimodality).
- Peirce, C.S. (1955). *Philosophical writings of Peirce*. New York: Dover Publications.
- Royce, T.D. & Bowcher, W.L. (Eds) (2007). *New directions in the analysis of multimodal discourse*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Stenglin, M. (2004). Packaging curiosities: Towards a grammar of three-dimensional space. Unpublished PhD Thesis, Department of Linguistics, University of Sydney.
- Stenglin, M. (2007). Making art accessible: Opening up a whole new world. *Visual Communication*, 6(2), 202–213.
- Ventola, E., Charles, C., & Kaltenbacher, M. (Eds) (2004). *Perspectives on multimodality*. Amsterdam: John Benjamins.
- van Leeuwen, T. (1999). *Speech, music, sound*. London: Macmillan.
- van Leeuwen, T. (2006). Towards a semiotics of typography. *Information Design Journal*, 14(2), 139–155.

Chapter 3

Grappling with a Non-speech Language: Describing and Theorizing the Non-verbal Multimodal Communication of a Child with an Intellectual Disability

Shoshana Dreyfus
University of Sydney

Introduction

This volume is indicative of an emerging breadth and depth of research into the diverse modes of communication that exist outside or alongside spoken and written language. While there has been, for example, a considerable focus on non-verbal modes of communication, such as gesture (see, for example, McNeill 1992, Kendon 1981, 2004), this focus is most often on the non-verbal communication of speakers¹. This chapter, however, focuses on the non-verbal modes of communication of someone who doesn't speak and who has a severe intellectual disability. People with intellectual disabilities and communication disorders exist in a particularly marginal space in society, relying on the beneficence of others for all their needs for the duration of their lives.

Focusing on the communication of this group of people marks a break, as it were, in the gaze of Systemic Functional theory in the sense that the theory was developed on/from the language and communication of people, be they infants or adults, with full intellectual capacity, who all either developed or already had the complete language system. While Halliday (see, for example, 1978) has always argued that SFL is a theory of semiosis, or meaning-making in general, the theory is most often applied to normative productions of language. Even when the theory has not been restricted to the description of language, and has been used to theorize other non-linguistic modes of communication such as art (O'Toole 1994), movement (McInnes 1998), space (Stenglin 2004 and Chapter 4) and gesture (Martinec 2000, 2004, Hood, Chapter 2), these are again the productions of humans with full intellectual capacity. Further, while the theory has also been used to describe and analyse the language of people with speech disorders (see, for example, Armstrong 1991, 1992, 1993, 2001, Togher 1998, 1999, 2000, 2001, Fine 1994, 2004), all the subjects of these studies do speak, even if it is in a 'disordered' manner.² Thus, within the field

of SFL, there is no research, other than my own (see Dreyfus 2007, 2008) that focuses on the communication of people who do not use speech as their main form of communication and who have intellectual disabilities. My work therefore begins, in a very tentative way, to expand into this area.

The purpose of this chapter is to highlight the complexities of both studying non-verbal multimodal communication in general and using SFL theory for this kind of study. The first section of the chapter examines the nature of the communication environment for a non-verbal multimodal communicator with an intellectual disability. The second section is a brief description of my study which examined this kind of communication. The third section expands on the issues arising from using a normative theory for the study.

The Communication Environment of Non-verbal Multimodal Communicators with an Intellectual Disability

The communication environment for intellectually disabled non-verbal multimodal communicators is different from the speaking environment in many ways. This is also the case for users of sign languages, such as deaf people and a comparison of these two types of language in relation to spoken language is useful. For deaf signers, Johnston (1996) characterizes this different communication environment as the 'semiotic umwelt' – noting that there is a symbiotic evolution of the (communication) environment and the organism. That is to say, the language and the environment evolve together. Similarly, for intellectually disabled non-verbal multimodal communicators, the communication environment evolves with the person. Thus, in this discussion, this communication environment will be compared to and contrasted with both the normative language environment and the sign language environment, where applicable.

Similar to sign languages, the non-verbal multimodal communication is a face-to-face system of communication using the visual-gestural medium (Johnston 1996). This use of gesture and the visual in space requires participants in the interactions to be co-present in the immediate environment; for example, the communication partner cannot see what a multimodal communicator is pointing at if they are not within the same physical space. Additionally, if being gestural includes the use of the hands (where possible), participation in communicative activity generally involves the cessation of most other non-linguistic activity. Sign languages are considered an oral tradition as there is no written form. The non-verbal multimodal communication under the microscope in this study also has no written form; however, as it is embedded within a spoken English milieu and conducted with speaking communication partners, it is based on a tradition that has its own written form.

Although embedded within a speaking community, similar to deaf signers, intellectually disabled non-verbal multimodal communicators live in a world quite separate from the speakers around them. While deaf signers at least have a community of other signers with whom to communicate in a shared language,

non-verbal multimodal communicators do not have their own communication community and interact mostly with speakers. This can be partly attributed to the idiosyncratic nature of their communication, which can be difficult to understand, and to the fact that as a result of the idiosyncrasies, they often need an informed communication partner and speaker as interpreter for their multimodally communicated meanings (Dreyfus 2007). There can sometimes be insurmountable difficulties in getting the simplest meanings across without speech, particularly in encounters with people who are not familiar with these idiosyncrasies. Therefore, intellectually disabled non-verbal multimodal communicators do not have a common language; they have an idiolect, communicating in their individual combinations of multimodes, while their communication partners typically communicate using speech. This communication environment is transmodal (Dreyfus 2007), where the meanings must traverse both non-verbal multimodal communication and speech.

In addition to being embedded (or isolated) within the larger speaking community, non-verbal multimodal communicators are also heavily dependent on speakers for their communication needs. This is not only for interpretation and mediation with uninformed or new communication partners, but also if they are to have multimodal resources that are external to their body, such as pictograph communication systems or electronic devices; it is incumbent upon their carers to provide them. This creates a power imbalance in the communication environment.

For both non-verbal multimodal communication and sign languages there is a blurring of the boundaries between semiotic and non-semiotic behaviour. For sign languages, the primarily temporal nature of the communication impacts on what constitutes semiotic behaviour (Johnston 1996). There is an absence of sound, but a presence of space, which fundamentally changes the nature of semiosis. For non-verbal multimodal communication, there is not necessarily an absence of sound, but an absence of verbal sounds, and there is the use of space, as modes such as gestures and actions come into play. Unlike the spoken language environment, where speech dominates, with non-verbal multimodal communication, all kinds of non-verbal behaviour can be seen to be semiotic, meaning semiotic behaviour is both different and less easily defined. Therefore, both the communication partner and the analyst need to broaden their view of what constitutes semiotic behaviour, and in many cases, be vigilant in watching out for it, particularly as research shows that many attempts at communication by non-verbal communicators are missed or misunderstood (Dreyfus 2007).

The use of space, rather than sound, also impacts on the way meanings occur. In English, meanings unfold syntagmatically in time (for speech) or on the page (for writing), trackable through an examination of the system of Theme. In other words, there is an order to the unfolding of meanings. If the order is disrupted, the meanings may not make sense. In face-to-face languages, the sequencing of meanings occurs differently as the use of space allows many meanings that would typically be expressed in more than one word to be expressed

altogether in one sign (Johnston 1996, Dreyfus 2007). For example, in the case of the boy reported on in this chapter, grabbing the driver's sleeve when travelling in the car, means 'Where are we going?' while flicking the door handle means 'Can I get out?'

In summary, the semiotic umwelt of the non-verbal multimodal communication of a person with an intellectual disability encompasses:

- A here-and-now context that includes the use of space, altering the syntagmatic relationships (the ways the meanings are ordered) in spoken language;
- a broader understanding of what constitutes semiotic behaviour; and
- a transmodal communication environment, where meanings have to traverse two kinds of expression: speech and multimodes;

The chapter will now turn to a closer examination of the study itself.

The Study

The study reported on in this chapter focused on the communication of a boy named Bodhi who has a severe intellectual disability and does not communicate using speech. Nevertheless, he can understand speech, is very communicative and uses a variety of other modes of expression to communicate his meanings. The study analysed these modes of expression and the meanings Bodhi made with them. The overall challenge was how to use a theory that has lexicogrammar as its way in when there is no lexicogrammar to study.

The study first proceeded to map the modes of expression Bodhi used to communicate his meanings (see Figure 3.1).

Modes of Expression

Bodhi's modes of expression have been clustered into six broad types based on the categories from another study of the communication of non-verbal children with disabilities (see Light et al. 1985). These modes are vocalizations, gestures, materials, actions, behaviours and eye gaze. However, the modes are invariably not communicated separately; that is to say, Bodhi generally does not communicate using only one mode at one time, but combines different modes in the one meaning-making move. For example, he may vocalize while pointing to something and looking at the communication partner. Each of the different modes carries different components of meaning. Together, these modes of expression give Bodhi a limited but nevertheless somewhat functional system of communication, with an informed or interested communication partner (Dreyfus 2007).

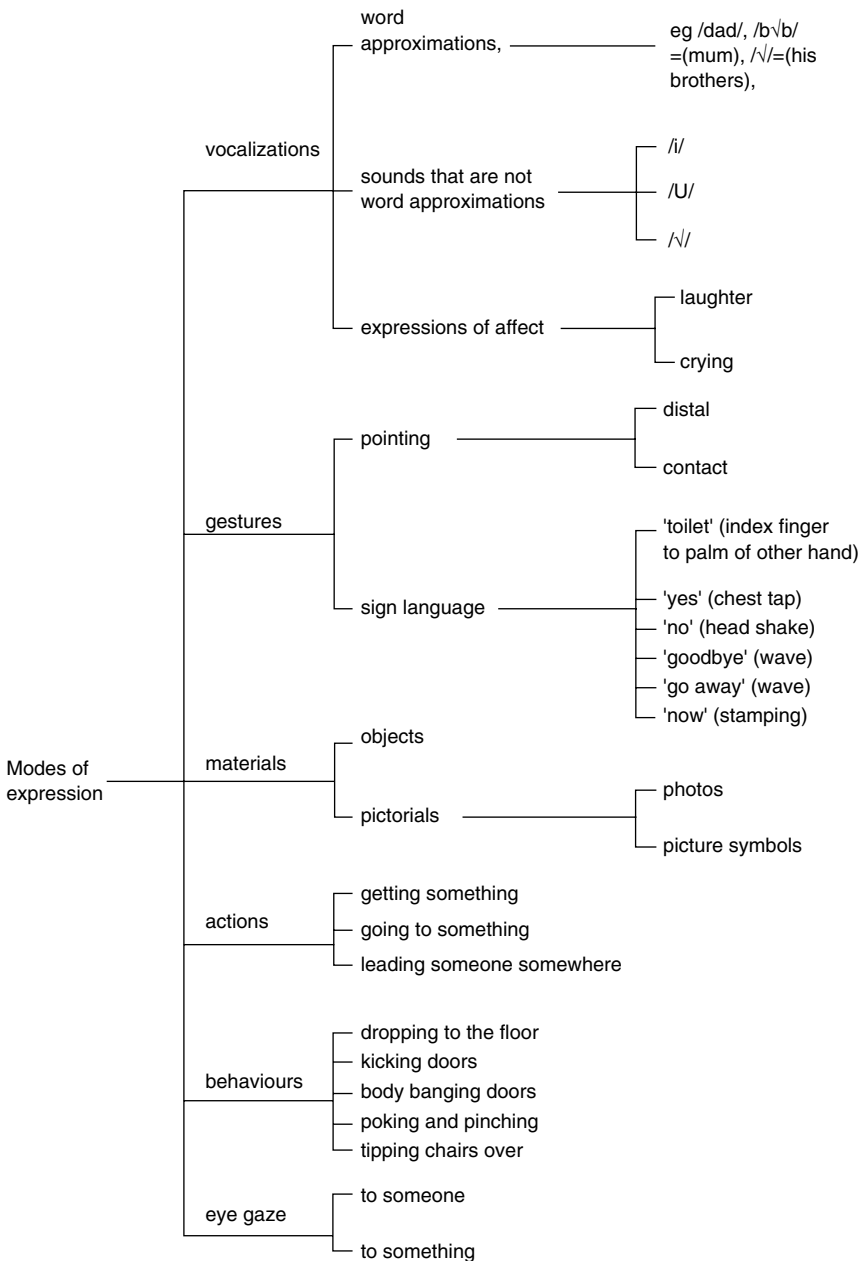


FIGURE 3.1 Taxonomy of Bodhi's modes of expression

In order to understand Bodhi's meanings, a number of research questions were posed that reflected the trinocular perspective articulated by Halliday (1996). Bodhi's communication was explored from above, taking into account the function and meaning of Bodhi's moves; it was explored from below, taking into account the modes of expression that he uses to express his meanings; and was explored from round about, taking into account the internal organization of his expressed meanings. The questions are as follows:

1. What kinds of meanings does Bodhi make?
2. What resources does he use to make those meanings?
3. What kinds of discourse roles does he take up?
4. Is he drawing on the same kind of language system that speakers use, and if not, what does his system look like?

While the analyses conducted in order to answer the questions are by no means exhaustive, they were still able to capture Bodhi's meaning-making abilities with different communication partners. In brief, the study showed that Bodhi is a very social being, and an active communicator who initiated more than half the conversation exchanges in the data. However, his meaning-making abilities are restricted with a limited number of experiential meanings and very few textual meanings expressed. In terms of the semiotic space Bodhi occupies, it is predominantly one of the concrete world, and not of the abstract world. With regard to textual meanings, Bodhi only communicates peaks of salience in his moves, which are constituted by new information. The rest of the meaning must be co-constructed by the communication partner using sources outside the text. Interpersonally, Bodhi deploys three of the four basic speech functions, although their expression is undifferentiated and the communication partner must work together with Bodhi to determine which speech function he means. To paraphrase Halliday (1975), Bodhi does not have an open-ended system with massive potential that 'can create indefinitely many meanings and indefinitely many sentences and clauses and phrases and words for the expression of those meanings' (pp. 35–36). The system is limited by both his lack of a lexicogrammar and his severe intellectual disability, but the way these two interact is hard to unravel and was not the task of the study. However, as the study showed, not having a lexicogrammar does not mean that Bodhi cannot make meaning. Indeed, he uses his multimodes to make a variety of meanings, albeit limited, in a process of joint construction with his communication partners. For a detailed description of the findings of the study, see Dreyfus (2007).

Issues Arising from the Study

As mentioned above, one of the issues encountered in my study was how to distinguish between semiotic and non-semiotic behaviour. In the field of

Augmentative and Alternative Communication (AAC), which is the main field that both studies and seeks to improve the communication of people with communication disorders, there is the view that when it comes to non-verbal communication, *all* behaviour can be communicative (Mirenda 1997). This contrasts with the SFL view as put forward first by Halliday (1985), which divides behaviour into symbolic and non-symbolic acts. This view is extended by Cléirigh (in preparation; see also Martin, Chapter 12) who divides behaviour into non-semiotic, non-linguistic semiotic and linguistic semiotic, which is inclusive of non-human and non-normative communication systems.

In contrast to the AAC view, as Cléirigh points out, and as evidenced by this study, not all behaviour is communication. There is a difference between behaviour that enacts social relationships and behaviour that does not. In terms of this study, there was, at times, an issue for both me as analyst, and also for communication partners, of how to decide whether what Bodhi was doing at any given moment did or did not constitute semiotic behaviour. In some cases this was quite clear-cut, but in others not so. With spoken and written language, determining when someone is being communicative or not is generally not difficult – the fact that someone is speaking typically indicates a communication act is taking place (other than talking to oneself, which, however, could also be deemed communicative). However, with regard to people who do not speak, determining the boundaries between non-semiotic behaviour (i.e. non-linguistic behaviour that is *not* communicative) and non-linguistic semiotic behaviour (non-linguistic behaviour that *is* communicative) can be difficult.

Therefore, following Cléirigh (in preparation; see also Martin Chapter 12) criteria were developed to determine the boundaries between semiotic and non-semiotic behaviour (see Dreyfus 2007). These were as follows:

1. The presence of certain markers that indicated semiosis, such as particular sounds and tones, pointing or eye gaze.
2. Whether the move was seen to be doing some communicative work and constituted a move in conversation; that is, whether the communication partner responded as if Bodhi was communicating something.

Bodhi's laughter is a prime example of this. That is to say, when Bodhi laughs, while he may not intentionally be communicating that he is happy, the communication partner responds to his laughter as a move communicating positive affect, therefore it constituted a move in the conversation, and was treated as such. The excerpt in Example 1 illustrates this:

- Mark: . . . Tomorrow Dad will take you to Saturplay and Bruce will come in the car too. We'll have Bruce in the car, with a banjo and a double bass. Yes . . .
- Bodhi:³ /l ye˘/ (smile)
- Mark: Yeah Bruce

- Bodhi: /1 he he´/ (laughish)
 Mark: You like Bruce?
 Bodhi: /1? hi-ye´/ (laughish)
 Mark: I think you do
 Bodhi: /h´h-h-h/ (laughs)
 Mark: Tomorrow. In the morning
 Bodhi: /´-e A/ (high pitched)
 Mark: yeah

Example 1

But it is not just *whether* Bodhi's behaviours are semiotic or not that is important; once it is decided that a behaviour is semiotic, there then needs to be some classification of that behaviour in terms of its semiosis. In other words, there can be degrees of semiosis, where some semiotic behaviours are more abstractly semiotic than others. That is to say, some of Bodhi's moves are strataally simpler, being much like a child's protolanguage in that they are simple content expression pairs where there is no intermediate (abstract) stratum of meaning (i.e. a lexicogrammar). These include modes such as behaviour and actions. For example, when he lies on the floor and kicks the door to communicate that he wants to go out. However, there are also moves that display more stratal complexity, having some kind of abstract layer sandwiched between the semantic and expression planes, such as signs and pictorials/pictographs. An example of this is when Bodhi makes the sign for 'toilet' (which is the index finger of one hand contact pointing the palm of the other hand) to communicate he wants to go or is going to go to the toilet.

Straddling the AAC view that all behaviour is communication and the SFL view that behaviour can be divided into semiotic and non-semiotic behaviour raises questions for the study of non-verbal multimodal communication, such as what constitutes a language or communication system, what the boundaries of these are, and how fixed or how flexible they can be. Do we classify these in developmental terms such as protolanguage or adult language? And does this kind of classification help with understanding and describing them? (See Dreyfus 2007 for a more detailed discussion of the issues associated with the classification of Bodhi's communication system.)

In order to capture this variance in type or degree of semiosis, a cline is posited to capture the varying degrees of semiosis, along the lines of both Halliday (1985) and Cléirigh (in preparation; see also Martin Chapter 12) (see Figure 3.2). At one end of the cline are the modes of expression that are content expression pairs, and more like the primary semiotic of protolanguage. At the opposite end are the most symbolic or higher order modes, which, of course, means the most linguistic modes.

In order to classify and analyse these different types of semiotic behaviour, their function needs to be determined. While within verbal language it is the

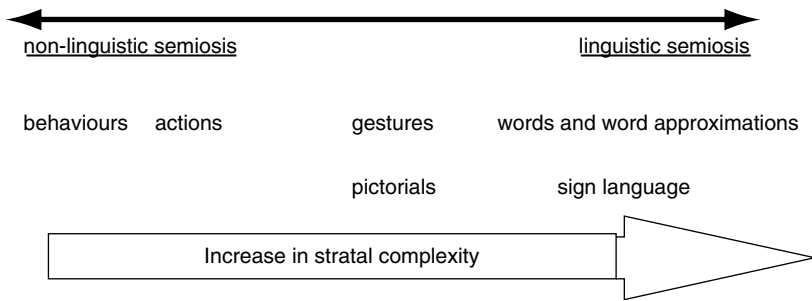


FIGURE 3.2 Cline of semiosis in Bodhi's communication

lexicogrammar that supplies the various (meta)functions, in non-verbal multi-modal communication it can be difficult to determine a function from within the move itself, because the move itself does not contain enough information. Further, it is in conjunction with the communication partner that Bodhi jointly constructs his meanings. Meaning is often constructed *across* the turns rather than *within* the turns. This can be seen in the exchange in Example 2 where Bodhi tries to tell his grandmother something about his breakfast bowl. At first she does not understand him, but as he replays his move, together they work towards a correct verbal interpretation/articulation of his move:

- | | | |
|--------|--------|--|
| Move 1 | Bodhi: | /5 i / (contact pointing ⁴ the bowl) |
| Move 2 | Dodo: | That's a lovely bowl, isn't it? |
| Move 3 | Bodhi: | /2 i / (contact pointing the bowl) |
| Move 4 | Dodo: | That's your bowl, yes. |
| Move 5 | Bodhi: | /2 i hi-hi /2 i hi-hi /2 i-hi /2 i / (contact pointing bowl) |
| Move 6 | Dodo: | Yes. D'you like that bowl? |
| Move 7 | Bodhi: | /ˈhˈh / (giggle) |

Example 2

As researchers within the field of AAC articulate, any study of non-verbal multi-modal communication needs to be able to take into account the contributions of the communication partner. Thus, a within-clause perspective needs to be supplemented by a beyond-clause or discourse semantic perspective. Exchange Structure Analysis (after Coulthard & Montgomery 1981, Berry 1981a, 1981b, 1981c, Martin 1992, Ventola 1987, 1988) offers this perspective and was therefore used to determine the function of certain moves by looking at the move in its dialogistic context. It is the subsequent communication partner's moves that can provide clues to the function of the multimodal move when the function of the move cannot be determined from within the move itself.

This means the function of the non-verbal multimodal move can be determined retrospectively. In the exchange above, it is not until the final move, where Bodhi giggles rather than replays, showing his satisfaction with Dodo’s response, that we are able to understand that Bodhi’s initial move meant he likes the bowl.

The discourse semantic (above clause) perspective was used in combination with a within-turn or clause perspective, in order to attempt to capture the instantiated meanings within each of Bodhi’s moves. All Bodhi’s moves were plotted onto a table that could reflect the metafunctional perspective. As shown in Table 3.1, there is a move of Bodhi’s that comes from an exchange where Bodhi is travelling in the car with his father. They are on their way to the chemist and Bodhi, who had an obsession with flushing toilets, asks if there are toilets there (see Example 3):

Bodhi: /2 i /2 i /2 i / (signs ‘toilet’)
Mark: No, there won’t be toilets. There’s no toilets. No toilet at the chemist and there’s no toilets at the fish shop

Example 3

The top left-hand corner with the B (standing for ‘Bodhi’), records the number of the turn in terms of where it is located in the transcript. The column below that lists all the possible modes of expression. The next column, the instance

Table 3.1 Instance table

70B	INSTANCE	SYSTEM		
Mode of expression	Realization	Possible meanings		
		Experiential	Interpersonal	Textual
Sounds	i i i	Existent or Goal	Demand info – polar question; or demand service	New
Tone	2 2 2			
Gestures	signs ‘toilet’			
Materials				
Actions				
Behaviours				
Facial expression				
Eye gaze	To Mark		Indicating who Bodhi is talking to	
GLOSS	Is there a toilet there? Or can I flush the toilet there?			

column, records which modes Bodhi actually uses in the instance of communication that is his turn. The following three columns are for recording the metafunctional meanings associated with the modes of expression. The bottom line 'gloss' refers to my interpretation of what Bodhi has tried to communicate. For those of Bodhi's moves that were able to be analysed for metafunctional meanings, the columns were filled in. For those that weren't, the columns were left empty.

Issues of the Gloss

The gloss is an important part in the deciphering of non-verbal moves, as we cannot make sense of Bodhi's moves unless we gloss them in our own terms. With regard to protolanguage, Halliday (1975) has said that everything is interpreted in terms of our own semantic system. Further, when discussing Auslan (Australian Sign Language), Johnston (1991) states that the gloss is a way of capturing the semantic content of a mode of communication, even if imperfect. However, Johnston (1991) also warns that:

no matter how frequently one may remind the reader that the gloss is no substitute for the sign, if there is nothing in the text that represents the sign *per se* (be it picture or script) the glossing may take on a life of its own. (p. 6)

If we consider both Bodhi's move and the gloss, it can be seen that the sounds and tone express some interpersonal content (some kind of demand); the sign 'toilet' expresses some experiential content; and together, they express a textual component that can be called New. One of the findings of this study was that Bodhi typically only expresses the New. Everything else must be gleaned from the context, where the communication partner must work together with Bodhi to jointly construct the meaning. However, having said that, for many moves I was not able to fill this table in because there was not enough information from Bodhi's move itself or from the move in its dialogistic context.

Limitations of the Networks: Expansion to the Speech Function Network

An additional issue resulting from this study concerns how Halliday's (1994) system of speech function was not able to accurately and delicately capture what was occurring in one of Bodhi's most important moves. The speech function network offers two possibilities of commodity that can be exchanged: information and goods-&-services. This study showed how the commodity of

goods-&-services needs to be expanded in order to capture the idiosyncracies of Bodhi's meaning-making. This is demonstrated in Example 2, where Bodhi is having breakfast with his grandmother, and tries to tell her something about the bowl he is eating out of by pointing at it while saying /i/ with a rising tone. In this exchange, Bodhi is doing two things simultaneously: first, he is giving information, making a comment about the bowl to his grandmother; and secondly, he is asking that she provide a particular response, evidenced by the fact that he replays his move after each of her responses until he gets the response he wants, at which point he shows his satisfaction by giggling, which completes the exchange. The giving information part is relatively straightforward; however, this does not capture all that is going on. What the data has shown is that where Bodhi initiates the giving of information, he simultaneously demands that the communication partner articulate that information back to him in words. It is as if he is saying, 'Tell me in words what I have just communicated to you multimodally'. This requires the communication partner to do two things: first, to understand the multimodal move; and secondly, to articulate what has been communicated multimodally in words.

It can be said, therefore, that Bodhi's initiating move realizes two speech functions simultaneously: giving information and demanding goods-&-services. However, the kind of goods-&-services he is demanding is a very particular type of service: it is a linguistic service rather than an action service, such as requesting someone get him a drink. Finer distinctions of types of service have been addressed by Ventola (1987) in her work on service encounters. In service encounters, Ventola (1987) shows how customers ask a particular type of question that is different from a demand for information, and is actually the demand for a linguistic service of the provision of information. She provides recognition criteria for these types of moves showing their difference from a straight demand for information. However, Bodhi's demand for a linguistic service is not the same as that identified by Ventola (1987). In Ventola's examples, the demander is asking for the provision of information they do not have. In Bodhi's case, he is asking for the articulation of information that he has just provided. In other words, the kind of linguistic service he is demanding is that of articulation. In order to be able to capture this, it is necessary to expand the speech function network as in Figure 3.3.

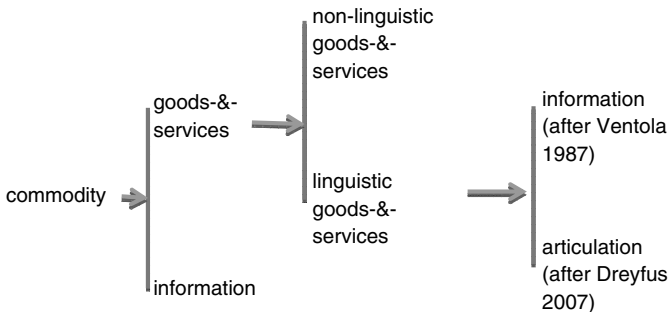


FIGURE 3.3 Revised commodity branch of speech function

Distinguishing between different types of services available for Bodhi then gives rise to a different set of options within the system network of move in dialogue (see Figure 3.4).

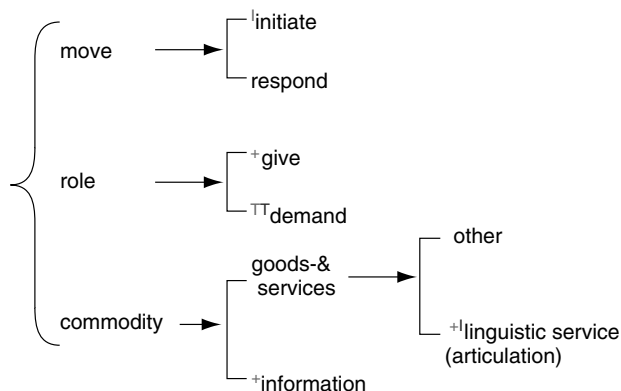


FIGURE 3.4 Bodhi's system of speech functions (after Halliday & Matthiessen 2004:108)

Key:

I = if, T = then

The case '^Iinitiate ⁺give ⁺information, ^Tdemand ⁺linguistic service (articulation), means, if Bodhi initiates with giving information, then he also demands the linguistic service of articulation of that information.

In the case of '⁺linguistic service (articulation) ^Tdemand, means if Bodhi expresses a linguistic service, it is always as a demand.

As the study used Exchange Structure to examine the kind of discourse roles Bodhi was able to take up, and the speech function moves correspond to conversational moves within an exchange, adapting the speech function moves also means it was necessary to expand the range of synoptic moves possible within the exchange.

To explicate, when Bodhi initiates an exchange to give information, taking up the Primary Knower role, he is simultaneously demanding an action, which is the Secondary Actor role. Further, the action is a particular type of action, the linguistic service of articulation. As each synoptic move in an exchange has a particular notation, and a move can only be one sort of move, that is, an information move (DK1, K1 or K2) or an action move (DA1, A1 or A2), the coding notation for moves needed to be altered to reflect the type of move Bodhi makes. Bodhi's initiating move becomes K1/A2:LS:A. K1 means Primary Knower; A2 refers to Secondary Actor; LS refers to linguistic service; while the final 'A' in the sequence refers to articulation. The response move of the communication partner is also expanded to include their response of giving Bodhi information while simultaneously providing the linguistic service of articulation. This is notated as follows: K2f/A1:LS:A. K2f refers to Secondary

Knower response or follow-up move; A1 refers to the Primary Actor (providing the service); LS: A also referring to the linguistic service of articulation. Therefore, the above exchange is notated as follows:

K1/A2:LS:A	BODHI: /2 i / (contact pointing the bowl)
K2f/A1:LS:A	DODO: that's a lovely bowl isn't it?
ch/rp	BODHI: /2 i / (contact pointing the bowl)
rrp	DODO: that's your bowl. yes
ch/rp	BODHI: / 2 i hi hi /2 i hi hi /2i hi /2 i /
K2f/A1:LS:A	DODO: yes, d'you like that bowl?
K1/A2:LS:Af	BODHI: /ˈhˈh / (giggly sound)

(ch/rp is challenge and replay; rrp is response to replay)

Conclusion

This chapter has attempted to highlight some of the issues arising from the study of the non-verbal multimodal communication of a child with an intellectual disability using systemic functional linguistic theory. The chapter explains how the differing nature of this kind of communication gives rise to a different communication environment – a transmodal environment where meaning is jointly negotiated and a variety of semiotic behaviour is brought into focus. The chapter also discusses how the boundaries between semiotic and non-semiotic behaviour are at times difficult to determine. We are at the edges of the theory in terms of being able to accurately describe this kind of communication using current networks. As such, an expansion of the speech function network is posited to capture the move that provides information while simultaneously demanding a linguistic service from the communication partner. Expanding the range of moves possible also has ramifications for the types of moves possible within the Exchange Structure model.

Notes

¹ This is not to say that there is no research into other sorts of non-verbal multimodal communication that uses systemic functional linguistic theory. Other such research includes studies of the communication systems of the primate species *Pan paniscus* in Benson and Greaves (2005) and Knight (2006 and Chapter 1).

² Trevor Johnston's (1991, 1992 and 1996) work applying SFL theory to Auslan does focus on a different kind of language – the non-verbal language of the deaf; however, this is again an intellectually able population, and as Johnston has noted, while sign languages have the added dimension of the visual-gestural medium, they are comparable to spoken languages in that they are seen to be tristratal languages, even if they are more iconic than spoken languages.

- ³ Bodhi's vocalizations are written phonetically. The numbers in front of Bodhi's vocalizations reflect Halliday's (1994) descriptions of the tones in spoken English.
- ⁴ Contact pointing refers to touching things when pointing to them (what we might call tapping). This is contrasted with distal pointing, which refers to pointing to things without touching them – these things are usually more than 6 inches away (Brady et al. 1995).

References

- Armstrong, E. (1991). The potential of cohesion analysis in the analysis and treatment of aphasic discourse. *Clinical Linguistics and Phonetics*, 5(1), 39–51.
- Armstrong, E. (1992). Clause complex relations in aphasic discourse: A longitudinal case study. *Journal of Neurolinguistics* 7(4), 261–275.
- Armstrong, E. (1993). Aphasia rehabilitation: A sociolinguistic perspective. In M.M. Forbes (Ed.), *Aphasia treatment: World perspectives* (pp. 263–290). San Diego, CA: Singular.
- Armstrong, E. (2001). Connecting lexical patterns of verb usage with discourse meanings in aphasia. *Aphasiology*, 15(10/11), 1029–1045.
- Benson, J.D. & Greaves, W.S. (Eds) (2005). *Functional dimensions of ape-human discourse*. London and Oakville: Equinox.
- Berry, M. (1981a). Systemic linguistics and discourse analysis: A multi-layered approach to exchange structure. In M. Coulthard and M. Montgomery (Eds), *Studies in discourse analysis*. London: Routledge and Kegan Paul.
- Berry, M. (1981b). Polarity, ellipticity and propositional development, their relevance to the well-formedness of an exchange. *Nottingham Linguistic Circular*, 10(1), 36–63.
- Berry, M. (1981c). Towards layers of exchange structures for directive exchanges. *Network*, 2, 22–32.
- Brady, N.C., McLean, J.E. & Mclean, L.K. (1995). Initiation and repair of intentional communication acts by adults with severe to profound cognitive disabilities. *Journal of Speech and Hearing Research*, 38, 1334–1348.
- Cléirigh, C. (in preparation) *The Life of Meaning*.
- Coulthard, M. & Montgomery, M. (1981). *Studies in discourse analysis*. London: Routledge and Kegan Paul.
- Dreyfus, S. (2007). When there is no speech: A case study of the nonverbal multi-modal communication of a child with an intellectual disability. Unpublished doctoral study, University of Wollongong.
- Dreyfus, S. (2008). A systemic functional approach to misunderstandings. *Bridging Discourses*. Online Proceedings of the Australian Systemic Functional Linguistics Association conference, University of Wollongong.
- Fine, J. (1994). *How language works: Cohesion in normal and nonstandard communication*. Norwood, NJ: Ablex Publishing Company.
- Fine, J. (2004). *Language in psychiatry: A handbook of clinical practice*. London: Equinox.
- Halliday, M.A.K. (1975). *Learning how to mean*. London: Edward Arnold.
- Halliday, M.A.K. (1978). *Language as social semiotic*. London: Edward Arnold.
- Halliday, M.A.K. (1985). *Spoken and written language*. Geelong: Deakin University Press.

- Halliday, M.A.K. (1994). *An introduction to functional grammar* (2nd edn). London: Edward Arnold.
- Halliday, M.A.K. (1996). On grammar and grammatics. In R. Hasan, C. Cloran & D. Butt (Eds), *Functional descriptions: Theory in practice* (p. 121). Amsterdam: John Benjamins.
- Halliday, M.A.K. & Matthiessen, C.M.I.M. (2004). *An introduction to functional grammar* (3rd edn). London: Edward Arnold.
- Johnston, T. (1991). Transcription and glossing of sign language texts: Examples from (Australian Sign Language). *International Journal of Sign Linguistics*, 2(1), 3–28.
- Johnston, T. (1992). The realization of the linguistic metafunctions in a sign language. *Social Semiotics*, 2(1), 1–43.
- Johnston, T. (1996). Function and medium in the forms of linguistic expression found in a sign language. In W.H. Edmonson and R.B. Wilbur (Eds), *International Review of Sign linguistics* 1(pp. 57–94). New Jersey: Lawrence Erlbaum Associates.
- Kendon, A. (1981). Introduction: Current issues in the study of ‘non-verbal communication’. In A. Kendon (Ed.), *Nonverbal communication, interaction, and gesture*. The Hague: Mouton de Gruyter.
- Kendon, A. (2004). *Gesture: Visible action as utterance*. Cambridge: Cambridge University Press.
- Knight, N.K. (2006). Appraisal in Bonobo-Human Culture: Negotiating social behavioural parameters through evaluation with bonobo apes. *Linguistics and the Human Sciences*, 2(3), 355–376.
- Light, J., Collier, B. & Parnes, P. (1985). Communication interaction between young non-speaking physically disabled children and their primary caregivers. Part 3 – Modes of communication. *AAC Augmentative and Alternative Communication* 1, 125–133.
- Martin, J.R. (1992). *English text: System and structure*. Amsterdam: John Benjamins.
- Martinec, R. (1998). Cohesion in action. *Semiotica*, 1/2, 161–180.
- Martinec, R. (2000). Types of process in action. *Semiotica*, 130(3/4), 243–268.
- McInnes, D. (1998). Attending to the instance: Towards a systemic based dynamic and responsive analysis of composite performance text. PhD Thesis, University of Sydney, Sydney.
- McNeill, D. (1992). *Hand and mind: What gestures reveal about thought*. Chicago, IL: University of Chicago Press.
- Mirenda, P. (1997). Supporting individuals with challenging behavior through functional communication training and AAC: Research review. *AAC Augmentative and Alternative Communication*, 13, 207–225.
- O’Toole, M. (1994). *The language of displayed art*. London: Leicester University Press.
- Stenglin, M. (2004). Packaging curiosities: Towards a grammar of three-dimensional space. PhD Thesis, Linguistics, University of Sydney, Sydney.
- Togher, L. (1998). Interpersonal communication skills in the traumatic brain injury population: An analysis across situations. School of Communication Sciences and Disorders, Faculty of Health Sciences. PhD Thesis, University of Sydney, Sydney.
- Togher, L. (2000). Giving information: The importance of context on communicative opportunity for people with traumatic brain injury. *Aphasiology*, 14(4), 365–390.

- Togher, L. (2001). Discourse sampling in the 21st century. *Journal of Communication Disorders*, 34(1–2), 131–150.
- Togher, L., Hand, L. & Code, C. (1999). Exchanges of information in the talk of people with traumatic brain injury. In S. McDonald, L. Togher & C. Code (Eds), *Communication disorders following traumatic brain injury* (pp. 113–145). Sussex: Psychology Press.
- Ventola, E. (1987). The structure of social interaction: A systemic approach to the semiotics of service encounters. London: Frances Pinter.
- Ventola, E. (1988). The logical relations in exchanges. In J.D. Benson & W.S. Greaves (Eds), *Systemic functional approaches to discourse: Selected papers from the 12th international systemic workshop XXVI* (pp. 51–72). Norwood, NJ: Ablex Publishing Company.

This page intentionally left blank

Part Two

Evolving Accounts of Space and Music

This page intentionally left blank

Chapter 4

Spaced Out: An Evolving Cartography of a Visceral Semiotic

Maree Stenglin
University of Sydney

Introduction

Space provides the setting in which we conduct all the activities that are part of our ongoing lives. These activities constitute our evolving ontogenesis and include working, learning, shopping, eating and resting. In western cultures, many of these activities occur in a built environment:

[. . .] everyone lives and works in buildings of some sort. We all react instinctively to the size and space relationships of architecture because buildings and their surroundings relate to the human figure as we walk in, around and through them. (Lumley 1992:v)

Because of its omnipresent nature, space is a modality of phenomenal importance – one that is also deeply visceral. Visceral because spaces can evoke powerful bio-chemical responses; responses that include palpitations, perspiration, increased heartbeat and even dizziness; responses that can be so intense they stimulate us to ‘fight’ or ‘flight’. Yet as museologists Falk and Dirking point out, the experience of space is often a tacit one, ‘. . . [Architectural] influences are at once the most subconscious and the most powerful, the hardest to verbalise but the easiest to recall’ (1995:31).

The tacitness in the way people experience space may help explain its relegation to the semiotic margins, especially for those who do not work within the fields of architecture or urban design. At the same time, however, the all-pervasiveness of this semiotic cannot be ignored as it impacts on every aspect of our lived experience. It is this anomaly that has motivated some social semioticians to begin explicating and articulating the meaning potential of this semiotic – a challenge that this chapter also pursues.

In meeting this challenge, the chapter will focus exclusively on houses. Two main reasons have motivated this choice. First, most of us have been reared in a house of one type or another, so they are familiar to us. Second, houses provide one of our most formative experiences of 3D space. Therefore,

this chapter aims to explore domestic architecture in order to demonstrate the importance of housing to our physical and emotional wellbeing.

Home Sweet Home: Thinking about Houses Semiotically

One way of thinking about a house is as a 3D structure that provides shelter from the elements and in whose rooms we conduct the practical daily activities of our lives. These activities include sleeping, eating, washing, resting, reading and so forth. Thinking about houses as functional structures in this way is ideationally tuned.

Houses also have an interpersonal dimension. Most people hopefully have their first personal experience of security in their homes¹ and grow to feel comfortable in particular configurations of spatial enclosure. This is more than a fleeting moment for as French philosopher Gaston Bachelard suggests, 'The house we are born in is physically inscribed in us' (1964:14). Australian writer, David Malouf, expresses similar views:

First houses are the grounds of our first experiences. Crawling about at floor level, room by room, we discover laws that we apply later to the world at large. And who is to say if our notions of space and dimension are not determined for all time by what we encounter there, in the particular relationship of living rooms to attic and cellar (or, in my case, under-the-house), of inner rooms to the verandas that are open boundaries. Each house has its own topography, its own lore: negotiable borders, spaces open or closed. (Malouf 1985:8–9)

In fact, when we talk about domestic architecture, we commonly refer to our houses as our 'homes'. It is not unusual to hear expressions like 'home sweet home', 'home is where the heart is', 'my home is my castle' and 'I feel at home here'. Significantly, the word 'home' operates both ideationally and figuratively. At a figurative level, the phrase 'I feel at home' has a metaphorical meaning: 'I feel comfortable and secure here'. A meaning that encapsulates a strong sense of inner peace, refuge and belonging. Such metaphorical meanings are interpersonally rather than ideationally tuned, and this chapter hopes to illuminate something about both **house** and **home** using Halliday's metafunctions (1978) as the heuristic.

The chapter has been organized into three distinct parts. The first section briefly outlines the tools for analysing space from a metafunctionally diversified perspective (Halliday 1978, 1985/1994, Halliday & Matthieson 2004). The second applies these tools to the analysis of phylogenesis of domestic architecture in Australia from a western, non-Indigenous perspective. It is not possible, however, to explore all choices for housing in this chapter. So the parameters have been narrowed to a focus on two specific regions of the continent: the

tropics and the cooler *southern parts* of Australia. These two locations have been chosen because these choices for housing offer strong contrasts, and in doing so, provide illuminating insights into the important role domestic space plays in all of our lives. The final section of this chapter moves beyond space grammar, to explore some of the other potential applications of the tools we will be discussing as well as articulating some of the remaining challenges.

Social Semiotic Tools for Analysing 3D Space

Social semiotics has been strongly inspired by the research of a group of linguists working with Systemic Functional (SF) theory, namely, Halliday (1978, 1985/1994), Halliday and Hasan (1976), Halliday and Matthiessen (2004), Martin (1992) and Matthiessen (1995). Not only has their work inspired a theory of language, but a number of social semioticians have also theorized a range of communicative modes:

1. visual images (Kress & van Leeuwen 1990, 1996/2006, O'Toole 1994)
2. movement (Martinec 1997, 1998a, 1998b, 2000a, 2000b)
3. speech, music, sound (van Leeuwen 1991, 1999)
4. architecture/three-dimensional space (O'Toole 1994, 2004, Kress & van Leeuwen 1990, 1996/2006, van Leeuwen 1998, 2005, Stenglin 2002, 2004, 2007, 2008a, 2008b, 2009a, 2009b, Martin & Stenglin 2007, Ravelli & Stenglin 2008).

All of these theoretical accounts have drawn on Halliday's metafunctional hypothesis (1978). This hypothesis states that a semiotic system simultaneously fulfils three communicative functions: first, an *ideational function*, which constructs representations of human experience; second, an *interpersonal function* which is concerned with social interaction and the expression of attitudes; and finally, a *textual function*, concerned with the organization of a text into a meaningful whole. Each metafunction will now be explained in relation to 3D space.

The Experiential Metafunction

The experiential metafunction is concerned with the ways we construe our experiences. In space, it has three aspects. The first is concerned with the practical functions or uses a space has been designed to fulfil. O'Toole's work in this area identifies the following functions of space (1994): private versus public; domestic versus utilitarian; either industrial, commercial, agricultural, governmental, educational, medical, cultural, religious or residential. These functions constitute a very useful way for thinking about space ideationally and

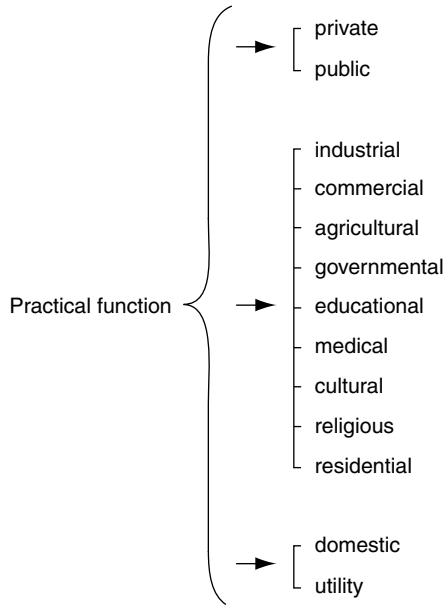


FIGURE 4.1 O'Toole's experiential functions of building as a system network (1994:86)

certainly one that is used by architects who focus strongly on form and function in their work.

From an SF perspective, is it possible to use a system network to represent the varying functions of space that O'Toole has identified (see Figure 4.1).

In applying this network to domestic Australian architecture, it is apparent that we are concerned with exploring spaces that are private, residential and domestic. However, O'Toole's public-private distinction is more complex than it initially appears. The reason being that the degree of privacy or public exposure tends to vary considerably from space to space. For example, *the most private* domestic spaces tend to be our bedrooms while the domestic spaces closest to the public end of the scale are the glass open-plan living areas that are popular in contemporary architecture and that passers-by can look into. These *appear to be public* as they are exposed to the external gaze but access and entry to them remain very much private and restricted. So, a more accurate representation would be *semi-public*. To accommodate such complexity, O'Toole's 'public-private' dimension could be represented as a sliding scale rather than a discrete set of choices (see Figure 4.2).

Using Martin's stratified model of context we can also project experiential meaning contextually, which yields a focus on field (Martin 1992:536). A field-focus means that experiential meanings can either have an object orientation

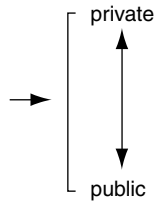


FIGURE 4.2 The public-private dimension

Table 4.1 3D space and field (activity and/or object orientations)

BUILDING: Field-focus	Activities	Objects
Learning (e.g. school, university, museum)	Teaching and learning (e.g. classroom)	Desks, chairs, books, pens, projectors, white/black boards, bins
Living/nesting (house, apartment)	Sleeping (e.g. bedroom)	Bed/mattress, sheets, pillows, blinds
	Cooking (e.g. kitchen)	Oven, stove, pots, pans, sink, utensils

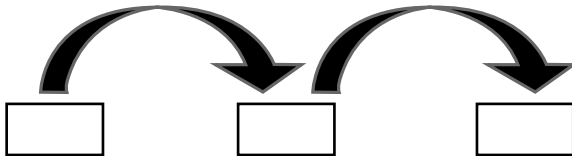


FIGURE 4.3 Serial structures

or an activity orientation. An activity orientation enables buildings to be classified in terms of doings or material processes. To briefly exemplify, we can distinguish between buildings for learning from buildings for living/nesting (see Column 1 in Table 4.1).

We can further classify the activity and object orientations of the spaces contained within these buildings. For example, if we look at the different types of activities that take place in the spaces of a house: we find spaces for cooking, sleeping and so forth (Column 2 of Table 4.1). In addition to these activities, a field-focus also requires us to account for the objects involved in each activity sequence (final Column of Table 4.1). In a kitchen, for example, the primary activity is cooking while the objects include the oven, stove, pots, pans and so on. In this way, a field-focus yields a finer-grained analysis of the functions each space has been designed to serve.

Ideationally, it is also possible to analyse space using particulate structure, in particular, orbital and serial structures. Serial structures are organized as a chain (Figure 4.3) such as the unfolding of rooms in a Victorian terrace house

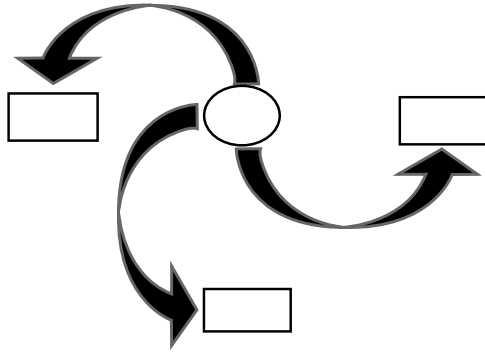


FIGURE 4.4 Orbital structures

along a central hallway. Orbital structures, on the other hand, are organized around a nucleus-satellite configuration (Figure 4.4). A typical example would be a house with a central courtyard functioning as the nucleus with all the rooms located around that courtyard.

The Interpersonal Metafunction

This section is concerned with tools that can help us understand the interpersonal dimensions that exist in the organization of space. One choice involves applying Kress and van Leeuwen's visual image tools to the third dimension (see Ravelli & Stenglin 2008). In particular, Ravelli and Stenglin have explored how the interpersonal resources for visual image analysis – power, social distance, involvement, contact and modality – can be used to analyse buildings such as Scientia at the University of New South Wales in Sydney, which has played a crucial role in reconstructing the visual identity of the university (Ravelli & Stenglin 2008).

A complementary perspective for analysing interpersonal meanings in 3D space is to take work on **feeling** as the point of departure. Drawing on Appraisal theory (Martin 1997, 2000, Martin & White 2005, White 1997, 1998), this approach has led to the development of two other semiotic tools: Binding and Bonding (Stenglin 2002, 2004, 2007, 2008a, 2008b, 2009a, 2009b). Binding is concerned with space and emotion, especially the security or insecurity dimension of affect. It is represented as a continuum. Choices for insecurity lie at both ends of the scale (Figure 4.5).

At one extreme of the Binding scale is the Too Bound dimension. This represents spaces that evoke feelings of claustrophobia and smothering while the other end of the scale contains the Too Unbound dimension. Spaces that

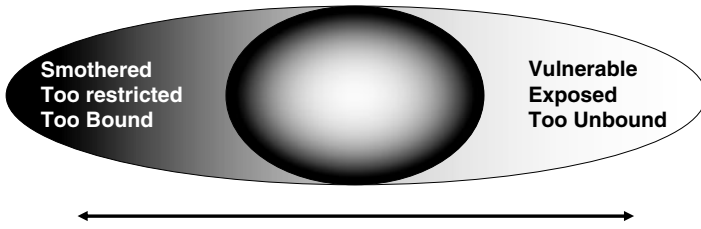


FIGURE 4.5 The Binding scale: Choices for insecurity

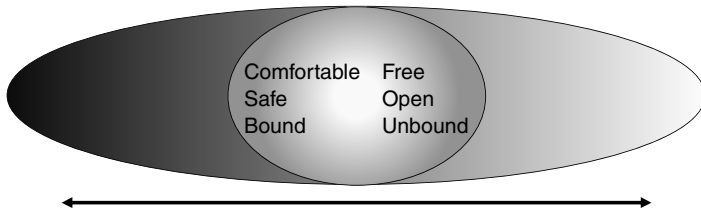


FIGURE 4.6 Binding: Choices for security

feel Too Unbound overwhelm their occupants by towering over them and not providing enough enclosure such as large public buildings and cathedrals.

The two other choices for emotion along the Binding scale are the *Bound* and the *Unbound* dimensions (Figure 4.6). Both are concerned with spatial security and located in the centre of the scale. Bound spaces are womb-like spaces that make users feel safe and secure while Unbound spaces also maintain a relationship of security with users by lessening the degree of spatial enclosure and making them feel freer and less enclosed.

Bonding is also concerned with interpersonal meaning in space but focuses on affiliation rather than in/security. It is a multidimensional resource concerned with aligning people into groups with shared dispositions. It explores ways of building togetherness, inclusiveness and solidarity through connection. There are at least four tools that materialize Bonding in the third dimension: Bonding icons, the attitudinal re/alignment of people around shared attitudes, classification and framing (Bernstein 1975).

Bonding icons are emblems of social belonging with the potential to rally people around shared values. They distil, crystallize and fuse interpersonal attitudes to ideational meanings. They include buildings (e.g. the Sydney Opera House), leaders (e.g. Nelson Mandela), songs (e.g. the Maori haka), symbols (e.g. Olympic rings) as well as medals, badges, trophies and even paintings (e.g. the Mona Lisa, which is a Bonding icon for the Louvre). They not only accrue values but radiate out for communities to rally around or reject (Stenglin 2008b, 2009a, 2009b, Martin & Stenglin 2007, Ravelli & Stenglin 2008).

Another resource for negotiating Bonding is attitudinal re/alignment. This is the process through which people are aligned into groups around shared attitudes (affect, judgement, appreciation). With regard to Attitude, Martin (2001) has demonstrated how shared affect has the potential to align people around shared emotions, shared judgement aligns them around shared principles and shared appreciation aligns them around shared tastes, preferences and values. Stenglin (2008b, 2009a, 2009b) has also demonstrated how attitudinal alignment unfolds logogenetically in museum exhibitions. Both Martin and Stenglin have shown that shared attitudes need to be evoked in response to some field, and in this way, attitudinal alignment involves the yin/yang coupling of ideational and interpersonal meaning, that is, field and attitude.

Another aspect of Bonding in space involves classification and framing as theorized by Bernstein (1975). According to Bernstein, strong classification means ‘things must be kept apart’ while weak classification means ‘they must be put together’. Things can refer to people, objects or spaces. In a strongly classified building, for instance, there are many rooms with discrete and separate functions and the objects found in each room are **not** interchangeable whereas weakly classified buildings tend to have spaces with many functions and highly interchangeable objects. Bernstein exemplifies this concept using a toilet. A strongly classified toilet only contains items related to personal hygiene: toilet paper, soap and a towel while a weakly classified toilet contains lots of objects: pictures, postcards, books and newspapers as well as ornaments.

Framing, as theorized by Bernstein, has two interrelated dimensions. The first involves the **physical boundaries** around a space. This fits well with the way van Leeuwen (2005) has defined the same concept in terms of textually connecting or disconnecting spaces. Second, Bernstein’s concept of framing refers to the **social interaction** between the participants in a space. For instance, in a strongly framed space the walls are impermeable and this decreases the potential for social *interaction* as participants are segregated from one another. Conversely, a weakly framed space has no partitions and this results in an open-plan enclosure with optimal potential for social interaction between the people occupying the space. These choices for classification and framing are summarized in Table 4.2 (also see Stenglin 2009a and 2009b for other examples of how classification and framing work in three-dimensional spaces).

Table 4.2 Classification and framing in 3D space (Bernstein 1975)

Strength	Classification	Framing	Potential for social interaction
strong	things kept apart	compartmentalized rooms	decreased
weak	things kept together	open-plan spaces	optimal

The Textual Metafunction

Textually speaking, there are several dimensions to consider. First, space involves movement as it unfolds along a **Path – Venue** trajectory and users are lured from one space into another through choices for **Prominence** such as gaze vectors, strong colours, the scent of freshly brewed coffee, flooring vectors and so forth. (See Stenglin 2004, 2008b, 2009a and 2009b; Martin & Stenglin 2007 for more applications of these semiotic tools to 3D spaces.)

Second, although space unfolds logogenitically it is actually stationary, so static tools are also needed for the textual analysis of space. Static tools include Theme-Rheme (Halliday 1985/1994) and information value, which Kress and van Leeuwen (1996/2006) have theorized for 2D visual image analysis, but which can be applied to some 3D spaces (see Stenglin 2009a for the application of Theme-Rheme and information value to the analysis of the external spaces of the Hyde Park Barracks Museum in Sydney).

Framing as theorized by van Leeuwen (2005) is also germane to our understanding of the textual organization of space. It explores the demarcation of one space from another. In particular, van Leeuwen makes a distinction between framing choices that segregate space into two distinct entities and those that separate one space from another through emptiness. His account of framing choices for 3D space is represented in Figure 4.7.

Given that there are many tools in the space grammar kit, it is not possible to apply all of them to the analysis of the domestic Australian architecture. So the following selections will be made. **Ideationally**, I will investigate the functions of domestic spaces. I will also explore field, that is, the activity and object orientations. **Interpersonally**, I will analyse for Binding and Bonding. The exploration of Bonding will focus on Bernstein's classification/framing, the coupling of field and Attitude as well as Bonding icons. In relation to classification/framing I will briefly touch on the contact dimension as developed by Kress and van Leeuwen (1996/2006). **Textually**, I will only focus on framing (van Leeuwen 2005:18) as the strength of boundaries connecting or disconnecting one

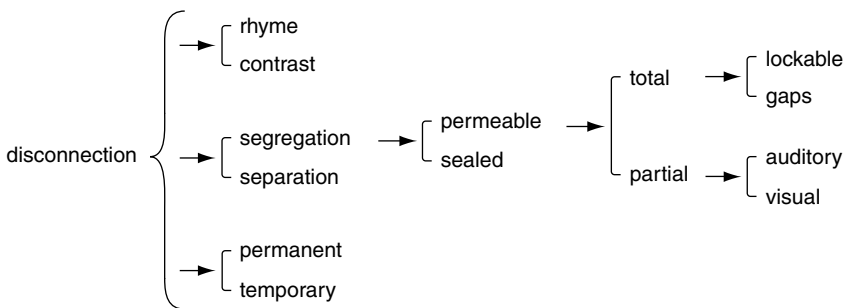


FIGURE 4.7 Framing system network (van Leeuwen 2005:18)

domestic space from another are extremely important in understanding how the cultural conceptualization of 'home' has changed over time in Australia. Having set the parameters, let us begin to explore domestic Australian housing starting at the end of the eighteenth century when the first permanent dwellings began to be constructed.

The Phylogenesis of Domestic Australian Architecture

The following account is based on a general survey of Australian architectural history from a western perspective. Surveying trends in the phylogenesis of Australian domestic architecture has involved reading the work of prominent architectural historians and reinterpreting key periods in the evolution of Australian housing using the theoretical tools discussed in the section on 'Social semiotic tools for analysing 3D space' of this chapter. More contemporary accounts have also been informed by current architectural writings but have involved a sampling of convenience, that is, the analysis and photography of domestic spaces I have had access to through personal networks of association.

Familiarity in a Hostile Alien Environment

At the beginning of European settlement, the earliest cottages in Australia were built of temporary materials but there was a preference for durability, and soon single storey houses were built based on the Georgian English cottage model (Broadbent 2001). These houses initially consisted of two rooms: a living room entered from the front door and a bedroom accessed from inside the living room (Archer 1987). **Textually**, both rooms were very firmly constructed through strong choices for framing. Thick boundaries constructed of non-porous bricks permanently segregated the spaces from one another. So there was strong disconnection between the interior spaces.

Experientially, the houses were residential, domestic and **private** in function. The bedroom was the most private space and used for personal activities of sleeping, washing and toileting while the more communal activities of daily life took place in the living room – a semi-private space that all visitors to the home occupied.

Interpersonally, an important aspect of Bonding is the way these English cottages were constructed through choices for familiarity in a harsh climate and hostile environment. These early homes exemplified what the new arrivals believed would make them feel 'at home' on the other side of the world – beliefs that were based on their experience of housing in the northern hemisphere (Brown 2000:108). Thus, early houses were replicas of familiar domestic dwellings overseas. Aesthetically, this was most evident through the choice of the Georgian architectural style.

The replication of a familiar aesthetic style also meant that the cottages were able to function as Bonding icons in an alien environment through the evocation of strong positive attitudes. These attitudes included positive appreciation of the Georgian aesthetic; feelings of positive affect, especially familiarity, security and happiness alongside positive judgements of, and confidence in, the British empire which was the home of the Georgian cottage. By functioning as Bonding icons in this way, the early cottages were able to palpably rally the immigrants around shared ideals of 'home and hearth' and sustain their connection to their country of origin to which they still emotionally belonged.

The early settlers living in Australia were not alone in clinging to such Bonding icons. According to architectural writer Balwant Saini:

All of the eighteenth and nineteenth century colonial domestic buildings had one thing in common. Their builders, whether they were British, Dutch, French or Portugese, wanted to keep the memory of their mother country alive; and so their preferences for certain prototypes was quite discernible. (Saini 1982:18)

Significantly, however, nobody seemed to realize that the **material** choices for housing they were replicating had actually developed in response to the climactic conditions of the northern hemisphere. Yet this had important implications for **Binding**. In particular, the warmth of the Australian climate made people living inside such houses feel smothered: 'The hot sun beating down on their bare facades made them unbearable in the summer' (Evans 1983:8). Thus, Handel Richardson also observes:

How glad he was to leave the tiny, sun-baked box that till now had been his home . . . It had neither blind nor shutter; and, on entering it of a summer midday, it had sometimes struck hotter than outside. (Richardson 1982:236)

People living in such houses clearly felt Too Bound, especially in the oppressive heat of summer. This is important as it points to the fact that the design of space is not a world-wide given in respect of what gives comfort and security to occupants. There is clearly cultural and climactic variation.

Respite from Oppression: Verandas

To alleviate these strongly negative interpersonal feelings and meet peoples' need for comfort and security, verandas became commonplace from 1792 onward. Interestingly the concept of a veranda was transported from India, also a British colony (Drew 1992) as seen in Image 4.1. So from very early on, the visual Bonding connection to the mother country was slightly weakened as the Georgian cottage aesthetic was permanently modified.



IMAGE 4.1 The veranda: Respite from oppression

In practical terms, verandas provided shade, shelter and protection from the elements. They also moderated the heat of the sun on the walls (Archer 1998). Initially a defence against the Australian climate, together with other choices for Binding such as shutters, thick walls and heavily lined curtains, they were used to deliberately exclude the sun as well as the **alien Australian landscape** from intruding into nostalgically furnished domestic interiors (Drew 1992). In this way Australian homes in the south became strongly Bound fortresses that shut out the threatening and unfamiliar elements, that is, the climate and the landscape.

Let us now explore how classification and framing impacted on social interaction in these Bound fortresses. As we have already seen, most homes in the south were initially built with two rooms. This meant that the initial classification of the spaces was weak. However, as soon as the materials and financial resources were available, additional rooms were added. Such expansion strengthened the classification of the house. Those with the financial means expanded into four and six room houses maintaining the English cottage plan as the model. The ideal was strong classification: one room per person and one room for cooking, a separate room for dining, a separate room for reading, sleeping in and so forth (Boyd 1952:12).

In such strongly classified and strongly framed spaces, furthermore, people have privacy. Privacy was important in the early cottages. The strong valuation of privacy actually began two centuries before Australia was occupied, when Queen Elizabeth of England proclaimed the principle of a private house for each family (Boyd 1952). Privacy is thus a relatively recent phenomenon in human history, and one that was transported to the Australian continent with British occupation/invasion.

Classification according to Bernstein is also about power. For example, in the dining room where formal meals were eaten in houses with enough space to accommodate this, strong classification and framing meant there were also strong boundaries around the behaviours and types of conversations that occurred in such spaces. In fact, 'etiquette' books articulated the behaviours expected of people at dinner parties: how to hold utensils, when to start eating, how to politely refuse a course and so forth (Kapetanios Meir 2005, Grylls 1994). Such books also listed the topics considered **inappropriate** for dinner conversations. They include servants, religion, politics, illness and sex. The advice was to avoid such topics as they could polarize guests and make them feel insecure. In these ways, the power dimension of strong classification and framing impacted significantly on behaviour and social interaction as the host and hostess were orchestrating and controlling the conversation to ensure that **only certain things could be spoken about in certain rooms.**

To summarize, ideationally, the spaces of the Georgian cottage were private or semi-private. Textually, the spaces were permanently segregated from one another and the addition of verandas constituted an additional layer of defence that deliberately excluded the sun and the alien Australian landscape from intruding into the interiors. Interpersonally, cottages were constructed to feel like Bound fortresses. Strong classification and framing meant the occupants had privacy but social interactions were formal and constrained while Bonding was directed outward to the mother country. Having found a baseline for security that was both comfortable and familiar, people living in the south maintained it for many, many years. Let us now turn to choices for housing in the tropics.

The Tropics

Houses built in the tropical regions of Australia also followed the Georgian English cottage plan. Initially they fought the land in the same ways as their counterparts in the south. Their interiors were also strongly classified and framed into 4 to 6 compartments for cooking, sleeping, dining and entertaining. They also quickly added verandas to provide them with much-needed shelter and shade. Verandas, moreover, typically covered all four sides of the house and not just the front as they did in the south.

However, in the tropics, roofs were made of corrugated galvanized iron as it was light and could be transported long distances at low costs. Metal, however, is a poor choice of material as it is a poor insulator and good heat conductor. So roofs heated up quickly and this heated everything below them. Temperatures inside these houses were commonly double the temperatures outside. Such soaring heat made living conditions intolerable. So the interpersonal relationship set up with the occupants was one of intense smothering. In response, windows became larger to allow the breeze inside and people began

using verandas more and more to provide respite from the heat and access to uninterrupted breezes.

Thermal comfort is clearly a very important part of Binding or feeling secure in a space. In fact, the hotter the climate, the more the veranda was used. Not surprisingly in the north of WA, Queensland and the NT the veranda became more than a shelter to the rooms – it became the main living area. In these regions, the width of the veranda physically expanded while the size of the rooms contracted. Soon the veranda was used not only for dining and sitting, but also for sleeping. In the dry season, verandas were accordingly furnished with tables, chairs, pictures and vases. The need for privacy was served by the interior of the house where Bound – securely enclosed and strongly framed rooms – were used for undressing.

In terms of field and activity orientation, dining, sitting and sleeping were not the only possibilities for action on the veranda. Verandas had many other uses as well:

Most family life took place on the verandah, which functioned as a dining room, a recreation centre, playground for the young on wet or scorching days, store room and vantage point for surveying the scenery or passers-by. Suspended from its rafters were the meat safe, the water bag, the clothesline in bad weather, swings for the children, bird cages, the Christmas hams and numerous pieces of wire or hooks on which to hang hats, bags and overcoats. At night it was the coolest place to sleep, with a mosquito net carefully tucked in for protection from the abundant tropical insect life. (Archer 1998:27)

In other words, the veranda was the space in which all the activities of daily life took place (Drew 1992).

From the perspective of Bonding icons, these houses represent an interesting development. Known as 'the Queenslander', they were one of the first vernacular housing styles to develop here from a western perspective and came to be a Bonding icon for the entire colony – one could even argue that they still are. In fact, the Queenslander is characterized by two features: a wide, all encompassing veranda and their elevation on stilts to increase airflow. Symbolically, they point to the fact that the interpersonal bond to the mother country has begun to weaken.

In terms of framing, the importance of the veranda in providing shade, shelter and access to breezes meant that solid walls could not be used to compartmentalize it in permanent ways. The occupants needed flexibility to be able to shift from one part of it to another as the direction of the sun changed during the day. This flexibility and minimal framing meant that people were able to complete the activities of their daily living comfortably. Also in interpersonal terms, it meant that it represents a significant shift to declassifying both the activities of domestic living and the spaces in which these activities took place.

The significance of this development has four dimensions. First, verandas broke down the barriers separating internal and external spaces. They did

this by extending the living area into the semi-outdoor realm. This meant that people living in the tropics did not shut themselves indoors as people in the south did. They actually lived **around** the house more than inside it. So there were two living areas in the tropics: the veranda, which was semi-public and strongly Unbound; and the compartmentalized interior which was private and strongly Bound.

Second, verandas began to dissolve the compartmentalization of living areas through weak classification and minimal framing. This forced the development of a more open and fluid lifestyle – one which was not characterized by the strong boundaries and classifications of housing in the southern part of the continent. This in turn meant a significant break down in the division of behaviours associated with the kitchen, the dining room and the parlour in the past. These behaviours now occurred simultaneously in one large space – the veranda.

In addition, once the barriers compartmentalizing living had broken down, it was not just the range of peoples' behaviours that increased. The range of interactions that were possible between the occupants of a house also increased accordingly as there was no longer a one to one relationship between a room, its function and behaviour. The potential for conversation on the veranda was therefore greater as the host/hostess were no longer in control; so the topics could range from the intimate to the more general.

Finally, weak classification and weak framing yields openness and freedom. This openness, however, is a double-edged sword. Its flipside is that it enables surveillance as discussed by Foucault in relation to the panopticon prison (1977/1991). This meant that the occupants of the Queensland veranda had a vantage point for looking at passers-by but they could also be continuously scrutinized. So textual choices for framing were strengthened by the addition of adjustable louvres and lattice (see Image 4.2). These optimize breezes and



IMAGE 4.2 Strengthening framing and increasing privacy

maintain thermal comfort, and together with the Bound rooms of the interior functioned to give the occupants of 'the Queenslander' their privacy.

In this way, Australia developed two very different baselines for domestic security: the Unbound in the tropics and the Bound in the south of the continent. But baselines for security are not static. They are dynamic and evolve in response to cultural changes, technological innovation and economic factors. As a result, houses in the south also moved towards the Unbound dimension of the security scale albeit over a much longer time period.

The Shift to Unbinding in the South

The first shift to Unbinding that occurred in the south was a direct consequence of increased material wealth after the gold rushes of the 1850s. During this time, a new middle class emerged. Not only did they build homes using a magnificence of scale that characterizes churches and other public buildings, they built on large estates in parkland surrounds (Fitzgerald 1999). These gardens were very important to their sense of security because they also vertically screened the house from passers-by. So the owners could install 9 feet high windows, feel strongly Unbound yet maintain both a sense of privacy and a relationship of security with their domestic spaces. Courtyards operate on the same principle: they Unbind by dissolving solid enclosures through choices like window walls/sliding doors and this extends the space outward and opens it up to the exterior but a more distant vertical boundary still provides privacy and a sense of enclosure.

Although Unbinding has a clear economic dimension here, the shift towards Unbinding in the wealthy homes of the nineteenth century did impact on spatial enclosure for ordinary citizens too. In particular, increased light and ventilation changed communal views on the desirability of such things, so much so that after 1900, and the outbreak of the Bubonic plague, their inclusion was seen as necessary for public health. As a result ventilation and lighting levels were incorporated into public legislation and Sydney residents were encouraged to move to the ring suburbs of Randwick, Bondi and Coogee (Fitzgerald 1999).

After Federation in 1901, the strong compartmentalization of domestic life in the south began to erode. At that point in Australia's history, domestic help was disappearing, so the kitchen, previously located at the rear of the house, was relocated **beside** the dining room so that meals could be more easily prepared and served. Next, choices for framing these two adjoining spaces weakened when 'serveries' – literally holes in the wall separating the kitchen and dining room – were introduced (Archer 1987). These weakened the framing of the two spaces by making them partially permeable.

After World War II, there was another major challenge for domestic spaces: materials were scarce and rooms diminished in size. The first major response to

this was the appearance of a **combined** living and dining room referred to as **the common room** (Boyd 1952). In addition, architects began using arches rather than solid walls to separate the **common** room from the sitting room. Textually, this meant that the flow of internal spaces was more continuous and integrated. Experientially, this represented a very profound shift in the culture as internal boundaries, driven by the English desire for privacy, had been firmly entrenched for 150 years and this shift had major interpersonal consequences.

First, the people had to learn to feel secure in houses that had less spatial enclosure. Rather than being strongly Bound they now felt minimally Bound. Second, the merging of rooms weakened classification and framing, and forced a change in social and cultural attitudes to domestic living. According to architectural theorist Robin Boyd, they 'required a degree of social informality contrary to the established concept of suburban life' (Boyd 1952:184).

The material shortages of life after World War II provided other challenges as well. In particular, the size of rooms diminished as a direct consequence of legislation restricting the size of houses to either 92 or 111 square metres. This provided architects with a deep challenge as reducing the size of a space makes it feel oppressive. To prevent people feeling Too Bound, they began to use windows more judiciously as they unbind occupants and provide a sense of spatial freedom:

. . . no one working at enclosing space, not even the most mercenary speculative builder, could have failed to note how the threat of claustrophobia in the diminishing rooms of the 1940s was counteracted by the increasing breadth of their windows. (Boyd 1952:186)

Unbinding by increasing the size and span of windows was thus deliberately adopted to compensate for the decrease in available space. Every window within reach of a corner ran into and turned it with a curve of glass. By 1950 the material shortages were less of a problem but the trend to Unbinding in the south continued more strongly than ever. In fact the shift to weakly classified internal spaces and Unbinding through windows paved the way for the open-plan 'glass house' living of today which was first introduced by Harry Seidler in 1948 at Rose Seidler House built for his mother in Turramurra, Sydney.

Rose Seidler House is a landmark in domestic Australian architecture in the south as it pushed all the boundaries towards Unbinding and weak classification and framing that had been occurring steadily since 1900. So much so that it was seen as radical and confronting and elicited strong feelings of insecurity from the general public. The main innovations were the use of movable floor to ceiling glass walls, known as sliding doors, and the removal of internal partitions; so sleeping, play and utility areas merged into one. Once again weak classification and weak framing delivered 'fishbowl living' and locals gathered outside on weekends to peer in at the occupants. Despite this,

Binding also takes into account individual as well as cultural variation and although many people felt insecure at the thought of fishbowl living, there are people who enjoy being on display (personal communication, resident, Harry Seidler's Horizon building, Sydney, 10 October 2005).

Nevertheless, the publicity Rose Seidler house received together with the work of other modernist architects meant domestic architecture became increasingly Unbound in the south. People became accustomed to houses such as those built by Seidler, and over time, such houses became a 'Given'. Not only that, they established a new and Unbound cultural baseline for security in the south with an emphasis on outdoor living. So much so that many houses are now designed to flow out into gardens (Image 4.3).

This is a strong trend especially in refurbished inner city terraces. Bi-folding doors are frequently used to optimize the permeability between internal and external spaces, and integrate them so that they flow seamlessly into one another. From the point of view of security, the occupants of such houses maintain their privacy as the courtyard provides external boundaries that screen them from the voyeuristic gaze of neighbours in the same way as gardens screened the internal spaces of the wealthy in the nineteenth century.

Another trend in contemporary Australian housing is to use glass window walls to extend the indoor spaces out onto panoramic vistas of the natural environment or urban skylines (Image 4.4). From the point of view of attitudinal re/alignment, this is a most significant cultural shift as it evokes a strongly positive valuation and appreciation of the Australian landscape together with a sense of pride and confidence in its city skylines, beaches, mountains and bush. It also reflects a strong love and affection for the land, which is now openly invited inside. This positive western valuation of the Australian environment



IMAGE 4.3 Unbinding and privacy: The courtyard



IMAGE 4.4 Unbinding to the external landscape

began with Harry Seidler and has become so strong in recent years that Renzo Piano, world famous Italian architect has said: 'I think in this country the sensitivity to nature, to breeze, to view, to sun is stronger than anywhere else' (Drew 1999:xv).

From the point of view of Bonding and social interaction, the refurbished Unbound terrace is often characterized by weak classification and weak framing especially in their living areas which are now designed to serve multiple functions. Open-plan in design, they function as a living room, library, home theatre, informal dining room, study and playroom. Exciting as it may sound, living in one room with many functions is not a positive experience for everybody. Sydney journalist Maggie Alderson describes it in the following way:

Lovely notion as it is to have one sprawling family area, with the youngest child doing homework at the kitchen table while Dad cooks a stir-fry listening to the cricket on the radio, Mum reads the paper and two teenagers play a violent video game, the reality is an imperfect experience for everyone involved. (Alderson 2007:41)

The dissatisfaction Alderson expresses seems to stem from the fact that weak classification and framing result in **too much** interaction. They also result in surveillance, which is why parents often like open-plan areas: they can easily keep an eye on their children. Teenagers, however, tend to respond negatively to continuous adult surveillance and seek refuge in the strongly classified and framed spaces of their bedrooms (Image 4.5). These provide them with sanctuary and escape.



IMAGE 4.5 Close me in and set me free: Strong classification and strong framing

Having swung from one extreme to the other, it seems that many of us need **both** openness and enclosure in our domestic spaces. Perhaps the capacity to provide both explains why the Queenslander has successfully survived as a housing choice for such a long time. It also explains why the house shown in Images 4.4 and 4.6 functions so effectively. The open-plan area is both a dining room and a lounge room. Large glass sliding doors dissolve internal-external boundaries but the inclusion of diaphanous curtains reduces the permeability of the space and provides the family with privacy whenever it is needed.

Privacy, for this family, is usually sought in the mornings. So the curtains remain drawn until the family is dressed and ready to open itself up to the light, activity and scrutiny of the world outside (Image 4.7). Once open, the curtains tend to stay open all day and all night. Regardless of the choice the family makes at any point in time, the curtains have the potential to either strengthen or weaken the framing of the space and give the family control over the openness of the house, that is, its privacy or degree of exposure.



IMAGE 4.6 Unbound terrace: Sanctuary, fishbowl or both?



IMAGE 4.7 Curtains: A strong choice for spatial security



IMAGE 4.8 A Bound nook for TV viewing

On the middle level of this terrace, where most of the communal activities take place, the house contains another living area (Image 4.8). This space is located adjacent to the open area just discussed but the choices for enclosure are very different. The second space is strongly Bound. It is mainly used for TV viewing and has become one of the family's favourite spaces. It is small and the furniture is laid out in such a way that everybody sits in very close proximity to one another but the TV is the focal point of the space, not the interaction between the occupants. Having the TV as the focus softens engagement and mitigates contact through oblique angles, to use Kress and van Leeuwen's tools (1996/2006), but still enables the entire family to participate in the shared activity of watching TV.

It is easy to see why this small, cocooned space is popular for family with two teenagers. It provides a secure environment that enables the family members to commune by drawing them into physically close proximity with one another and then allowing them to simply 'be' without pressure to overtly share feelings or attitudes. Bonding in this space is not about valuing the external Australian landscape. It is about feeling part of an important social unit and enjoying the deep attachments that form between family members without demanding any of the intimate familial social interactions that many teenagers seem to struggle with, and rebel against. It is therefore not surprising that the mother of the family says she would never demolish that wall and turn the living area into one seamless space (personal communication, 5 November 2007). It also seems that the nature of the Bonding, the physical and emotional connection family members desire changes over time, and our spaces can be designed to accommodate those changes and facilitate the types of interaction that are sought.

Returning to Bonding icons, this analysis so far has been strongly oriented to the way whole houses crystallize interpersonal attitudes to the land. At the heart of this discussion has been a consideration of who and what is allowed to enter into the domestic space and who or what is deliberately excluded. If we link the idea of inclusiveness and exclusiveness to the notion of a hierarchy of Bonding icons, it seems that households, including couples, may develop personalized (as opposed to rallying national) icons. The family we have just been discussing, for instance, have a collection of teddy bears, which lives on the sill in the Bound TV nook (Image 4.9). These teddy bears evoke feelings of security, love, warmth, tenderness, affection, happiness and intimacy. Each member of this family has their own personal teddy with their name sewn on it and all the members of the family are represented: past and present, nuclear and extended. Significantly, these teddies are not a rallying icon like the Olympic torch (see Stenglin 2008b for a detailed analysis of the Olympic torch as a Bonding icon). They are much more privileging – you have to be part of this family to belong to that ledge.

So it seems that we can grade Bonding icons along at least three dimensions: local icons (teddy bears) and international icons (the Olympic flame). Second, each Bonding icon has the potential to evoke an affectual charge that varies in its intensity: at times it can be so strong that it moves you to tears, at other times it may just evoke a feeling of warmth. This means the intensity of the affectual charge can be graded along a continuum of minimal to maximum. Finally, the function of Bonding icons varies considerably: some rally while others privilege.



IMAGE 4.9 Local privileging bonding icons

Regarding the home and house distinction established at the beginning of this chapter, an interesting question to ponder is when does a house become a home? When do we stop thinking about a structure ideationally and start feeling it interpersonally? Perhaps the answer can be found in Bonding: it's the myriad of intimate social interactions that our spaces facilitate with those whom we love and with whom we share our lives that make a home. It's about the quality and nature of the contact that takes place in our familial spaces, the types of interpersonal connection that our choices for classification and framing materialize and the myriad of shared activities and shared attitudes that evolve over time. Above all, it seems to be about affect, about that emotional charge that can interpersonally move us to tears. Affect seems to be the glue that binds it all together. It is not surprising then that home is where the heart is.

Beyond Space Grammar

Moving beyond the space grammar and domestic Australian architecture, the final section of this chapter concludes by extending the application of the tools we have been exploring as well as identifying those dimensions that need further consideration. First, the contact dimension as theorized by Kress and van Leeuwen (1996/2006) in relation to visual images needs to be incorporated into the space grammar. At this point, it appears to sit most comfortably within Bonding. The reason for this being that Bonding is concerned with social interaction and contact constitutes an important dimension of that. If a space has weak classification and weak framing, for example, it is open to surveillance. At the heart of such surveillance is contact and it raises some interesting anomalies: living in a fishbowl appears to mainly be an offer but it can also be a demand . . . how do we reconcile these choices?

In addition, contact may need to be adapted in relation to 3D space. In 3D space, for example, it is not just contact that is important but the directionality of the contact: is it one-way or two-way? If one way, is it 'in-out' as we saw with lattice-enclosed verandas or is it 'out-in' as is the case at night when outsiders look into plate glass enclosures. Control also seems to be an element for consideration. To what extent can the occupants control the extent of the contact through choices for screening offered by curtains, blinds and shutters? Another related dimension seems to be the participants: the 'who' or 'what' the contact is with. Is it with the natural environment, passers-by or other members of the household? As the participants may be inanimate objects there is also a need to theorize furnishings and the ways we interact with objects as the discussion of the TV and contact in the contemporary Glebe terrace pointed out. Schefflen and Ashcroft's (1976) work on territoriality would constitute an important starting point here but one would clearly need to go beyond and consider the social implications of such interaction on Bonding.

In addition, Binding and Bonding seem to have broader applicability than 3D space. They can also, for example, be applied to 2D representations of space such as those we find in advertisements. They equally apply to static images in picture books and moving images in films and 2D artworks. Significantly, the chapter by Painter, Martin and Unsworth (Chapter 6) in this volume extends 'bound' and 'unbound' to the analysis of textual meanings in 2D picture books. This is an interesting development as they are not only being applied to a different semiotic but their metafunctional complementarity is also being explored.

Furthermore, two systems have been developed to account for the materialization of Binding (Stenglin 2004): permeability and ambience. Permeability is concerned with the degree to which a space can be penetrated by the elements. Underlying this is the concept of a space as a membrane, a covering or a shell. Clearly, the concept of membrane relates to other things as well like the skin on our bodies and our clothing, which many refer to as our second skin.

This similarity between skin, clothes and space seems to suggest that the concepts of Binding and Bonding apply not only to space but to clothing as well. To explore this possibility in relation to clothing, the Too Bound dimension is evident in past practices such as foot binding. It is also often evoked by suits and ties, which many experience as smothering. Feeling Bound on the other hand involves feeling comfortable, for example, being snugly dressed to suit cold weather while Unbound clothing involves peeling off the layers. The Too Unbound dimension seems to involve feeling vulnerable and too exposed. It therefore seems to relate to exposing more of oneself than one feels comfortable with. Intercultural variation is pivotal here.

Clearly, much remains to be explored in relation to the grammar of three-dimensional space including one final dimension, its multisensory nature. The work presented in this chapter has only taken **vision** into account yet the experience of space involves all of the senses: sight, sound, smell and touch – and most of these are largely unexplored.

Conclusion

In closing, the phylogenetic analysis of domestic architecture has shown space to be an all encompassing semiotic of monumental importance because from the moment of our conception in the womb to the moment of our burial, we are continuously enveloped by it. In fact, the ubiquitous nature of space gives it an omnipresence that is in our interests to understand as it has the potential to deeply affect the quality of our lived experience. We should therefore reclaim it from the semiotic margins. How? By using the tools explored in this chapter as the starting point for creating spaces that free us to engage in the activities we value most, and facilitate the interactions we desire with all those with whom we share our lives.

Acknowledgements

First, I would like to thank Joan Rothery for opening up and guiding my exploration of architectural spaces, especially housing on the Australian continent. Second, I would like to thank my friends and family – Carolyn, Clare, Daniel, Kelvin, Stephanie, Vanessa and Yarro – who trusted me to photograph their private spaces, analyse them and share my thoughts in the public domain. I also owe an enormous debt of gratitude to my critical readers – Chris Cléirigh, Sally Humpreys, Shooshi Dreyfus Michele Zappavigna and Ahmar Mahboob – for their insightful and encouraging comments. Finally, I thank Roland Stocker for his generosity in assisting with the diagrams and supporting me in every aspect of this work.

Note

¹ See Stenglin (2009b) for an alternative exploration of domestic security – one that occurs in homes characterized by abuse – verbal, physical or sexual. This account is based on an analysis of an exhibition called ‘Scumbag’ by renowned Australian artist and photographer, Ella Dreyfus.

References

- Alderson, M. (2007). Makeover madness. *Sydney Morning Herald, Good Weekend Magazine*, June 30, p. 41.
- Archer, J. (1987). *The great Australian dream: The history of the Australian house*. Sydney: Angus & Robertson.
- Archer, J. (1998). *Your home: The inside story of the Australian home*. Port Melbourne, Victoria: Lothian Books.
- Bachelard, G. (1964). *The poetics of Space* (M. Jolas, trans.). New York: Orion Press.
- Bernstein, B. (1975). *Class, codes and control: Volume 3* (2nd edn). London, Boston, MA and Henley: Routledge & Kegan Paul.
- Boyd, R. (1952). *Australia's home*. Melbourne: Melbourne University Press.
- Broadbent, J. (2001). The colonial bungalow. Unpublished talk, 31 July, Australian National Trust Centenary of Federation Lecture Series, Sydney: SH Irwin Gallery.
- Brown, N. (2000). Making oneself comfortable, or more rooms than persons. In P. Troy (Ed.), *A history of European housing in Australia* (pp. 107–124). Cambridge: Cambridge University Press.
- Drew, P. (1992). *Verandah: Embracing place*. Sydney: Angus & Robertson.
- Drew, P. (1999). *Touch this earth lightly: Glenn Murcutt in his own words*. Potts Point, Sydney: Duffy & Snellgrove.
- Evans, I. (1983). *The Australian home*. Sydney: Flannel Flower Press.
- Falk, J. & Dierking, L. (1995). *Public institutions for personal learning: Establishing a research agenda*. Washington, DC: American Association of Museums.

- Fitzgerald, S. (1999). *Sydney: A story of a city*. Sydney: Pot Still Press.
- Foucault, M. (1977/1991). *Discipline and punish: The birth of the prison* (Alan Sheridan, trans.). London: Penguin Books.
- Grylls, D. (1994). Victoriana. *Essays in Criticism*, XLIV (4), 346–352.
- Halliday, M.A.K. (1978). *Language as a social semiotic: The social interpretation of language and meaning*. London: Edward Arnold.
- Halliday, M.A.K. (1985/1994). *An introduction to functional grammar*. London: Edward Arnold.
- Halliday, M.A.K. & Hasan, R. (1976). *Cohesion in English*. London: Longman.
- Halliday, M.A.K. & Matthiessen, C.M.I.M. (2004). *An introduction to functional grammar* (3rd edn). London: Edward Arnold.
- Kapetanios Meir, N. (2005). A fashionable dinner is arranged as follows: Victorian dining taxonomies. *Victorian Literature and Culture*, 33(1), 133–148.
- Kress, G. & van Leeuwen, T. (1990). *Reading images*. Geelong: Deakin University Press.
- Kress, G. & van Leeuwen, T. (1996/2006). *Reading images: The grammar of visual design*. London: Routledge.
- Lumley, A. (1992). *Sydney's architecture*. Melbourne: Longman Cheshire.
- Malouf, D. (1985). *12 Edmondstone Street*. London: Chatto & Windus.
- Martin, J.R. (1992). *English text*. Philadelphia, PA/Amsterdam: John Benjamins Publishing Company.
- Martin, J.R. (1997). Analysing genre: functional parameters. In F. Christie & J. R. Martin (Eds), *Genres and Institutions: Social Processes in the Workplace and School* (pp. 3–39). London: Cassell.
- Martin, J.R. (2000). Beyond exchange: APPRAISAL systems in English. In S. Hunston & G. Thompson (Eds), *Evaluation in text* (pp. 142–175). Oxford: Oxford University Press.
- Martin, J.R. (2001). Fair trade: Negotiating meaning in multimodal texts. In P. Coppel (Ed.), *The semiotics of writing: Transdisciplinary perspectives on the technology of writing* (pp. 311–338). Brepols: Semiotic and Cognitive Studies X.
- Martin, J.R. & Stenglin, M. (2007). Materialising reconciliation: Negotiating difference in a post-colonial exhibition. In T. Royce and W. Bowcher (Eds), *New directions in the analysis of multimodal discourse* (pp. 215–238). Mahwah, NJ: Lawrence Erlbaum Associates.
- Martin, J.R. & White, P.R.R. (2005). *The language of evaluation: Appraisal in English*. London: Palgrave.
- Martinec, R. (1997). Towards a functional theory of action. Paper presented to the Multimodal Discourse Analysis workshop, University of Sydney, 15–17 December.
- Martinec, R. (1998a). Cohesion in action. *Semiotica*, 120(1/2), 161–180.
- Martinec, R. (1998b). Interpersonal resources in action. *Semiotica*, 135(1/4), 117–145.
- Martinec, R. (2000a). Rhythm in multimodal texts. *LEONARDO*, 33(4), 289–297.
- Martinec, R. (2000b). Construction of identity in Michael Jackson's jam. *Social Semiotics* 10(3), 313–329.
- Matthiessen, C.M.I.M. (1995). *Lexicogrammatical cartography: English systems*. Tokyo: International Language Sciences Publishers.
- O'Toole, M. (1994). *The language of displayed art*. London: Leicester University Press.

- O'Toole, M. (2004). Opera Ludentes: A systemic-functional view of the Sydney Opera House. In K. O'Halloran (Ed.), *Multimodal discourse analysis* (pp. 11–27). London/New York: Continuum.
- Ravelli, L.R. & Stenglin, M. (2008). Feeling space: Interpersonal communication and spatial semiotics. In E. Ventola & Gerd Antos (Eds), *Handbook of Applied Linguistics, Volume 2, Interpersonal communication* (pp. 355–393). Berlin: Mouton de Gruyter.
- Richardson, H. (1982). *The fortunes of Richard Mahoney*. Ringwood, Victoria: Penguin Books Australia.
- Saini, B. (1982). *The Australian house: Home of the tropical north*. Sydney: Lansdowne Press.
- Schefflen, Albert E., & Ashcroft, Norman (1976). *Human territories: How we behave in space-time*. Englewood Cliffs, NJ: Prentice-Hall.
- Stenglin, M. (2002). Comfort and security: A challenge for exhibition design. In L. Kelly & J. Barrett (Eds), *UNCOVER: Volume 1* (pp. 17–22). Sydney: Australian Museum.
- Stenglin, M. (2004). Packing curiosities: Towards a grammar of three-dimensional space. PhD Thesis, University of Sydney.
- Stenglin, M. (2007). Making art accessible: Opening up a whole new world. *Visual Communication*, Special edition, *Immersion* 6(2), 202–213.
- Stenglin, M. (2008a). Binding: A resource for exploring interpersonal meaning in 3D space. *Social Semiotics*, 18(4), 425–447.
- Stenglin, M. (2008b). Olympism: How a Bonding icon gets its 'charge'. In L. Unsworth (Ed.), *Multimodal semiotics: Functional analysis in the contexts of education* (pp. 50–66). London/New York: Continuum.
- Stenglin, M. (2009a). Space odyssey: Towards a social semiotic model of 3D space. *Visual Communication*, 8(1), 35–64.
- Stenglin, M. (2009b). Space and communication in exhibitions: Unravelling the nexus. In C. Jewitt (Ed.), *Routledge handbook of multimodal analysis* (pp. 272–283). Oxford, UK: Routledge.
- van Leeuwen, T. (1991). The sociosemiotics of easy listening music. *Social Semiotics*, 1(1), 67–80.
- van Leeuwen, T. (1998). Textual space and point of view. Paper presented to the Museums Australia State Conference, Who sees, who speaks – voices and points of view in exhibitions, Australian Museum, 21 September.
- van Leeuwen, T. (1999). *Speech, music, sound*. London: Macmillan.
- van Leeuwen, T. (2005). *Introducing social semiotics*. London: Routledge.
- White, P. (1997). Death, disruption and the moral order: The narrative impulse in mass-media 'hard news' reporting. In F. Christie & J.R. Martin (Eds), *Genre and Institutions: Social processes in the workplace and school* (pp. 101–133). London: Cassell.
- White, P. (1998). Telling media tales: The news story as rhetoric. PhD Thesis, University of Sydney.

Chapter 5

Dealing with Musical Meaning: Towards an Embodied Model of Music

Edward McDonald
University of Auckland

Introduction: Approaching Musical Meaning via Analyst Talk

The question of musical meaning is one of the great practical and philosophical cruxes of the Western tradition especially since the rise of autonomous instrumental music in the eighteenth century broke the hitherto unquestioned links between musical performance and its verbal texts, and the propagation of the notion of absolute music in the nineteenth century detached music-making from its immediate social contexts. At the same time, however, whether from the viewpoint of what the medievals dubbed *musica theoretica*, or its less respectable cousin *musica practica*, the question of **what** music means, or **how** it means, paradoxically has been not so much raised as begged. Indeed such are the problems evoked by the notion of musical meaning, and how it relates to musical form, that a recent study explicitly drawing **on a social-semiotic model** (Halliday 1978, Hodge & Kress 1988, Kress & van Leeuwen 1996), van Leeuwen's *Speech, Music, Sound* (1999), deliberately declines to use the key social semiotic concept of metafunction in order to analyse various semiotic uses of the modality of sound, including music. van Leeuwen chooses **not** to adopt the so-called metafunctional hypothesis (Halliday 1967, 1978) whereby the expression plane of language is related to its interpretation plane(s), and through them to the social context, in terms of the three abstract generalized functions of ideational, interpersonal and textual meaning, concentrating instead on the materiality of sound, on the one hand, and its ideological implications, on the other. It is interesting to note how he justifies his decision, contrasting his earlier analysis of language and vision with that of sound and music:

The resources of sound simply did not seem as specialized as those of language and vision, and the mode of sound simply did not seem so clearly structured along metafunctional lines as language and visual communication. [In analysis] I always ended up feeling that a given sound resource

(say pitch or dynamics) was used both ideationally and interpersonally, or both ideationally and textually and so on. (van Leeuwen 1999:190)

van Leeuwen goes on to suggest that both the use and analysis of sound are perhaps not yet culturally developed enough to be susceptible to systematization – in his terms, they are still a ‘medium’ rather than a ‘mode’:

[T]he semiotics of sound cannot be approached in quite the same way as the semiotics of language and or of images. It is not, or not yet, a ‘mode’, and it has therefore not or not yet reached the levels of abstraction and functional structuration that (written) language and image have reached, as a result of their use in social crucial ‘design’ processes. (van Leeuwen 1999:192)

While the implications of this ‘argument from design’, and its relationship to analysis, are not clearly spelled out by van Leeuwen, it seems to me that what is needed for a social-semiotic treatment of any particular modality is a kind of triangulation between the analysis of its texts, the theoretical frameworks that have been applied to it, and the social meanings it has for its communities of users. It is not enough to have just one or two of these: the theoretical and social without the textual leaves the analysis ungrounded, with no way of understanding in detail how analysts have come up with their interpretations; the social and the textual without the theoretical traps analysts in the (unexamined) presuppositions of their commonsense (or ‘intuitive’) viewpoints; the textual and theoretical without the social makes analyses ultimately only personal ones – insightful, perhaps, but in the end only one individual interpretation.

The current chapter makes no claim to have achieved a comprehensive model of musical meaning of this kind within a social semiotic framework. What is attempted here is the more modest aim of what Mao Zedong referred to as ‘reactionary editing’: in other words, putting several different kinds of discourse about the phenomenon side by side and seeing what emerges from the mix. It seems to me there is an enormous untapped resource for analysts not just in ‘analyst talk’, as it were, but also in ‘audience talk’ and ‘performer-composer talk’: the latter is particularly useful because it is based on a (literally) hands-on experience of the modality – including the crucial – though again often ‘intuitive’ alas – sense of the probabilities involved: how common is a particular feature? What sort of regularities is it playing off? – in other words, how can it be contextualized logo-genetically?

This chapter makes a start on this kind of project with an analysis of selected analyst talk (a preliminary attempt at incorporating performer-composer talk is made in McDonald, 2010). It begins with an analysis of two recent textbooks dealing with music, van Leeuwen (1999) already discussed above, and Vella (2000), then moves on to the semiotic approach of Nattiez (1990), whose model explicitly includes the viewpoints of both the composer-performer (‘poietic’) and listener (‘esthetic’), ending up with the phenomenological

approach of Burrows (1990). The conclusion that emerges from this discursive journey is that the crucial factor required for any conceptualization of musical meaning is **embodiment**: in other words, that the ultimate locus of musical meaning must be the signifying human body (Thibault 2004).

Making Musical Meaning Accessible: The Approach from Textbooks

The pioneering work of Thomas Kuhn (1962) has alerted us to the role played by textbooks in establishing and reflecting the current consensus of what he called 'normal science'. In the humanities, or what in more equal contrast to the natural sciences may be called the '**semiotic sciences**' (Halliday 2005), that is, those branches of knowledge that deal with meaning rather than matter, there is less agreement about what constitutes the 'basics'. In these areas, textbooks are less likely to reflect a consensus, since on few, if any, issues can consensus really be said to exist. Textbooks here tend to be more like what Kuhn describes as the situation for the natural sciences prior to the Scientific Revolution of the seventeenth century: each textbook representing an individual attempt to 'start from the beginning', and lay down some basic principles that should guide the understanding of the phenomenon under discussion. An examination of textbooks in these areas, therefore, allows us to see what are taken as the main issues to be explained, even if the explanations themselves tend to vary widely.

The two textbooks examined here provide interestingly different perspectives on the problems of musical meaning. From a composer's point of view, Richard Vella's *Musical Environments* (2000) explores the listening experience in detail and uses exercises in improvisation and composition to analyse and reflect on how music utilizes sound in different ways. From a film maker's perspective, Theo van Leeuwen's *Speech, Music, Sound* (1999) places sound in a multimodal context, where it not only has linguistic and musical functions, but also works alongside images in order to express complex understandings of the world. Both textbooks can be seen as continuing the methodology of Canadian composer R. Murray Schafer (1977) who, in clear contradistinction to the 'absolute' tradition of musical description deriving from Hanslick (1976[1854]), puts music firmly back into the context of the whole sound environment and the lived experience of its listeners.

These textbooks began their lives as part of tertiary education programs which had a focus on interdisciplinarity, and both initially derive, coincidentally, from the same academic institution, Macquarie University in Sydney. The two books have, however, very different theoretical underpinnings, and approach the description of music from very different angles. Vella's treatment of music is based on constructions of time and space and their contribution to listening. It aims to avoid virtuosic performance definitions of music so that

a model for sound and music can be discussed on the same level and not based on hierarchies of ability. With a similar inclusive purpose, van Leeuwen's treatment places spoken language, music and sound effects on the same plane, and discusses commonalities of meaning and expression across all three modalities. Both, however, in effect avoid what for the analyst is the key question of musical meaning: how it is that particular patterns of sound **expression** relate to their different levels of **interpretation**.

Vella's book, *Musical Environments*, is based on a course he taught for some years in the School of Mathematics, Physics, Computing and Electronics, and is very much the embodied 'hands-on' and 'ears-on' approach of a performer-composer, as he explains:

[This book] emphasises improvisation, listening and composition as a way of developing both the creative and conceptual skills needed for music-making and understanding sound relationships . . . Part 1 of the book examines different applications of sound in space. The reader is required to listen to the environment, explore making sounds in space, appreciate the migration of sounds around geographical space, experience the psychoacoustic sensations of register shift . . . Part 2 . . . examines sounds in time . . . in terms of abstract perceptual models based on texture in which the placement of sounds within each model creates different listening strategies . . . analogous with the way individuals and communities express their relationships with each other and their environments. (Vella 2000:7)

van Leeuwen's *Speech, Music, Sound* derives from courses he taught first in the Department of Media Studies at Macquarie, and then at the London School of Printing, and in contrast to Vella's approach, takes more of the semiotic analyst's point of view:

This book tries to do on a *theoretical* level what many contemporary musicians, poets, film-makers, multimedia designers and so on, already do in practice (and what children have always done): *integrate* speech, music and other sound. It tries to foreground the integration of these three, rather than talk about their specifics, and to contribute to the creation of a vocabulary for talking about this integration, and for exploring its ramifications and potentials. Above all, it tries to make you listen. Listen to the city as though it was music and to music as though it was the city, or to speech as though it was music and to music as though it was speaking to you. This listening . . . is not always going to be easy . . . Not through our own fault we have, most of us, ill-educated ears. Re-educating them may take some effort. (van Leeuwen 1999:4)

van Leeuwen introduces here what can only be called a moralistic note, a prominent sub-theme in his book, which both presents and demonstrates a moral purpose, to 're-educate' our 'ill-educated' ears. In this, van Leeuwen is

in the tradition of writers on music such as composer R. Murray Schafer, and his near namesake Schaeffer, composer and theorist of *musique concrete* (Schaeffer 1967), as well as earlier examples such as the Futurists of the early twentieth century, one of whose purposes was to awaken our senses to the range of sounds around us, beyond what is usually thought of as 'music' as such.

Vella and van Leeuwen place their exploration of music in relation to quite distinct professional and academic enterprises. For Vella, himself a well-known composer experienced in film, theatre and concert music, his concerns are in the first instance those of a **composer**, **improviser** and then of the **listener** (Vella 2000:9):

This book has been conceived from a composer's point of view which is fundamentally a sensory relationship to the creation and ordering of sounds. As soon as sounds are placed next to each other, we as listeners automatically invent relationships between the sound events and therefore meaning.

Here Vella gives a preliminary definition of musical meaning as 'relationships between sounds' and locates such relationships in the organizing ear of the listener. He follows this by placing his exploration of musical organization within the field of **music cognition**:

The perception of texture and musical events falls within the domain of music cognition, a complex field of study drawing together the two disciplines of music and psychology. It is largely concerned with the way we, as listeners, perceive musical structures, differentiate and organise sonic information, remember, predict and reject musical events, internalise larger formal structures, and create relationships between sounds. (Vella 2000:10)

van Leeuwen's background is in film making and jazz performance, as well as in linguistics. His textbook, in a lengthy explanation from which I have extracted the main points, brings the concerns of **user** and **analyst** together under the banner of **semiotics**:

[T]his book is about semiotics. But what is semiotics? Or rather, what do semioticians do? Three things, I think. Semioticians describe the semiotic *resources* people use in communication . . . the semiotics of sound concerns itself with describing what you can 'say' with *sound*, and how you can interpret the things other people 'say with sound' . . . Describing semiotic resources provides the means for describing and explaining how these resources are actually used . . . when I try to formulate the semiotic value of the 'choices', I do not provide a *code*, with definite and fixed meanings, but a *meaning potential* which will be narrowed down and coloured in the given context . . . There is yet another contribution semioticians can make: they are particularly well placed to explore how semiotic resources can be *expanded*, so as to

allow more options, more tools for the production and interpretation of meaningful action . . . (van Leeuwen 1999:4–10)

Although, as noted above, neither book sets out to formally define the concept of the **meaning** of music, each does nevertheless provide a clear working definition of musical meaning for their very different purposes. For Vella, in line with his contextual ‘hands-on’ approach, meaning is created through the listener’s experience of sound:

As we listen to sounds, we allocate meanings to them. We need to do this for them to make sense. However, a sound might have a completely different meaning to two different people. This is why context and our relationship to the sound event are important. The process of listening has 3 aspects:

1. the music itself
2. its context and
3. its meaning.

The music itself includes all its auditory qualities; its context is defined by where or how the music is positioned in relation to the listener and its purpose; listening to music through a pair of headphones, for example, is a very different experience from hearing it in a concert hall; and the meaning of music is determined by who is listening and the cultural experiences and associations of its audience. (Vella 2000:24)

Such an approach is in direct contradistinction to a long tradition of ‘absolutist’ approaches to musical meaning as put forward by theorists like Hanslick (1976[1854]) and composers like Stravinsky, or more recent formalist descriptions of music such as those of Lerdahl and Jackendoff (1983). In contrast to such approaches, which are highly influential in musicology, Vella is concerned to stress how music does not ‘contain’ meaning in itself, but rather takes on meanings from its contexts of performance. From the point of view adopted here, Vella’s approach stresses the **social** aspect of music-making, and how it takes on significance for its users, whether performers, composers or listeners, in its specific contexts of performance. Given the audience for which this textbook was originally developed, general education students from across the whole university with an enthusiastic performing interest in music, such an approach can also be seen as a way of stressing the grounded nature of musical meaning and therefore the accessibility of music as a social activity to all members of a society, not just the ‘professionals’.

van Leeuwen also stresses the different contexts of music, but does so from a text-based **semiotic** viewpoint. His book aims to set out the ‘semiotic resources’ available in music, representing these resources by way of a ‘system network’, both concept and formalism deriving from systemic functional linguistics

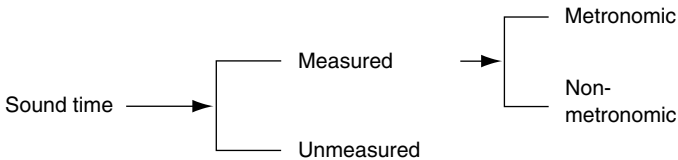


FIGURE 5.1 Fragment of a system network for sound

(Halliday 1978, Halliday & Hasan 1985), a framework within which van Leeuwen, wearing his linguist's hat as a phonologist, has himself worked. Figure 5.1 illustrates his representation of the semiotic resources available in the area of musical timing (van Leeuwen 1999:6–7):

This fragment of a system network is concerned with a particular sound 'resource'. It tries to map how sound events can be structured in terms of their timing . . .

van Leeuwen thus represents musical meaning in terms of the meaning-making resources available to members of a musical community, but in a distinct further step, seems to understand musical meaning largely in terms of the **ideological** significance of particular sound features:

Measured time is time you can tap your feet to . . . The physical reaction to unmeasured time is more likely to be a slow swaying of the body . . . ' . . . metronomic' and 'non-metronomic' time form a subdivision of 'measured time'. 'Metronomic time' is governed by the implacable regularity of the machine, whether or not a metronome (or a drum machine or a stopwatch) is actually used. It is the time of the machine, or of soldiers on the march. 'Non-metronomic time' is also measured, but it subverts the regularity of the machine. It stretches time, it anticipates or delays sounds and so on. It is the time of human speech and movement, or of Billie Holliday singing a slow blues while 'surfing on the beat' . . .

It is, in a sense, unfair for textbooks to be subjected to this kind of analysis: treating them, in effect, as the type of academic monographs they were never intended to be. However, apart from the fact that such a procedure allows us to highlight the problems attendant on defining such a slippery concept as musical meaning, it also sensitizes us to the fact that *all* such presentations, whether intended for a more popular or a more technical audience, have specific goals, goals which inevitably shape their treatment of the substantive issues they address. The following sections will go on to examine some more technical treatments of musical meaning in order to gain a clearer idea of what is at stake in putting forward theories of musical meaning.

A Typology of Stances Towards Musical Meaning

Musical semiotician Jean-Jacques Nattiez's *Music and Discourse: Towards a Semiology of Music* (1990) is a more technical kind of 'textbook', intended not for the interested practitioners – in the widest sense of that term – who form Vella's and van Leeuwen's main audience, but for those with a professional interest in the analysis of music. Nattiez's work first appeared in French as *Musicologie generale et semiologie* [General Musicology and Semiotics] (1987), which in turn was a revision and extension of an earlier work *Fondements d'une semiologie de la musique* [Foundations of a Semiotics of Music] (1975). The title of this early work points clearly to what was identified above as the textbook-like aim of 'starting from the beginning', an aim which is continued in his later works.

The main title of Nattiez 1990, *Music and Discourse*, gives the hint as to his approach, a version of what I referred to in a previous paper as the 'identification of music with language' and of 'linguistics with musicology' (McDonald 2005). Nattiez takes this project further, and treats it more explicitly, than do most such accounts, undertaking a comprehensive critique of the whole notion of music as a 'language'. Allied to his analysis of what he calls the 'semiology of the musical fact' (1990: Part I), is a detailed critique of the 'semiology of the discourse of music' (1990: Part II); and this two-pronged approach performs an enormously valuable task by allowing him to set out the linguistic and musicological assumptions that usually lurk, more or less unexamined, beneath the surface of most accounts of how music works. Nattiez's exegesis thus provides a handy way of navigating and contextualizing the large body of literature in this area.

Nattiez starts his treatment of musical meaning by taking over a characterization first put forward by Leonard Meyer (1956) (see Figure 5.2):

For Meyer, there are on one side *absolutists* who believe that meaning is based exclusively on the relationships between the constituent elements of

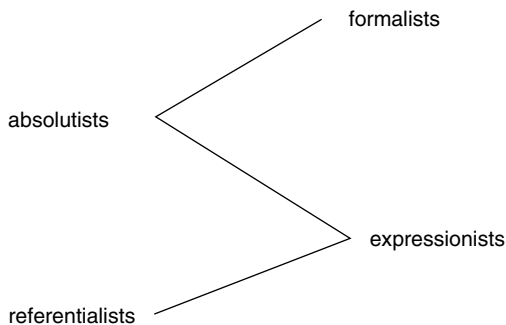


FIGURE 5.2 Stances towards musical meaning

the work itself, and on the other *referentialists* for whom there cannot be meaning in music, except by referring to an extramusical universe of concepts, actions, emotional states, and characters (1956:1). But this first dichotomy is mirrored in another than does not exactly correspond to it: the *formalists* who (according to Meyer) do not acknowledge that music can provoke affective responses (it has an intrinsic significance given to it by the play of its forms), and the *expressionists* who acknowledge the existence of feelings. But though formalists are necessarily absolutists, expressionists will be absolutists if (for them) the expression of emotion is contained in music itself, and they will be referentialists if the expression is explained in terms of music referring to the external world. (Nattiez 1990:108–109)

Nattiez then combines this typology with his own basic theoretical framework of what he calls the ‘semiotic tripartition’, 3 ‘dimensions’ of any ‘symbolic phenomenon’, which he defines as follows:

1. *The poietic dimension*: even when it is empty of all *intended* meaning . . . the symbolic form results from a process of creation that may be described or reconstituted.
2. *The aesthetic dimension*: ‘receivers’, when confronted by a symbolic form, assign one or many meanings to the form; the term ‘receiver’ is, however, a bit misleading . . . we do not ‘receive’ a ‘message’s’ meaning . . . but rather construct meaning, in the course of an active perceptual process.
3. *The trace*: the symbolic form is embodied physically and materially in the form of a trace accessible to the five senses . . . An objective description [of the trace] can always be proposed – in other words an *analysis* of its immanent and recurrent properties. This is referred to . . . as ‘analysis of the neutral level’. (Nattiez 1990:11–12, original emphasis)

What Nattiez is attempting here, very ambitiously, is to bring together three major traditions in musicology: the composer-focused study of the compositional process (‘poietic’); the listener-focused study of audience reception (‘esthetic’); and the text-based analysis of musical structure (‘neutral’). Arguing against a naïve type of intentionality whereby the composer would be understood as attempting to ‘communicate’ his or her meaning to listeners, Nattiez assigns equal weight to the listener’s interpretative role as to the composer’s creative function. However, with no explicit theorization of the social context at all, and with composer and listener understood as relating to musical patterning – significantly characterized as ‘neutral’ to composer or listener concerns – in quite distinct ways, Nattiez remains unable to characterize where exactly musical meaning should be understood as residing: in terms of the current framework, what exactly is the nature of the relationship(s) between musical expression and musical interpretation.

In the light of Meyer's classification, Nattiez then goes on to treat what he calls the 'two extremes' of thinking on 'musical aesthetics', in other words, stances towards musical meaning, which he sums up as follows:

- The formalist-absolutist position: music means itself.
- The expressionist-absolutist position: music is capable of referring to the non-musical. (Nattiez 1990:110)

We will now use this classification, and the theorists Nattiez quotes, to give a brief guided tour through a range of stances on musical meaning.

Formalist-absolutist Stances to Musical Meaning

The *locus classicus* of absolutist theories of music is undoubtedly Hanslick's (1976[1854]) treatise *On the Beautiful in Music*, whose main thrust may be summed up in Hanslick's own words as follows:

. . . the beauty of a musical work is specifically musical – i.e. it inheres in the combinations of musical sounds and is independent of all alien, extramusical notions. (Hanslick 1976[1854]:12)

In assessing Hanslick's position, and stances towards musical meaning more generally, Nattiez, puts forward a useful distinction between **empirical** claims and **normative** positions (Nattiez 1990:109):

In order to understand Hanslick's position, we need to distinguish what is for him an *empirical* claim: 'far be it from us to underrate the deep emotions which music wakens from their slumber' (Hanslick 1976[1854]:26), and a *normative* position: 'if contemplation of something beautiful arouses pleasurable feelings, this effect is distinct from the beautiful as such'. (Hanslick 1976[1854]:18)

Nattiez then uses this distinction between empirical and normative, together with his notion of the semiotic tripartition referred to in the previous section, to carry out an admirably clear dissection of Hanslick's complex position:

The empirical point of view:

- (1) from the poietic side, emotion exists in the composer, but does not manifest itself except in a purely musical form;
- (2) on the immanent level, music's content is its form;
- (3) from the esthetic side, emotion is the result of the form's effect, and its origin must be sought in the music itself

The normative point of view:

- (1) Poietic: one should not write program music, or imitative or sentimental music. In opera, music should occupy the predominant position.
- (2) Immanent level: 'the Beautiful is nothing more than form' (Hanslick 1976[1854]:16).
- (3) Esthetic: perception is not exempt from emotions, but it must try to elevate itself to pure contemplation of forms. (Nattiez 1990:109–110)

Without going into a detailed discussion of Hanslick's arguments, it does I think need to be acknowledged that the problems of defining musical meaning give a much firmer basis to arguments *against* music having any meaning external to itself than is the case with language. Here is musicologist Victor Zuckerkandl grappling with the problem of musical 'indicating' or 'pointing':

Tones too [like words EMcD] indicate, point to something. The meaning of a tone, however, lies not in what it points to but *in the pointing itself*, more precisely, in the different way, in the individual gesture, with which each tone points toward the same place. The meaning is not the thing indicated but the manner of indicating (otherwise all tones would mean the same thing, namely, $\hat{1}$ [tonic EMcD]) . . . In the strictest sense . . . what the tone means is actually and fully contained in the tone itself. Words lead away from themselves; but tones lead into themselves. Words only point toward what they mean, but, beyond that, leave it, so to speak, where it is . . . Tones, on the other hand, have completely absorbed their meaning into themselves and discharge it upon the hearer directly in their sound. (Zuckerkandl 1956:67–68, original emphasis)

Zuckerkandl's claim rests on his identification of differing degrees of 'tension' between musical tones, with the 'tonic' functioning as the point of rest and return for all the other tones: thus the 'meaning' of each tone consists in its relationship to the tonic.

If this explains the underlying harmonic logic of music, at least in tonal music, linguist and semiotician Roman Jakobson takes a more textual approach, considering how the different parts of the musical text hang together:

. . . instead of aiming at some extrinsic object, music appears to be *un langage qui se signifie soi-même* ['a language which signifies itself' EMcD]. Diversely built and ranked parallelisms of structure enable the interpreter of any immediately perceived musical signans [signifier] to infer and anticipate a further corresponding constituent . . . and the coherent ensemble of these constituents. Precisely this interconnection of parts as well as their integration into a compositional whole acts as the proper musical signatum [signified]. (Jakobson 1970:12)

In other words, in order to understand what music means, one simply needs to be able to identify relationships between the different parts of the musical text, and link them into a coherent whole. Linguist Robert Austerlitz makes a similar point, using the notion of pointing or ‘deixis’, normally used of linguistic elements which refer outside language to aspects of the situation. In the case of music, according to Austerlitz, any musical deixis is text-internal or ‘cataphoric’, that is, referring forward to subsequent musical patterns:

[T]he meaning that is conveyed by a musical text is basically deictic, cataphoric, in the sense that it is *prediction*. The musical text makes reference to the future, in that it challenges the listener to predict the shape of the musical substance to come in the immediately impending future – on the basis of the musical substance perceived in a given moment. (Austerlitz 1983:4)

Musicologist Nicolas Ruwet, one of the earliest scholars to apply the then new ideas of transformational-generative grammar to the analysis of music, basically updates Hanslick by claiming that it is only an analysis of the ‘syntax’ of music, that is, its internal patterns of organization, that can give us any ‘access’ to the ‘study of musical meaning’:

Music’s meaning cannot manifest itself except in descriptions of music in itself . . . the signified (the ‘intelligible’ or ‘translatable’ aspect of the sign) is, for music, conveyed by the description of the signifier (the palpable aspect). Our only means of access to a study of musical meaning is, indeed, a formal study of musical syntax, and a description of the material aspect of music on all levels where music has a real existence. (Ruwet 1967:91)

However, as Nattiez comments, the notion of music ‘pointing to itself’ still doesn’t completely account for the effect of musical form on the listener. For this, some kind of ‘expressionist’ stance is necessary.

Expressionist-absolutist Stances Towards Musical Meaning

In fact, as noted by Nattiez, even those like Ruwet who are concerned to argue for a kind of ‘musical autonomy’ in relation to meaning, often end up invoking some basis for musical meaning in *non*-musical experience:

Linguists, structuralists, and practitioners of generative grammar have taught us that internal examination of a work is more important than examinations of its psychological or physiological circumstances . . . analysis . . . would enable us to entangle those musical structures that are homologous with other structures, those arising from reality or lived experience; it is in this

homological correspondence that the 'sense' of a musical work in unveiled . . . Suppose there is a fragment of tonal music made up of two parts, A and A'; A ends with an interrupted cadence, and A' begins the same way but ends in a perfect cadence. Within the framework of the tonal system, the first part will obviously be interpreted as a movement, directed toward a certain point, but interrupted or suspended; the second as repetition of the same movement, this time continued to its end. (Ruwet 1972:13–14)

Ruwet's notion of a 'homological correspondence', whereby, for example, an 'interrupted cadence' will be interpreted as 'suspended' movement, as opposed to the a 'perfect cadence' in which the movement is 'continued to its end', locates at least part of the meaning of music, albeit at a very abstract level, in a relationship between 'musical structures' and 'other structures . . . arising from reality or lived experience'.

A more recent attempt to grapple with this relationship points out that some aspects at least of musical meaning are ineffable, and thus cannot be captured in language, and draws a useful distinction between the portion of musical meaning that is 'articulated' and thus explicitly describable, and that which is simply 'presented':

. . . part of the meaning of a work of art music is articulated in the very structure of the work taken as a 'significant form', whereas another part of its meaning is merely presented, but not articulated. In other words, each work of art music presents some meaning in an articulated form and about this meaning we can speak. But through this meaning it also makes present a certain 'excess of meaning' which cannot be articulated . . . [but which] reaches far beyond the meaning that is actually present in the work's significant form. This excess of meaning refers to a world. This world cannot be that of the artist, because then the work would merely address the artist's contemporaries. It cannot be the world of the listeners either, because this world changes considerably over time. This world is rather a 'universally' livable world, a world that *in principle* is a world of possibilities for everyone. Yet this world can be presented only through concrete presentations of the work. (Kockelmans 1999:184–185)

So what exactly is this 'world' that music 'presents'? The term 'presents' suggests that what is going on in music is some sort of 'performance', not just in the obvious sense that music is a performance art, but in some deeper sense that involves the music and all the participants in some sort of shared experience. However, this still leaves unproblematised the exact nature of musical presentation, and how it is that the participants interpret what the 'performance' means. Thus, we seem to be pushed towards attempting to define some sort of notion of how music 'refers' to something outside itself: in a broad sense, how it is that the 'world of music' relates to the 'world of lived experience'.

Expressionist-referentialist Stances Towards Musical Meaning

One answer to this question is given by philosopher Stephen Davies, who puts forward a theory of music's expressiveness which depends on an analogy with human facial appearance, body language and gait, whereby certain 'emotion characteristics' are publicly displayed and can thus be interpreted by an onlooker 'in appearances':

Our experience of musical works and, in particular, of motion in music is like our experience of the kinds of behavior which, in human beings, gives rise to emotion characteristics in appearances. The analogy resides in the manner in which these things are experienced rather than being based on some inference attempting to establish a symbolic relation between particular parts of the music and particular bits of human behavior. Emotions are heard in music as belonging to it, just as appearances of emotions are present in the bearing, gait, or deportment of our fellow humans and other creatures. The range of emotions music is heard as presenting in this manner is restricted, as is also true for human appearances, to those emotions or moods having characteristic behavioral expressions: music presents the outward features of sadness or happiness in general. (Davies 1994:239)

The analogy works here via a long tradition of identification of the 'motion' of the body with the 'emotion' of the feelings (the terms themselves, derived from Latin for 'to move' and 'to move out of', respectively, show how this metaphor is embedded in many Western European languages at least). Davies argues that just as we can interpret how someone is feeling by the way they move, using 'move' in the broadest possible sense to include facial expressions and gait, we can also interpret musical form by 'establishing a symbolic relation between particular parts of the music and particular parts of human behaviour'. Thus, music 'expresses' by 'referring', by presenting symbolic forms which can be interpreted as expressing emotions.

A complementary approach to explaining what music refers to is taken by Niall Griffith, who again stresses the performative, dynamic nature of music, in this case in contrast to the analytic, static nature of language:

Music while [like language EMCD] involving categorial perception . . . does not attach meaning to the arbitrary signs that it identifies. While language breaks down, analyses and labels and separates itself from the flow – moulding experience *with the pale cast of thought*, music . . . adopts a complementary strategy to language – allowing the use of sound to represent change and causation by using it in direct physical metaphors that maintain their origin in action and change. (Griffith 2002:200)

Both these types of 'referring', the emotional and the causational, can be brought together in a model put forward by Michel Imberty in his notion of

'dynamic vectors', by which an 'affect', an emotion, is linked to a 'vector', a musical element carrying significations of change:

The notion corresponding to that of vitality affect in music is undoubtedly what, on the basis of experiments on the semanticization of musical experience, I suggested characterizes the dynamic and temporal aspects of forms: it is the notion of a dynamic vector. Dynamic vectors are musical elements that transport temporal significations of orientation, progression, diminution or growth, and repetition or return. Perceived and felt change is thus a dynamic vector that orients the listener's perception, anticipation, and internal representations. The quality of orientation depends on what the dynamic vector refers to, assimilated here to the set of vitality affects that the subject experiences or relives immediately in listening. (Imberty 2000:459)

Thus, music operates through elements of 'perceived and felt change' that refer to 'vitality affects', emotional responses that have been developed on the basis of previous experiences. Ruthrof's *The Body in Language* (2000) puts forward a similar model for language, in which he attempts to put bodily experience back into theories of language via a process of psychological imagery:

In itself . . . language as an ordered sequence of words is . . . empty. It is mere syntax, mere sequences of words. Only when language is combined with something other than linguistic signs is it able to mean. This Other of language is not the world as a set of unmediated data, but rather a fabric of [embodied EMCD] nonverbal signs out of which cultures weave the world the way they see it. (Ruthrof 2000:31)

So it begins to seem as though it is the **body** that must be the ultimate locus of attempts to ground musical meaning in something external to itself. In the final section, I go on to deal with this notion of embodiment in relation to music and show how it makes us rethink the whole nature of musical experience and musical meaning.

The Locus of Musical Meaning: The Signifying Body

In a book produced almost a decade before van Leeuwen and with almost the same title, musicologist David Burrows in his *Sound, Speech, and Music* (1990) sets out to understand the characteristic uses of sound in language and music in a way that stems from its fundamentally embodied nature. Burrows emphasizes the semiotic affordances provided by sound, claiming that 'sound is far more to speech than a passive conveyance' but rather the means through which human thought has evolved, exploiting 'the unique capacity of vocal sound for rapidity of articulation in detachment from the world of enduring spatial

objects' (1990:9). Burrows develops a 'three field' model of body-mind-spirit in order to understand sound both as a basic element of human perception, and, in its paradigmatic form of the human voice, as something which human beings themselves produce in ways that are meaningful to their fellows:

[T]he voice is . . . the most intimate and powerful human exploitation of the potential in sound, a means of displaying mood and attitude and a way of bonding separate individuals and negotiating their mutual interests . . . If speech is a displacement of the mutual awareness of speaker and listener from Field 1, the here and now conveyed by the senses, and into the metasensory domain of Field 2, then the speaking self is defined by its relationship to shifting possibilities outside the actuality of the moment, possibilities at best indirectly verifiable. This means that a corresponding quality of contingency and provisionality must characterize that range of identity which is at the focus of speech and speech-related thought. Music is seen . . . as one of a range of activities that help compensate for this debilitation of identity by moving the participants' orientation towards that of Field 3. (Burrows 1990:11–13)

Burrows' model provides a rationale for the existence of music in every human society in a way that allows for its extra-bodily dimensions but does not detach them from the fundamentals of bodily experience.

From a philosopher's point of view, Stephen Davies in his *Musical Meaning and Expression* (1994) gives a long discussion of various attempts within philosophy and musicology to capture how it is that music expresses meaning. His very wide-ranging study puts forward, among others, the very useful concept of music expressing 'emotion characteristics in appearance' already referred to above. But although he spends almost the whole book trying to define 'meaning', he takes the equally slippery concept of 'music' as a given, and this causes him enormous problems: for example, in trying to justify how a 'non-sentient' phenomenon like music can have 'feelings' (1994:163).

In fact as musicologist Christopher Small points out, such questions are basically pseudo-questions: '[t]here is no such thing as music':

Music is not a thing at all but an activity, something that people do. The apparent thing 'music' is a figment, an abstraction of the action, whose reality vanishes as soon as we examine it at all closely . . . If there is no such thing as music, then to ask 'What is the meaning of music?' is to ask a question that has no possible answer. Scholars of Western music seem to have sensed rather than understood that this is so; but rather than directing their attention to the activity we call music, whose meanings have to be grasped in time as it flies and cannot be fixed on paper, they have quietly carried out a process of elision by means of which the word *music* becomes equated with 'works of music in the Western tradition'. Those at least do seem to have a real existence, even if the question of *how* and *where* they exist does create

problems. In this way the question 'What is the meaning of music?' becomes the more manageable 'What is the meaning of this work (or these works) of music?' – which is not the same question at all. (Small 1998:2–3)

Both Nattiez and Davies fall into this trap of trying to locate the meaning of music in particular musical works, a trap that has been carefully prepared for them, and many others, by a whole tradition of thinking about music in the European cultural sphere with the development over the last millennium of ever more sophisticated and comprehensive forms of musical notation. Having the music 'in black and white' on the page not only means that the focus of attention then becomes interpreting and explaining the notation, but that music quite naturally comes to be seen as an object, removed from its immediate context, and thus amenable to 'scientific' study. Small quotes 'the doyen of contemporary German musicologists, Carl Dalhaus' in a pithy summary of this attitude in which he states that 'the concept "work" and not "event" is the cornerstone of music history' (Dalhaus 1983, in Small 1998:4).

Thus, Nattiez's long discussion of the musical fact and Davies' equally long struggle with musical meaning both suffer from a type of misplaced concreteness, because they both take it for granted that such a thing as 'music' exists, and that it exists in musical 'works', not musical 'events'. Small's study attempts to redress this imbalance, by looking not at this pseudo-object 'music' or the musical work, but at the musical event, a process to which he gives the name 'musicking'. The fact that he does this by a detailed ethnographic study of the Western concert hall and what goes on there is a nice riposte to the tradition represented by Dalhaus which places actual performance outside the ken of observation or theorization.

And as soon as we start focusing on performance, as soon as we bring back living breathing people into our conception of music, more specifically the musical event, we are also perforce required to focus on people's **bodies**. The notion that the body is central to musical meaning is not a new one, as the analogy between 'motion' and 'emotion' shows, though overshadowed by another long tradition in European thinking about music, dating back at least to Pythagoras, which places music in some sort of abstract, celestial realm beyond our everyday mundane lives. But the relationship between bodily movement and musical expression was already being emphasized by French psychologist Frances half a century ago:

The kinship between rhythmic and melodic pattern in music and the patterns of gestures that accompany behaviour, represents one of the basic elements of music's expressive language . . . The basic psychological states (calm, excitation, tension, relaxation, exaltation, despair) normally translate themselves as gestural forms that have a given rhythm, as tendencies and ascents, as modalities for organizing fragmentary forms within global forms . . . the transpositions of these rhythms, tendencies, and modalities of movement

into the sound structure of music constitutes music’s basic expressive language. (Frances 1958:299, quoted in Nattiez 1990:118–119)

Even earlier, Swiss music educator Jacques-Dalcroze was stressing the inherent link between movement and music by identifying the crucial links between time, space and muscular energy:

Musical rhythm consists of linking up durations, geometry consists of linking up fragments of space, while living plastic movement links up degrees of energy . . . Any movement we have to perform in a given tempo requires further muscular preparation if we wish to repeat it in a different tempo. A line traversed by a limb in a given space and time becomes shorter or longer according to the degree of energy required to make the movement. A duration of time occupied by a limb moving at a given rate of muscular energy becomes prolonged or shortened according to the length of space to be traversed. (Jaques-Dalcroze 1930:10–11)

In fact the whole body of theory and practice of music learning which stems from Dalcroze’s work, known as ‘Eurythmics’, works precisely by maintaining the link between body movement (‘rhythm’) and aural training at every step of the music learning process. Table 5.1, simplified and adapted from a

Table 5.1 Embodiment and cognition in music learning

LISTENING		
Enactive mode through action & manipulation	Affective (emotional) Cognitive (intellectual)	Iconic Mode through organisation & imagery – aural, visual, tactile & kinaesthetic
MOVING		
Start and Stop natural movement – adjustment, flexibility, energy, elevation	Coordination of mind and body	Momentum preparation of the body – control, relaxation, stamina
MUSICAL DEVELOPMENT		
Time / Space / Energy	Motivation Communication	Phrasing and Form
SYMBOLISING		
Memory	Symbolic mode – through words & notation Decision making Self Discipline Time / Space / Energy	Concentration

Source: Vanderspar 1984/1992:25–26

pedagogical work in this tradition, shows how such a link can be made in terms of a process of music learning that involves listening, movement, cognition and symbolization (Vanderspar 1984/1992:25–26).

The sort of detailed repertoire of practice as developed by Eurythmics, allied to an embodied conception of musical meaning such as put forth by Frances, could begin to really get to grips – to use another embodied metaphor – with the nature of musical meaning as grounded in the embodied context of human semiosis more generally (Thibault 2004). Such a model has been adumbrated in several publications by myself in collaboration with a voice expert colleague (Callaghan & McDonald 2002, 2003, McDonald 2002, forthcoming), but the full working out of a social semiotic model of music, including both text analysis and ethnographic studies, remains in the form of a promissory note. The present chapter, through its critique of a range of instances of analyst talk, has hopefully managed at least to point one possible way out of the haze of misconceptions with which the topic has been obscured. So it is perhaps fitting to end this patchwork of quotations with a self-quotation which sums up the general attitude taken here towards this complex topic:

The human animal uses its body to dance and sing and move and speak, to model and (re)enact the processes and interactions of its material and social worlds, as well as to create verbal and musical texts that embody (pun intended) its semiotic worlds. How much longer can musicologists – or linguists for that matter – ignore the fact that they are dealing with an embodied social-semiotic system? (McDonald 2002:305)

References

- Austerlitz, R. (1983). Meaning in music: Is music like language and if so, how? *American Journal of Semiotics*, 2(3), 1–12.
- Burrows, D. (1990). *Sound, speech, and music*. Amherst, MA: The University of Massachusetts Press.
- Callaghan, J. & McDonald, E. (2002). Expression, content and meaning in language and music: An integrated semiotic approach. In P. McKeivitt et al. (Eds), *Language, Vision, Music* (pp. 205–230). Amsterdam: John Benjamins.
- Callaghan, J. & McDonald, E. (2003). The singer's text: Music, language and the expression of meaning. *Australian Voice*, 9, 42–48.
- Dalhaus, C. (1983). *Foundations of Music History* (J.B. Robinson, trans.). Cambridge and London: Cambridge University Press.
- Davies, S. (1994). *Musical meaning and expression*. Ithaca, NY: Cornell University Press.
- Frances, R. (1958). *La perception de la musique*. Paris: Vrin.
- Griffith, N. (2002). Music and language: Metaphor and causation. In P. McKeivitt, C. Mulvihill & S.O. Nuallain (Eds). *Language, Vision & Music* (pp. 191–220). Amsterdam: John Benjamins.
- Halliday, M.A.K. (1967). Notes on Transitivity and Theme in English, Part 1. *Journal of Linguistics*, 3(1), 37–81.

- Halliday, M.A.K. (1978). *Language as social semiotic*. London: Edward Arnold.
- Halliday, M.A.K. (2005). On matter and meaning: The two realms of human experience. *Linguistics and the Human Sciences*, 1(1), 59–82.
- Halliday, M.A.K. & Hasan, R. (1985). *Language, context, and text: Aspects of language in a social-semiotic perspective*. Geelong: Deakin University Press.
- Hanslick, E. (1976[1854]). *On the musically beautiful* (G. Payzant, trans.). Indianapolis, IN: Hackett Publishing.
- Hodge, B. & Kress, G. (1988). *Social semiotics*. Cambridge: Polity.
- Imberty, M. (2000). Innate competencies in musical communication. In L. Nils Wallin, Bjorn Merker & Steven Brown (Eds), *The origins of music* (pp. 449–462). Cambridge, MA: The MIT Press.
- Jakobson, R. (1970). Language in relation to other communication systems. *Linguaggi nella società e nella tecnica* (pp. 3–16). Milan: Edizioni di Comunità.
- Jacquès-Dalcroze, E. (1930). *Eurythmics, art and education*. Reprinted 1976. New York: Arno Press.
- Kockelmans, J.J. (1999). Why is it impossible in language to articulate the meaning of a work of music? In Kemal, Salim & Ivan Gaskell (Eds), *Performance and authenticity in the arts* (pp.175–194). Cambridge: Cambridge University Press.
- Kress, G. & Van Leeuwen, T. (1996). *Reading images – The grammar of visual design*. London: Routledge.
- Kuhn, T. (1962). *The structure of scientific revolutions*. Chicago, IL: University of Chicago Press.
- Lerdahl, F. & Jackendoff, R. (1983). *A Generative Theory of Tonal Music*. Cambridge, MA: MIT Press.
- McDonald, E. (2002). Through a linguistic glass darkly: A critique of theories of music from a social semiotic perspective. In C. Stevens et al. (Eds), *Proceedings of the 7th International Conference on Music Perception and Cognition, Sydney 2002* (pp. 303–306). Adelaide: Causal Publications.
- McDonald, E. (2005). Through a glass darkly: A critique of the influence of linguistics on theories of music. *Linguistics and the Human Sciences*. 1(3), 463–488.
- McDonald, E. (2010). Creating the classical gig: Exploring new contexts and values for performing classical music. MMus Thesis, School of Music, University of Auckland.
- Meyer, L. (1956). *Emotion and meaning in music*. Chicago, IL: University of Chicago Press.
- Nattiez, J. (1990). *Music and discourse: Towards a semiology of music*. Princeton, NJ: Princeton University Press.
- Ruthrof, Horst (2000). *The body in language*. London: Cassell.
- Ruwet, N. (1967). Musicologie et linguistique. *Revue internationale des sciences sociales*, 19(1), 85–93.
- Ruwet, N. (1972). *Langage, musique, poésie*. Paris: Seuil.
- Schafer, R. (1977). *The tuning of the world*. Toronto: McLelland & Stewart.
- Small, C. (1998). *Musicking*. Hanover, NH & London: Wesleyan University Press.
- Thibault, P.J. (2004). Brain, mind and the signifying body: An ecosocial semiotic theory. London: Continuum.
- van Leeuwen, T. (1999). *Speech, music, sound*. London: Macmillan.

- Vanderspar, E. (1984). *Dalcroze handbook: Principles and guidelines for teaching eurhythmics*. Updated edition 1992. Launceston, Cornwall: The Dalcroze Society.
- Vella, R. (2000). *Musical environments: A manual for listening, improvising and composing*. Sydney: Currency Press.
- Zuckerkandl, V. (1956). *Sound and symbol: Music and the external world*. Princeton, NJ: Princeton University Press.

This page intentionally left blank

Part Three

Intermodality between the Visual, Verbal and Aural

This page intentionally left blank

Chapter 6

Organizing Visual Meaning: FRAMING and BALANCE in Picture-Book Images

Clare Painter

University of Sydney

J.R. Martin

University of Sydney

Len Unsworth

University of New England

Introduction

The social semiotic analysis of visual texts has made considerable progress in the past decade since the publication of Kress and van Leeuwen's (2006) *Reading Images: The Grammar of Visual Design*, which makes use of M.A.K. Halliday's (1978) theory of 'metafunctions' to identify three distinct but coexisting kinds of meanings that interplay within any text. This chapter aims to develop further the social semiotic analysis of visual images within one of these metafunctions and in relation to one particular source of data – a corpus of children's narrative picture books. The region of meaning under focus is that of the 'textual' metafunction (Halliday 1978; Halliday & Matthiessen 2004) or 'composition' (Kress & van Leeuwen 2006), within which a number of visual choices will be identified and their meanings discussed. Picture book narratives have the advantage as data for visual analysis that they are 'apprenticing' texts in terms of visual as well as verbal literacy, thus making the relevant visual choices salient as they guide readers/viewers into an understanding of the meanings made.

Halliday's notion of metafunctions as regions of meaning has been developed as part of systemic-functional linguistics for the explanation and analysis of verbal texts. Whereas the 'ideational' metafunction is concerned with the content or topic of a text, and the 'interpersonal' metafunction with attitudes, stances and relations of power and social distance between reader and writer (or between characters in a fictional work), the 'textual' metafunction is said to be concerned with the organization of both ideational and interpersonal meanings. On the one hand the textual metafunction of language involves 'cohesion' by such means as ellipsis, lexical chains or pronominal reference,

which create links across different parts of a text, while on the other hand it concerns the **staging and packaging of ideational and interpersonal meanings by such means as choice of initial clause or paragraph element (Theme) and the organization of given and new information.** It is with the visual equivalent of this latter aspect of meaning that the current chapter will be concerned. That is, in considering the visual systems of textual or compositional choices found within picture books, the focus will not be on colour 'rhymes' or visual repetitions that achieve cohesion across the narrative, but rather on the way visual elements are 'packaged' on the page, on questions of the separation or integration of elements and the training or direction of attention.

It is proposed here that the textual, or compositional, metafunction as it applies to children's picture-book images principally involves three systems, or **sets of options: those of FRAMING, BALANCE and INTERMODAL INTEGRATION, the first two of which will be described and discussed in this chapter.** These systems have been inferred from an examination of a corpus of over 50 narrative picture books including many prize-winning texts. In such texts, the visual choices made are highly systematic and contribute to creating the thematic significance of the story for the young reader. For example, one of the most popular and acclaimed books aimed at the preschooler, Maurice Sendak's (1963) *Where the Wild Things Are*, immediately foregrounds the issue of how a story image is framed by the white border or margin of the page (FRAMING) and how the placement of the verbal text relates to the visual (INTERMODAL INTEGRATION). This is because the first five page openings have verbal text on the left-hand page and on the right-hand page an image surrounded by a white margin, but with each succeeding image larger than the previous one, until the 6th image fills the entire page, expunging the margin, and the 7th begins to transgress across the gutter to encroach on the left-hand (text) page. By the 9th spread the image has extended right across both pages so that the text has to appear beneath, rather than facing the image, following which the text is entirely ousted and there are three spreads entirely filled by images that extend to the page edge.

These choices are far from arbitrary when considered in relation to the ideational and interpersonal meanings in the story. The young protagonist, Max, is shown initially in smaller pictures 'hemmed in' and constrained by the surrounding white margin. At this point, he is full of aggression, getting up to serious 'mischief' and being sent to his bedroom as a punishment. Once there he begins to use his imagination, the room expands and transforms into a forest and Max sets off in his boat on an adventure to the land of the 'wild things'. The gradual expansion of the images to the edge of the page and beyond clearly symbolizes the liberating quality of Max's imagination. The central wordless set of spreads depict him having a 'wild rumpus' with the wild things, after which the verbiage and the margin are reinstated by degrees as Max's emotional storm subsides and he gradually returns 'home' in a calmer, happier and more reflective frame of mind. As various critics have noted

(e.g. Nodelman 1988), the framing choices are an important means of conveying Max's imaginative and emotional journey, and suggest clearly the kind of meaning made by 'margined' images (i.e. with a border) as against those that bleed to the edge of the page.

In Figure 6.1, which gives the entire FRAMING system, this choice of the presence or absence of a margin of space around the story image is indicated by the 'features' [bound] versus [unbound], with the meaning residing not in the label but in the contrast between the options (the network does not imply that options are consciously taken up by the artist). All such pairs of features in the diagram are to be read as either/or options, the selection of which may lead to further options as the figure is read from left to right (the visual realization of each feature or sub-feature is indicated by the downward sloping arrows on the diagram). In some picture books a consistent choice of one of these meaning features will be made throughout – for example, every image will be similarly [bound]¹ or not, while in other cases, like *Where the Wild Things Are*, the shifting of choices across the course of the narrative is what proves most relevant in terms of the narrative theme.

While the changing margin in Sendak's story is most significant for referencing Max's emotional state, Kress and van Leeuwen (2006) have also pointed out the role of frames, borders and white space in separating elements out, or conversely (where absent) in creating greater connections. In a picture-book image, this characteristic helps explain an important aspect of an unbound image – the fact that the lack of an intervening white margin between the image and the page edge reduces as far as possible the boundary between the reader's world and that of the story, inviting the reader to connect and feel part of that world. Again, consistent choices may be found within a particular story, or in other cases, frames and margin borders may be present on many pages but removed at key points where the reader is 'invited in' (see, for example, the final image of Anthony Browne's (2004) *Into the forest*, where a previously anxious child protagonist is greeted by a close up of the smiling mother welcoming him with open arms. The absence of any boundary between the story world and the reader's world encourages the child reader at this point to participate in the welcome and share in the strong positive affect created here).

Thus, a choice of [unbound] for the image avoids fencing in the character and also avoids holding the reader at any distance. Such images may still vary, though, in whether the story-world setting fills the entire page or whether the characters are simply shown without context on a white page background (see Image 6.1).

This is the choice of [contextualized] versus [decontextualized] for an unbound image, an option which Kress and van Leeuwen (2006) treat as an interpersonal one relating to the relative 'realism' of the image. However, in picture books, the removal of a depicted context seems most significant as a means of making the behaviour or attitudes of the depicted character much more salient (Nodelman 1988), thus triggering an evaluative response in the

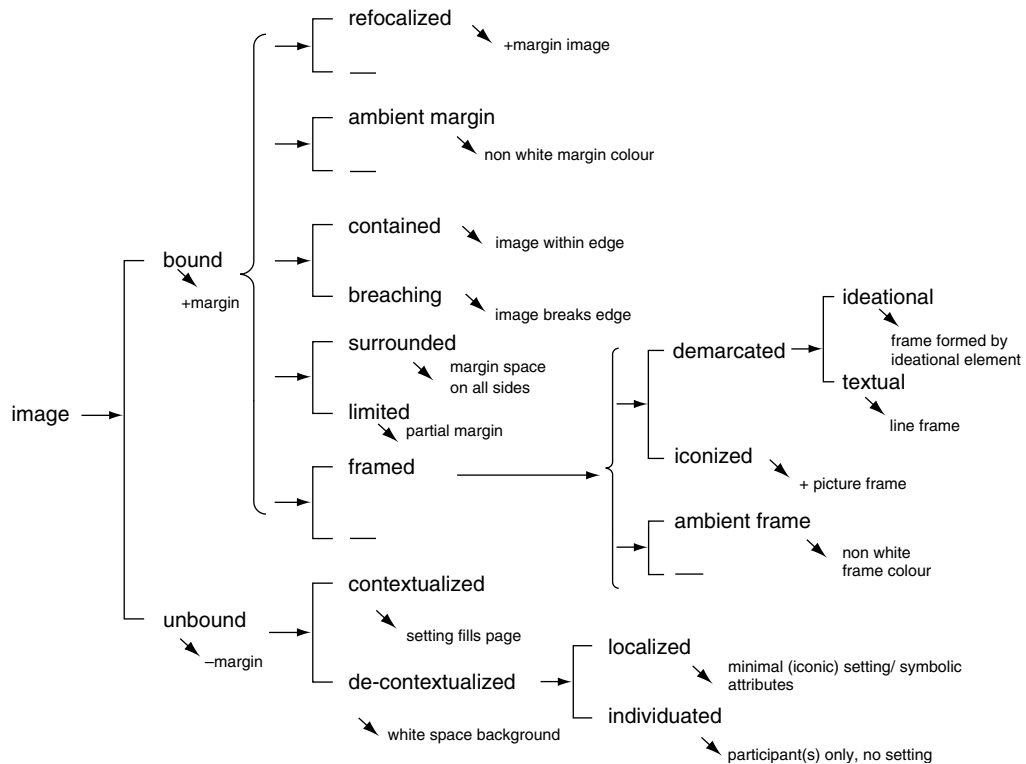


FIGURE 6.1 Choices in FRAMING



(a) [unbound: contextualized]



(b) [unbound: decontextualized]

IMAGE 6.1 Contrasting [unbound] images: [contextualized] and [decontextualized]

reader. Because of this, in some books, such images may occur at particular moments in the story when the reader is strongly invited to empathize with or judge the character (positively or negatively). Browne's (1996) *Piggybook* has variation of this kind, containing just a few [unbound: decontextualized] images in the course of the story at key moments in the generic structure (see Painter 2008 for discussion of how they contribute to the theme of the narrative). In other cases where the entire book comprises decontextualized images, the effect is to make the character/s rather than story world itself the focus of attention, thus achieving the more generic status for those participants that Kress and van Leeuwen observe for such images. An example is Machin and Vivas's (1991) *I Went Walking*, where the preschooler's attention is to be focused on the increasing number of animal participants at each stage of the simple story rather than on any fully realized alternative imaginative world.

Where a largely decontextualized image nevertheless includes a very limited local context, (the [decontextualized: localized] option), that context is likely to include what Kress and van Leeuwen (2006) refer to as 'symbolic attributes' associated with the character. For example, in Fox and Vivas' renowned tale of *Possum Magic* – a story in which a wise old possum makes her baby grandchild invisible to keep him safe from predators and then faces a dilemma when he wishes to return to visibility – we see Grandma Poss making baby Hush invisible in an image which also contains the minimal context of a shelf of magic books. The page is not filled out with any depiction of the background setting, but the shelf of books provides just enough localized setting to symbolize Grandma's special knowledge and power.

In sum, then, the most general option for framing a picture-book image involves the presence or absence of a margin to 'hold' the image within the page. Where the margin is absent and the page edge is the only limit to the image, it is [unbound] in two senses. The depicted characters are less constrained by their circumstances and the story world is more opened up to the reader. Where such an unbound image of the characters is decontextualized, attention is focused on the behaviour or nature of the depicted character/s, which when used selectively, has the potential to trigger an evaluative response at particular moments in the story or, where just a few iconic elements of the setting are provided, to assist the symbolic 'reading' of the character.

When it comes to bound images (i.e. those with a margin), there are a number of possible meaningful choices simultaneously available, as shown by the brace enclosing five different sets of oppositions in Figure 6.1. The first two of these relate to the way the margin may afford interpersonal meaning. This may occur first through the use of colour in place of the default choice of white for the margin. Colour in a picture-book image is a crucial means of creating 'ambience' or mood (Painter 2008) and can be carried by the margin as well as the image itself. A nice example is provided by the Australian picture book *Lucy's Bay* (Crew & Rogers 1992), relating the story of a young boy coming to terms with his sister's death. All pictures are [bound] and the surrounding margin is a soft, light peach colour, providing a warm ambience that plays an important role in avoiding a dark, depressing atmosphere for such a sombre theme.

A second way in which the margin may afford interpersonal meaning is by the depiction of characters in the margin itself. This is very rarely done, but has been used to great effect in the subtle and sophisticated Australian picture book *Hyrant and B* (Caswell & Ottley 2003). This story of two discarded teddy bears is narrated by one of the bears, but where another character's experiences are related, the presence of that character in the margin signals a re-focalization, such that the depicted image bound by the margin is read as that other character's memory or experience. This text in fact makes use of both [bound: ambient margin] and [bound: refocalized] options to great interpersonal effect.

As well as these possibilities for managing interpersonal meaning, there are a number of other options for bound images to be considered. First of all, the margin may surround the image on all four sides [bound: surrounded] as in the opening images of *Where the Wild Things Are*, or it may extend from only a single picture edge, thus limiting the image on the page but not enclosing it entirely [bound: limited], as also occurs in *Wild things* as the image expands (see Figure 6.2 for a schematic representation).

Then, in either case the image may be entirely contained by the margin [bound: contained] or may transgress the edge created by the margin [bound: breaching]. The choice of [breaching] is one quite frequently taken up, usually providing an iconic way of suggesting that the depicted character has too much energy, presence or momentum to be entirely constrained or bound by

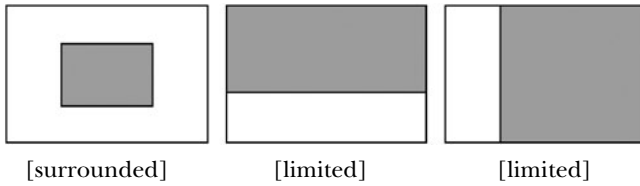


FIGURE 6.2 Options for bound images: [surrounded] and [limited]

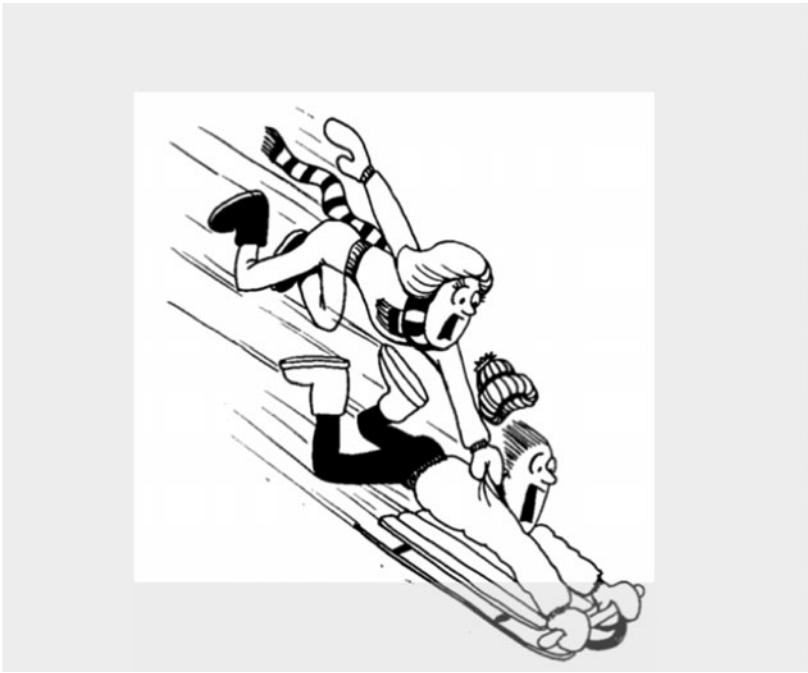


IMAGE 6.2 An example of [bound: breaching]

the margin, as shown in Image 6.2. Examples of a character breaking the frame in this way can be found in several of Browne's books, including *Piggybook* and *My Dad*, while in Sendak's *Where the Wild Things are*, by contrast, it is the gradually expanding setting that breaches the margin as Max's imaginative world expands even beyond the confines of a single page. The [breaching] option thus signifies in general terms the transgressing of a border, underlining the overall meaning of the [bound] option.

The final set of options for bound images relates to the presence or absence of lines creating a defined frame around the image.



(a)



(b)

IMAGE 6.3 [bound] images: With and without frame

If the images in Image 6.3 are compared, it can be seen that the effect of a frame is to make the image more overtly a 'picture', an option often favoured in more traditional illustrated stories (and like the margin itself, a frame can be coloured and thus contribute to the ambience created by the picture). While most books make the same choice throughout in terms of framing the images or not, Browne's (1994) *Zoo* is an example of one that exploits the possibility of variation to great effect. The book tells the humorous story of a family's day out at the zoo to make a moral point about the inhumanity of caging and objectifying animals. The book is laid out with a small unframed picture of members of the family on each left-hand page (together with the verbal text), and a large, beautifully rendered, framed picture of the animals on the right-hand page, a contrast in framing which quietly emphasizes the way the animals are 'a sight' displayed for human enjoyment. In the book, the reader/viewer is gradually moved to take on the animals' perspective, and as part of this process there comes a point where the left-hand image of the family as part of an unpleasant crowd of zoo patrons is enclosed in a frame, emphasizing how they appear as an ugly sight from the animal's point of view.

Browne's picture books in fact make considerable clever use of the [framed] option. While the frames in *Zoo* are of the most straightforward kind, explicitly rendered by a black or coloured line ([demarcated: textual]), in other books there are images where it is the ideational content of the image that creates the frame. For example, in *Voices in the Park* (Browne 1994), there is an image where the playground apparatus being enjoyed by the children serves as a frame to the image (the [demarcated: ideational] option), and in *My Dad* (Browne 2000), the edge of the blackboard on the wall behind Dad-as-teacher creates an ideational frame within the textual one. In such cases where an ideational element is used to demarcate the frame, the frame appears to serve as a symbolic attribute: signifying the playfulness of the children in the first instance (*Voices in the Park*) and the authority and knowledgeability of Dad in the second (*My Dad*). Finally, another of Browne's books, *Piggybook* (1996), illustrates an additional option that can be taken up. This is the possibility of elaborating the

frame so that the image is enclosed in what appears to be a mounted 'picture frame'. In this text it occurs when the male chauvinist pig of a father is shown at the window as a reformed character, happily washing up, in a view which looks like a fully framed picture. As well as drawing attention to the image as a picture, the effect of this [framed: iconized] option is to turn the character into an ideal – here of the model domestic male. This meaning of iconization is also evident in Jeannie Baker's (1991) ecologically themed book, *Window*, where the view of natural world that is being idealized for the reader is depicted through an old-fashioned window that serves as a frame and mount for the view beyond (see Image 6.4).

Options within the system of framing therefore play a considerable role in organizing visual meaning, inviting the reader either to enter the story world [unbound: contextualized] or to contemplate it [bound: framed], to focus on the fictional world in its entirety or more specifically on the behaviour and attitudes of its characters [unbound: decontextualized]. Framing choices can also emphasize the character's sense of being contained or constrained [bound], the emotional or physical dynamism of the character [bound: breaching] or his or her key symbolic attributes [unbound: decontextualized: localized], [bound: framed: ideational]. Readers can additionally be guided to see the character (or the setting) as an ideal [bound: framed: iconized], to recognize a particular



IMAGE 6.4 The [framed: iconized] option: Initial spread of J. Baker's (1991) *Window*

relevant point of view [bound: refocalized] and to respond to the ambience implied by the colour choices of margin and/or frame.

While the FRAMING network is concerned with the various options for creating a border **around** the image, it is of course also possible for ideational elements such as doorways to create internal frames **within** a picture, and this needs to be considered in relation to the general possibilities for organizing and arranging the depicted content within the image. To discuss this, reference will be made to Figure 6.3, detailing the system of BALANCE, which attempts to outline the most general options deployed. The work of Arnheim (1982) and Dondis (1973) on visual art and perception have been useful sources of interpretation here and also the work of Caple (2009) on the composition of news photographs. However, it should be noted at the start that picture-book illustrations are more complex and varied than is captured by the options shown in Figure 6.3, which attempts to represent only certain general and repeated archetypal patterns observed in the corpus.

When it comes to verbal language, the dynamics of sequence and of tonic placement allow information to be packaged according to the speaker/writer's chosen point of departure (Theme) and chosen focus of hearer/reader attention (New), creating 'periodic' or 'wave-like' textual structures (Halliday 1979). Moreover, Theme choice within each clause is complemented by 'hyper' Theme choice within a paragraph, such that topic sentences predict and prepare for the rest of the paragraph, while text introductions offer a 'macro' Theme to the entire piece and text conclusions accumulate points into a section of macro New. These layers of organization build what Martin (1996) refers to as a 'hierarchy of periodicity' in a verbal text, especially an expository one. It remains an open question whether a visual image, which does not unfold in time in a comparable way, can realize any meaning equivalent to this hierarchy of periodicity, but there is no question that visual information is packaged and organized in ways that draw attention towards or away from different visual areas and depicted ideational elements.

As shown in Figure 6.3, two basic and contrasting options of BALANCE are for a composition to be placed in or balanced around a centre in various ways [centrifocal] or else to have same/similar ideational elements repeated in a series across the image [iterating]. In the latter case, the elements are nearly always organized in fairly regular 'lines', whether vertical, slanted or horizontal, as when a row of child or animal characters is depicted. This is the [iterating: aligned] option as opposed to that of [iterating: scattered], found only once in the picture-book corpus (on the endpapers of Marsden and Tan's (1998) *The Rabbits*, where a representation of nature in the form of a scattering of birds, leaves, twigs etc. suggests the unregimented random nature of the wilderness). The [iterating: aligned] choice is favoured for displaying a series of characters, such as the monsters who greet or farewell Max in *Where the Wild Things Are*, or where the child narrator gradually accumulates more and more animal companions in Machin and Vivas's (1991) *I Went Walking*.

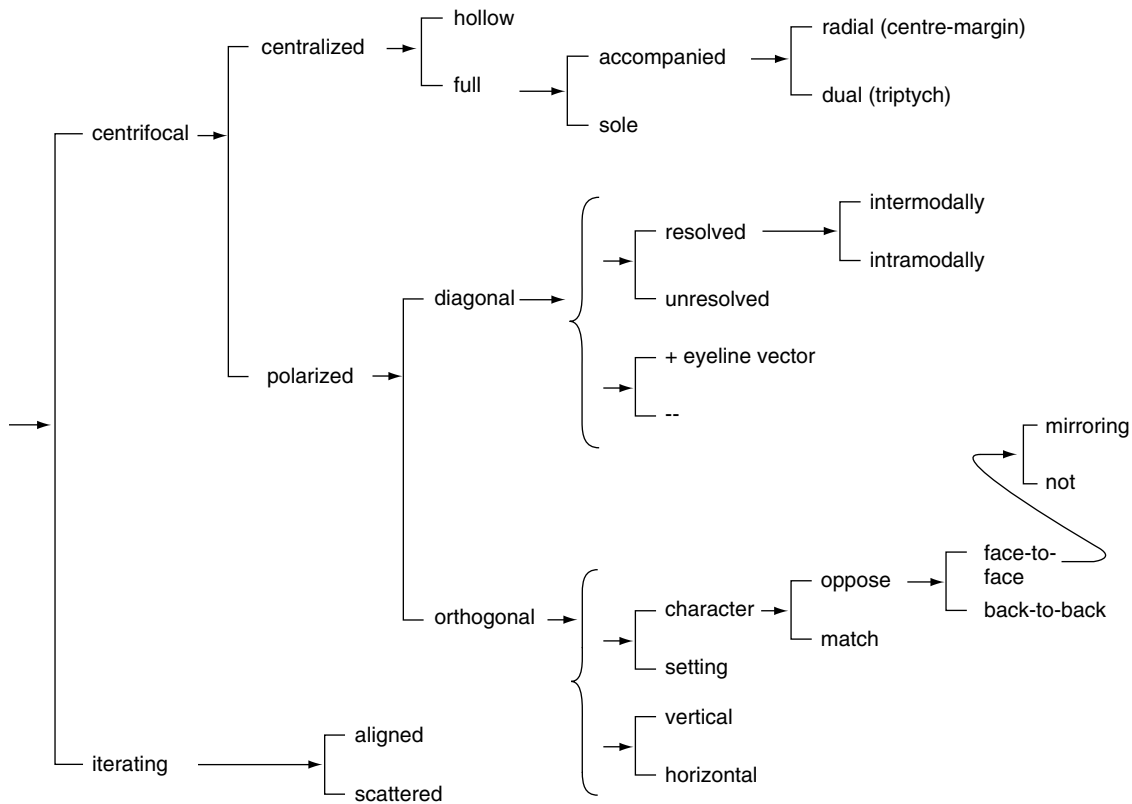


FIGURE 6.3 Choices in BALANCE

A centrifocal image can take a number of forms, with the principal contrast between a [centralized] and [polarized] composition, as shown in Figure 6.4.

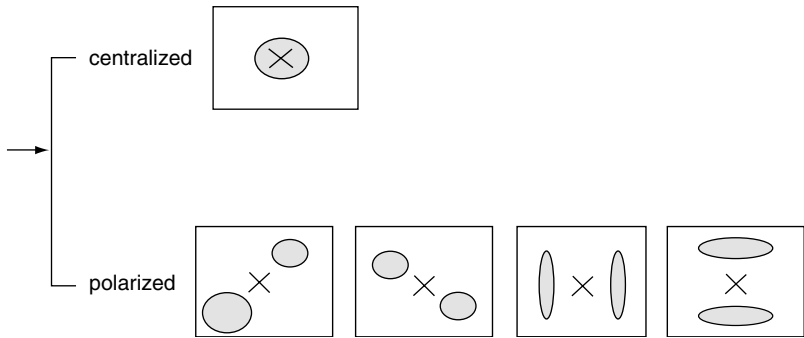
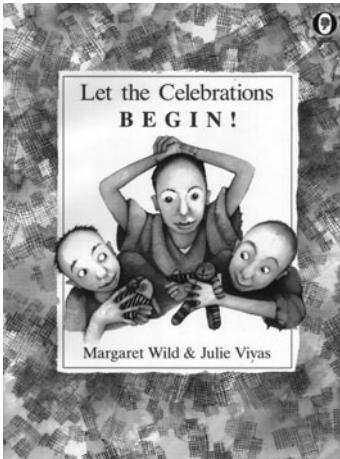


FIGURE 6.4 Centrifocal options: Balanced on or around a centre (indicated by X)

The most straightforward form for a centralized image to take is for the centre of the page to be filled, usually by a single central character or group, drawing the gaze to that participant in an unambiguous way. This is the option of [centralized: full] – a kind of bullseye composition that may be used to create a moment of stasis in the momentum of the narrative. Balance can also be created by ranging the participants around the centre of the page in a circle, [centralized: hollow], a rare choice in narratives, though common for life-cycle depictions in information books.

Where the centre is filled, the centralized participant may be accompanied either by a pair of other elements [accompanied: dual] or by an encircling ring of other participants [accompanied; radial]. These are compositional layouts noted for other kinds of material by Kress and van Leeuwen (2006), who refer to them as ‘triptychs’ and ‘centre-margin’ compositions, respectively. Two different book covers can illustrate these patterns: that of Wild and Vivas’ (1991) *Let the Celebrations Begin* (about the sense of community possible even in a concentration camp) shows the narrator accompanied by a companion on either side, while Lunn and Pignataro’s (2002) *Waiting for Mum*, about an overanxious child, shows the protagonist encircled by her worries. Thus, one cover uses the [accompanied: dual] composition to signify lack of aloneness as a positive feature, while the other uses the [accompanied: radial] option to thematize the negative situation of feeling surrounded and besieged on all sides (see Image 6.5).

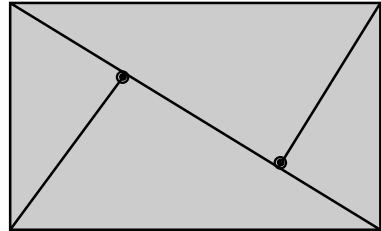
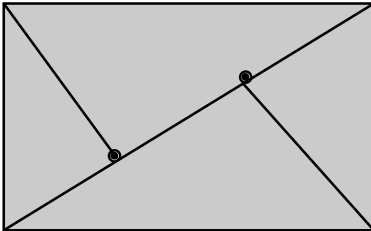
The other group of [centralized] images are those taking up the [polarized] option and balancing different depicted elements around a space. Budding photographers are advised to place elements of their composition in this way on a diagonal axis to create a sense of balance without filling the centre (Präkel 2006) (see Figure 6.5 for the points favoured for creating a balanced composition around a vacant centre (Dondis 1973)).



(a) [accompanied: dual]



(b) [accompanied: radial]

IMAGE 6.5 Options for [centralized] images**FIGURE 6.5** Placement of pictorial elements for a [polarized: diagonal] composition

Where this is done, the polarization is [diagonal] and [resolved]² (i.e. achieving balance), which is a very common choice for picture-book images, usually with characters as the polarized pictorial elements. Balance of this kind may be further enhanced by mutual gaze between the depicted characters, which strongly guides the reader to view the polarized composition as a cohering unity (the [+eyeline vector] option in Figure 6.3). Sometimes however, the balance is only achieved intermodally by opposing a pictorial element against a verbal text element which participates in the composition ([resolved: intermodal]).

On occasions, of course, it is preferable not to resolve the polarization, but rather to create an **unbalanced** effect in order to encourage page-turning or to foreground narrative complication. While this option is not taken up as often as might be predicted for narratives, the preschooler text *I went walking*

(Machin & Vivas 1991) provides a repeated series of two somewhat unresolved images followed by a fully balanced third, matching the wording of 'I went walking/ What did you see?/ I saw a [animal] looking at me'. These choices encourage the novice reader to turn the page to arrive and pause at the balanced image, helping to create a pattern that can be broken at the climax, and to pace the text into a series of comparable incidents, introducing the pre-reader into some very fundamental aspects of literary form.

Polarization in picture-book images occurs not only on a diagonal axis but also on a vertical or horizontal one [polarized: orthogonal], where a balance may be created in relation to either the setting or the characters. For example, polarization of setting may occur where a clump of trees on the left is balanced against a building on the right or where the image is split into a dark and a light half, as in Browne's (1998) *Voices in the Park*, where the cheerful child sits on a bench in a sunny, summery setting next to the nervous and cowed child in a more gloomy setting. The choice of [polarized: opposed] here organizes the interpersonal ambience and helps the novice reader to read the symbolic significance of setting, teaching another fundamental aspect of narrative. Less commonly, two similar (rather than contrasting) elements of the setting – for example, a pair of beach umbrellas on one of the pages of *Possum Magic* – may be balanced against each other in a choice of [polarized: match].

While Kress and van Leeuwen (2006) see left-right polarization as signifying a Given-New relation, and polarization on the vertical axis as signifying an Ideal-Real relation, these interpretations were not found to be very convincing for images in the picture-book narratives. In fact the most frequent kind of orthogonal polarization is the depiction of two characters on a vertical or horizontal axis, so as to enable the image composition to organize interpersonal meanings. Narratives are primarily concerned with interpersonal relationships between characters and these are readily signalled by the placement of characters on the page and their orientation to one another. Where the stance and posture of characters 'match' one another, some form of solidarity and likeness is foregrounded, as in the example from *Let the Celebrations Begin*, where two of the camp inmates are shown sitting side by side with similar poses (see Image 6.6).

If characters are depicted in a [polarized: orthogonal: **opposed**] composition on the other hand, whether on a vertical or horizontal axis, the nature of the interpersonal relation will vary according to their bodily orientation. If the characters are face-to-face, they are in contact, possibly in dialogue, with proximity, stance and expression indicating the intimacy and affect of the contact. On the other hand, if back-to-back with one another, disconnection or conflict is signalled, as in the central spread of John Burningham's (1984/1988) *Granpa*. Here the separated back-to-back image of child and old man is accompanied by the snatch of dialogue 'That was not a very nice thing to say to Granpa', evoking with both force and economy the temporary rupture in the familial relationship, without any need for an intervening narrative voice.

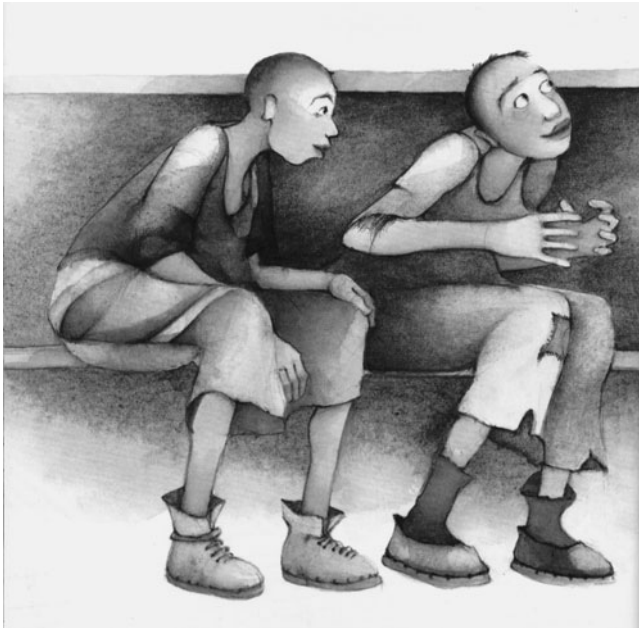


IMAGE 6.6 [polarized: orthogonal: match/character] (Wild & Vivas 1991, *Let the Celebrations Begin*)

The face-to-face option has the further possibility of an exact ‘mirroring’, where a character looks at their reflection, a choice that signals that issues of identity and self-worth are at stake in one way or another. A clever example is provided in Fox and Vivas’s *Possum Magic* where Grandma peers into a pool and sees her own reflection but grandchild Poss fails to see his, hinting at the problem for him of being invisible as he grows up. Another instance is found in Browne’s *Voices in the Park* when the browbeaten child, Charles, has his first taste of independence and adventure by playing with another child at the park. As he sits at the top of the slide, his lack of self-confidence is neatly captured by his tiny mirrored facial reflection represented as a version of Munch’s famous painting *The Scream*.

This image demonstrates that it is a considerable over-simplification to suggest that most picture-book images do in fact have only one compositional principle. The picture is essentially a [centralized: full] composition with the playground slide with children aloft filling the centre. But closer inspection reveals that Charles and his reflection are in a [face-to-face: mirroring] relation. Thus, while there is an overall [centralized] ‘gestalt’, the image incorporates an additional view or focus of attention, such that the ‘alternatives’ shown in the network in Figure 6.3 are not accurately represented as exclusive to one

another. While some pictures are indeed arranged in only one of the basic idealized layouts described by the network, very many in fact combine different principles within the one image.

Another example is to be found in *Possum Magic* (Fox & Vivas 1983) on a page where the two main characters are on the bottom right of a spread in an arrangement that is [polarized: diagonal: resolved]. Alongside them a balance is provided intermodally by several lines of text, but above them there are nine people sitting on benches with their backs to the reader, 'extras' in the scene. Taken as a whole, this upper group realizes the option [iterated: aligned], but considered more closely there are, within that, pairs of people in either matching or opposed face-to-face orientations. Examples such as these are relevant to the question of whether there is a visual equivalent of the linguistic notion of 'hierarchy of periodicity', where a verbal text sets up a higher order structure. A visual text differs from a verbal one in that it does not unfold in time but has all its levels of organization available to the viewer simultaneously. Rather than a macro-Theme unfolding hyper-Themes which in turn predict succeeding Themes, the 'layers' of a visual text are all present simultaneously; an image offers what might be thought of as an 'array of foci'. Information is visually packaged so that 'at first glance' one particular kind of organization is most dominant as a general compositional principle, but closer scrutiny is possible, allowing additional patterns to be attended to.

In organizing a complex composition of this kind, different artists may prefer different means of training and guiding the viewer's attention. Vivas in *Possum Magic* tends to create salience through the use of size and subtle colour choices in order to offer a number of potential foci within an overall view. By contrast, Anthony Browne is an artist who makes heavier use of internal frames to provide an array of foci. For example, in one image from *Gorilla* (1983/1992), the protagonist Hannah is in the centre of the page in a [centralized: accompanied: dual] composition, where she stands between two large male father figures (one a gorilla in coat and hat and the other the father's outdoor clothes hanging on a hook). Within this overall composition, a door jamb provides an internal frame which allows us to notice the gorilla and Hannah as a distinct pair in a balanced composition of [polarized: diagonal: character: face-to-face]. The six panes of the window set in the door offer further frames for additional foci of attention though these will not necessarily be attended to at first. Indeed Browne famously hides visual elements on the page by making them non-salient on the viewer's first overall 'take' as guided by the organizational balance of the image, so that they are revealed only on closer scrutiny or subsequent readings. The possibility of doing this depends on managing the viewer's attention in the first place to take in a view which foregrounds certain depicted elements to create an initial compositional 'take'.

Where Browne's images usually offer a clear overall principle of balance 'at first glance', McKee's (1982) *I hate my teddy bear* is interesting for its distracting and somewhat confusing images in which the two child protagonists are rarely

centre stage or made particularly salient in any way. This is in keeping with the book's metafictional nature, which frustrates our expectations of a simple narrative line with main characters, offering instead a myriad of potential, but incomplete visual stories. Thus, there are on most pages several competing foci of attention, with little sense of any single overarching compositional principle to guide the reader. By disturbing our expectations in this way McKee makes clearer what is going on in the more typical case.

The two systems of meaning that have been presented here, those of FRAMING and BALANCE, are proposed as sets of semiotic choices within only one of the three metafunctions into which meaning is organized. The textual metafunction of language is sometimes described as a 'derived' function in comparison with the ideational and the interpersonal. That is to say, it is brought into being by the presence of ideational content (talking about something) and interpersonal meaning (enacting social relations), structuring these meanings into coherent and cohesive discourse. Similarly, the visual textual metafunction described here (usually referred to as the compositional metafunction) serves to organize ideational and interpersonal meanings of picture books, and in-text interpretation needs to be considered in relation to those other metafunctions (see Painter 2007 for some account of these in picture books). Indeed, in the explanation of the meaning potential of the various systems, it has been necessary to discuss such matters as how relations between characters may be organized, how readers' attention may be constrained, how readers may be positioned in relation to the story world and how dynamism or stasis may be enabled by compositional choices. The two systems of FRAMING and BALANCE are not proposed as exhausting the meaning potential of the textual metafunction, since the various ways that the verbiage may (or may not) be visually integrated into the image also needs to be taken into account, together with a recognition of the way choices in colour, shape, setting and framing may contribute to cohesion over the course of a complete narrative. However, the two systems play a key role in organizing visual meaning within the page or spread, and while our descriptions have been informed by pioneering work by Kress and Van Leeuwen (2006) on visual grammar, the exploration of children's picture books has also indicated the value of focusing on one particular register of texts for further developing our understandings of visual semiotics.

Notes

¹ The terms 'bound' and 'unbound' were first introduced by Stenglin (2004) as semiotic resources for analysing interpersonal meaning in 3D space. This chapter extends their use to textual meanings in 2D visual images.

² The term 'resolved', after Caple (2009), borrows from Gestalt theories of perception in which perceptual 'resolution' or closure is achieved when information is organized around the balance points shown in Figure 6.5.

References

- Arnheim, R. (1982). *The power of the center: A study of composition in the visual arts*. Berkley, CA: University of California Press.
- Baker, J. (1991). *The window*. London: Julia MacRae Books.
- Browne, A. (1992). *Gorilla*. London: Walker Books. First published London: Julia MacRae Books, 1983.
- Browne, A. (1994). *Zoo*. London: Red Fox. First published London: Julia MacRae Books, 1992.
- Browne, A. (1996). *Piggybook*. London: Walker Books. First published London: Julia MacRae Books, 1986.
- Browne, A. (1998). *Voices in the park*. New York: DK Publishers.
- Browne, A. (2000). *My dad*. London: Doubleday.
- Browne, A. (2004). *Into the forest*. London: Walker Books.
- Burningham, J. (1988). *Granpa*. London: Puffin Books. First published London: Jonathan Cape, 1984.
- Caple, H. (2009). Playing with words and pictures: Text-image relations and semiotic interplay in a new genre of western news reportage. PhD Thesis, University of Sydney, Sydney.
- Caswell, B. & Otle M. (Illus.) (2003). *Hyrax and B*. Sydney: Hodder Children's Books.
- Crew, G. & Rogers, G. (Illus.) (1992). *Lucy's bay*. Nundah, Qld: Jam Roll Press.
- Dondis, D.A. (1973). *A primer of visual literacy*. Cambridge, MA: MIT Press.
- Fox, M. & Vivas, J. (Illus.) (1983). *Possum magic*. Adelaide: Omnibus Books.
- Halliday, M.A.K. (1978). *Language as social semiotic*. London: Edward Arnold.
- Halliday, M.A.K. (1979). Modes of meaning and modes of expression: Types of grammatical structure and their determination by different semantic functions. In D.J. Allerton, E. Carney & D. Holdcroft (Eds), *Function and context in linguistic analysis* (pp. 57–79). London: Cambridge University Press.
- Halliday, M.A.K. & Matthiessen, C.M.I.M. (2004). *Introduction to functional grammar* (3rd edn). London: Edward Arnold.
- Kress, G. & Van Leeuwen, T. (2006). *Reading images* (2nd edn). London & New York: Routledge. First published in 2001.
- Lunn, H. & Pignataro, A. (Illus.) (2002). *Waiting for mum*. Sydney: Scholastic Australia.
- Machin, S. & Vivas, J. (Illus.) (1991). *I went walking*. Norwood: Omnibus Books.
- Marsden, J. & Tan, S. (Illus.) (1998). *The rabbits*. Port Melbourne: Lothian.
- Martin, J.R. (1996). Waves of abstraction: Organising exposition. In T. Miller (Ed.), *Functional approaches to written text: Classroom applications* (pp. 87–104). Paris: TESOL France & US Information Service.
- McKee, D. (1982). *I hate my teddy bear*. London: Andersen.
- Nodelman, P. (1988). *Words about pictures: The narrative art of children's picture books*. Athens, GA: University of Georgia Press.
- Painter, C. (2007). Children's picture book narratives: Reading sequences of images. In A. McCabe, M. O'Donnell & R. Whittaker (Eds), *Advances in language and education* v.2 (pp. 40–59). London: Continuum.
- Painter, C. (2008). The role of colour in children's picture books: Choices in AMBIENCE. In L. Unsworth (Ed.), *New literacies and the English curriculum: Multimodal perspectives* (89–111). London: Continuum.

- Präkel, D. (2006). *Composition*. London: AVA.
- Sendak, M. (1963). *Where the wild things are*. New York: Harper & Row.
- Stenglin, M. (2004). Packaging curiosities: Towards a grammar of three-dimensional space. PhD Thesis, University of Sydney, Sydney.
- Wild, M. & Vivas, J. (1991). *Let the celebrations begin*. Norwood: Omnibus.

Chapter 7

Integrating Visual and Verbal Meaning in Multimodal Text Comprehension: Towards a Model of Intermodal Relations

Eveline Chan

University of New England

Introduction

In contexts for literacy teaching and learning, the contribution of non-language modalities in the construction of meaning is widely recognized and officially acknowledged by education departments and curriculum authorities such as those in the United Kingdom, South Africa, Canada, Singapore and Australia. For example, the Australian national curriculum for English includes in its definition of texts: written, spoken and multimodal texts – which combine language with visual images and sound, in print or digital/online forms (ACARA, 2009). However, in the large-scale assessment of student literacy, tests of reading comprehension remain strongly oriented towards the written mode, despite the inclusion of various types of images in the test materials. This is the case with the National Assessment Program Literacy and Numeracy (NAPLAN) materials in Australia (MCEETYA 2007). Nevertheless, some group test materials over time did increasingly target the reading of images in the test items. The NSW state government Basic Skills Tests (BST) and English Language and Literacy Assessment (ELLA), now replaced by the National Testing Program, are examples of such tests. The Program for International Student Assessment (PISA) in recent years has also incorporated ‘non-continuous texts’ such as charts, graphs, maps and diagrams into its assessment of reading (OECD 2006). Still, a discrepancy exists between contemporary reading practices in schools and how these are formally assessed; one of the reasons suggested for this curriculum/testing disjunction is the lack of a substantive account of the ways in which different kinds of images interact with language in different kinds of texts in models of reading comprehension (Unsworth 2006, 2008, Unsworth et al. 2004).

The purpose of this chapter is to explore a tentative framework for modelling image-text relations (earlier versions appear in Unsworth 2006, 2008) which describes the extent to which visual and verbal elements contribute to the overall ideational meaning in multimodal texts and the nature of the relationships

among the elements. It is intended that such a model will contribute to a richer understanding of students' reading of multimodal texts, while offering a systematic approach to describing inter-semiotic relations in a way that is both useful and accessible to teachers and test-writers. To test the efficacy of the model, the framework has been applied to the analysis of data from a project investigating multimodal reading comprehension in group literacy tests administered by a state government education authority (Unsworth et al. 2006–2008). The questions explored in this research relate to how image and verbiage interact in the test stimulus materials and how students interpret meanings involving image-text relations.

One of the goals of the project was to develop an account of the kinds of image-text relations students are likely to encounter in curriculum materials, tested in the first instance with the data from this study. The modelling of these relations, while initially derived from theory and research on multimodal analysis from a social-semiotic perspective, is also very much data-driven and draws on 3 sets of data gathered for this project:

1. Stimulus texts from the reading comprehension section of the Basic Skills Tests (BST) for students in primary Years 3 and 5 in 2005 and 2007, and the English Language and Literacy Assessment (ELLA) for students in Year 7 in 2007 (NSW DET 2005a, 2005b, 2007a, 2007b, 2007c);
2. student results on questions involving images from the literacy (Reading) component of the BST and ELLA for the state test populations, and post-test performance on the same subset of items for individual student participants in the study; and,
3. participants' verbalizations of their understandings of the images and texts in the test stimulus materials, and their strategies for responding to test items related to these texts – these were audio recorded in post-test interviews.

The first section of this chapter presents an approach to describing the relationships between printed text and still images, drawing on related work on modelling image-text relations in social semiotics. An account of image-text relations that may be applied to the comprehension of multimodal texts is then explored, examining a framework of relations through the analysis of the data collected for the study. Examples from the test materials are used to illustrate inter-semiotic relations in ideational, or representational meaning (Kress & van Leeuwen 2006), and to explore briefly how this may interact with compositional meaning; excerpts from interviews with participants are presented to highlight how children integrate meanings from image and text, and the difficulties they may experience with this. In light of the findings from the project, the chapter then revisits some questions on the nature of image-text relations and the implications for multimodal reading comprehension and its assessment.

Describing Image-text Relations

Connections between the pictorial and verbal elements of a text may be described in terms of the relative contribution of each mode to the overall meaning and purpose of the text or the 'division of semiotic labour' (Matthiessen 2007). For example, in looking at how words and pictures combine to create meaning in hybrid texts such as comics, McCloud (1994) identifies a range of image-text relations in terms of their equal/unequal contributions to meaning:

- a. word specific, where pictures illustrate but do not significantly add to a largely complete text;
- b. picture specific, where the picture dominates and words do not add significantly to the meaning of the image;
- c. duo specific, where words and pictures send essentially the same message;
- d. additive – words amplify or elaborate on an image or vice versa;
- e. parallel – words/image follow different courses without intersecting;
- f. montage – words are treated as integral parts of the picture;
- g. interdependent – image/words together convey an idea that neither could convey alone.

As a generalized approach to describing the distribution of meaning across semiotic modes, such an account has immediate appeal in that an impression of the overall visual-verbal balance of meaning may readily be formed. Efforts to theorize the description of such relations have been substantial in the work on multimodal analysis from a social-semiotic perspective (e.g. volumes edited by O'Halloran 2004, Royce & Bowcher 2007). Martinec and Salway's (2005) system of equal and unequal status is derived from Halliday's (1994) account of inter-clause dependency in language, and maps this notion of dependency onto image-text relations. Inter-semiotic relations may also be described in terms of how these elements connect to form a single, cohesive text. For example, Royce (2007) examines how visual message elements and language complement each other through an analysis of cohesive relations across semiotic modes, drawing on categories of lexical cohesion developed by Halliday and Hasan (1985).

Another perspective on the connections between image and language can be gained by examining the logical relations that extend across semiotic modes. It has been demonstrated that the logico-semantic relations of expansion and projection derived from the grammar of language (Halliday 1994, 2004) may also be applied to the relations between the visual and verbal elements of a multimodal text (e.g. Djonov 2005, Martinec & Salway 2005, Unsworth 2006, 2008, van Leeuwen 2005). Logico-semantic relations have also provided a way of interpreting inter-semiotic rhetorical relations, for example, in linking text and image on the printed page, and intra-semiotic image sequences in film

(Matthiessen 2007). The advantage of approaching image-text relations from this perspective is that logical relations are not confined to language – they are defined broadly enough to describe connections across different semiotic modes, yet are quite specific in the nature of the relations they delineate, thus providing a framework that works as well at the level of ‘grammar’, that is, between clauses or between parts of an image, as for relations at the level of whole text/image or discourse semantics. This versatility becomes apparent when attempting to account systematically for relations that operate at multiple levels of a text across units of different rank.

Two perspectives on intermodal relations are gained by identifying cohesive links and examining logical relations in ideational meaning. The first is one of *co-variate unity*, which reveals ‘(thematic) continuity across structural-unit boundaries of cohesive chains’, which may be semantically or grammatically interconnected (Lemke 2006:50); this view is oriented towards relations of similarity across semiotic modes. The second is oriented towards *multivariate unity*, which reveals the ‘functional complementarity of structural syntagmatic units’ (Lemke 2006:50). For the purposes of this chapter, I focus on the relations between the visual and verbal elements of a multimodal text from these two views, with a more detailed exposition of relations which may be broadly described in terms of ideational *concurrence*, where ideational meaning corresponds across semiotic modes (co-variate unity); and ideational *complementarity*, where visual and verbal modes each contribute different aspects of ideational meaning to the multimodal text (multivariate unity). I examine how these two broad types of image-text interaction are realized through the logico-semantic relations of expansion: elaboration, extension and enhancement (Halliday 1994, Halliday 2004:395ff).

Analysis of Visual-verbal Relations in the Test Stimulus Materials

In describing the image-text relations in the corpus of test materials (NSW DET 2005a, 2005b, 2007a, 2007b), the purpose for which the texts were being read was an important consideration. In a testing situation, students with good test-taking strategies will very often read the questions first then scan for information relevant to formulating an answer – the wording of a question guides the reading of the text and/or image by setting up a demand for information and constraining the subsequent response. Thus, when the test task is taken into account, different aspects of the text and/or image become important. With this in mind, we identified aspects of the language and images in the stimulus materials that were at stake in answering the test questions, and the image-text relations that had to be negotiated in order to achieve the correct response. This necessitated a consideration of the relationships among elements of a

Table 7.1 Levels of analysis: Units and rank

SEGMENTS	LEVELS OF ANALYSIS (RANK)	LABEL
Unit:	complete multi-semiotic text (single/double page spread)	M-S FRAME
Sub-units:	multi-semiotic frame	FRAME A, B, C
Verbal elements:	text	I, II, III
	clause	1, 2, 3
Visual elements:	image	I, II, III
	image part (figure/member)	a, b, c
	embedded image	i, ii, iii

multimodal text at different levels: the relations between multi-semiotic frames; inter-semiotic relations among images and text, and/or their sub-units; and intrasemiotic relations between smaller units of text (clause level analysis) and image (elements within images).

The stimulus texts were segmented into sub-units indicative of the levels of meaning expressed through visual and verbal modes (Table 7.1). To capture the relationships among these elements, the concept of *rank* was applied to the segmentation of texts. For example, at a primary level of analysis, a multi-modal text may be composed of multiple *frames*, each comprised of either verbal text or images or a combination of both. These *multi-semiotic frames* may be connected to one another cohesively through abstract relationships at some higher level of organization. Within each frame, the multimodal elements may play quite a different role in relation to one another in the meanings they signify. Text segments were also analysed at clause level in order to specify the relations between language and image in terms of their participant-process configurations. Continuous stretches of cohesive text representing a complete discourse event [MAIN-TEXT] were distinguished from text accompanying (e.g. captions) or embedded within image frames (e.g. labels) – these were coded as [SUPPLEMENTARY-TEXT].

For example, the stimulus text ‘Telling time using water’ (NSW DET 2005b:14) consists of two main frames composed of language and images entitled with the captions, *An Egyptian water clock* [FRAME A] and *A Greek water clock* [FRAME B], respectively. Both images can be described as conceptual [ANALYTICAL] images (following Kress & van Leeuwen 2006) displaying the whole-part structure of the water clocks. The ideational content of the two frames stand in a co-hyponymic relationship with each other, each representing sub-types of water clocks. The main title, *Telling The Time Using Water*, classifies these types. The frames are oriented horizontally, one above the other (Figure 7.1).

In Frame A, the supplementary text labels (III) and the parts of the image (II) to which they refer via the arrows are connected by correspondence

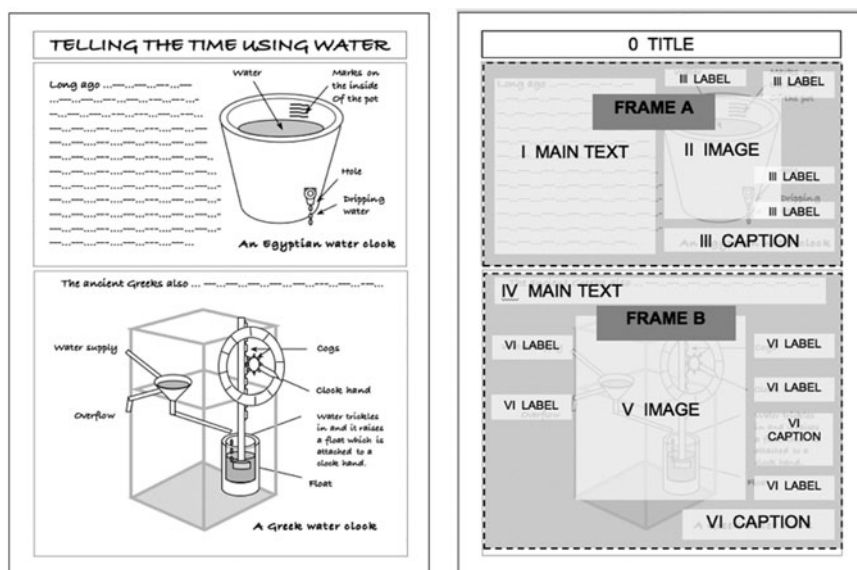


FIGURE 7.1 A schematic representation of elements in *Telling The Time Using Water*

or equivalence of meaning across semiotic modes in that they are mutually identifying. This is a very different type of relationship from that between the image (II) and the juxtaposed main text (I), which contextualizes the image historically and complements the meanings represented in the image by providing an explanation of how the water clock was used by the Egyptians to tell the time. Each set of relations contributes a different level of meaning to build the overall meaning of the text (Table 7.2).

Similarly, the arrangement of Frames A and B on the page simultaneously sets up implicit relationships of historical sequence and technological development in that the Egyptian clock (top) precedes the Greek water clock (bottom), and this order represents visually a progression from a simple to a more sophisticated device. This complex layering or multiplication of meaning through different modes of expression invites further examination, which is the focus of the section to follow. The image-text relations encountered in the data will be described in terms of their function in the expansion of ideational meaning across semiotic modes.

Ideational Concurrence: Image-text Relations of Elaboration

Ideational *concurrence* (Gill 2002, Unsworth 2006) may be described as a correspondence of ideational meaning across semiotic modes. Where meanings across modes are similar, meaning is not simply repeated or duplicated

Table 7.2 Levels of analysis in stimulus text in Figure 7.1

Level	Sub-unit	Type	Feature
A	0	SUPPLEMENTARY	[TITLE]
	TEXT		
	M-S FRAME		
	I	MAIN	[EXPLANATION]
	TEXT		
	II		[CONCEPTUAL: ANALYTICAL]
	IMAGE		
	III	SUPPLEMENTARY	
	TEXT		
	An Egyptian water clock		[CAPTION: HEADING]
B	1. Water		[LABEL]
	2. Marks on the inside of . . .		[LABEL]
	3. Hole		[LABEL]
	4. Dripping water		[LABEL]
	M-S FRAME		
	IV	MAIN	
	TEXT		
	V	SUPPLEMENTARY	[CONCEPTUAL: ANALYTICAL]
	IMAGE		
	VI		
	TEXT		
	Greek water clock		[CAPTION: HEADING]
	1. Water supply		[LABEL]
	2. Overflow		[LABEL]
	3. Cogs		[LABEL]
	4. Water trickles in . . .		[CAPTION: EXPLANATORY]
	5. Float		[LABEL]

however – the different sets of semiotic resources employed by each mode enable distinct affordances (Lemke 1998). More often than not, we find relationships of similarity where one mode elaborates on the meanings of the other by further specifying or describing while no new ideational element is introduced by the text or image. This relationship is similar to that of ‘elaboration’ between clauses in language (Halliday 2004:396).

A number of sub-types of these elaborative relations can be identified from various analyses of image-text relations. *Equivalence* is a feature where ideational content corresponds across modes in the participant-process-circumstance configuration of an image and its accompanying text (following Gill 2002), resulting in some degree of redundancy in meaning. Equivalence between image and text is also seen in keys and legends, or between parts of diagrams

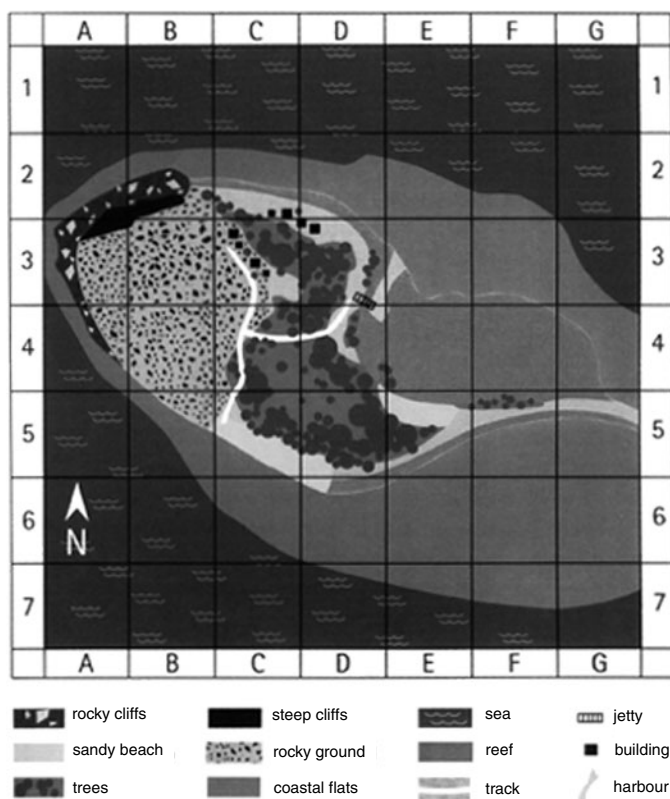


FIGURE 7.2 Equivalence in *Mapping Islands*. From *The Earth: Oceans and Sea* by Wendy Blaxland, © Macmillan Education Australia, 2000:27. Reproduced by permission of Macmillan Education

and their labels, where there is a one-to-one correspondence between an image or symbol and the word or phrase that identifies it. For example, in the Year 5 test stimulus ‘Mapping Islands’ (Figure 7.2), a key appears below a map with labeled symbols – language (word/group) and image (figure/member) at this level are mutually identifying.

An example of equivalence at clause rank can be seen where a descriptive caption provides the same information as depicted in an image, such as in the Year 3 BST text ‘Water Animal Records’ (NSW DET 2005:3). In this instance, a diagram depicting a large shark on one side of a beam balance and seven elephants on the other side is accompanied by the caption, ‘One whale shark weighs the same as 7 elephants’ – both image and language represent the same participant-process-participant configuration.

Exposition is another sub-type of elaborating relationship where image and text reinforce each other by restating or reformulating meaning in some way¹.

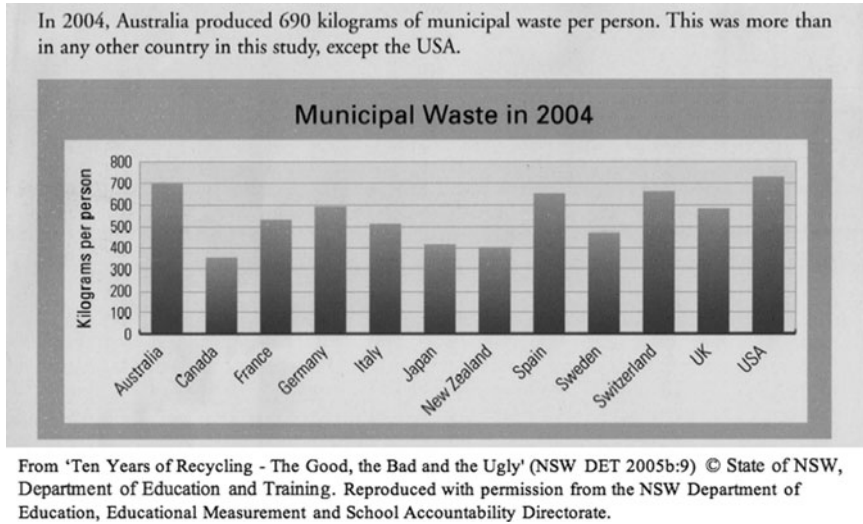


FIGURE 7.3 Exposition in *Ten Years of Recycling*

An example of exposition where the image elaborates on aspects of the text and vice versa, can be seen in the Year 5 BST stimulus (NSW DET 2007b:9), an extract from '10 Years of Recycling' (Figure 7.3). The two sentences above the image provide a direct commentary on the data displayed in the bar graph. In the image, the vertical axis represents the amount of waste produced per person in hundreds of kilograms while the main text specifies '690 kilograms' (language more specific). Similarly, the commentary states 'This was more any other country . . . except the USA' while the graph specifies the individual countries compared in the study (image more specific).

In *exemplification*, image exemplifies text or text exemplifies image; text and image represent different levels of generality (Martinec & Salway 2005), and this is realized by a class-member relationship, where specified members are not an exhaustive set but rather, represent examples from that class set. In the stimulus 'Secret Life' (Figure 7.4), the main text refers to the desert and 'its plants' – there is no other mention of plants in the text. The images below the text, however, display five examples of desert plants accompanied by captions giving the name of each plant in indigenous languages together with their common English counterparts.

Another type of concurrence between image and text can be described as *homospatality*, where different semiotic modes co-occur in one spatially bonded homogenous entity (Lim 2004). Examples of homospatality in the test stimulus materials were rare; one instance was found where the words of the poem 'Stingray' (BST5 2005:14) were arranged in the shape of a stingray.

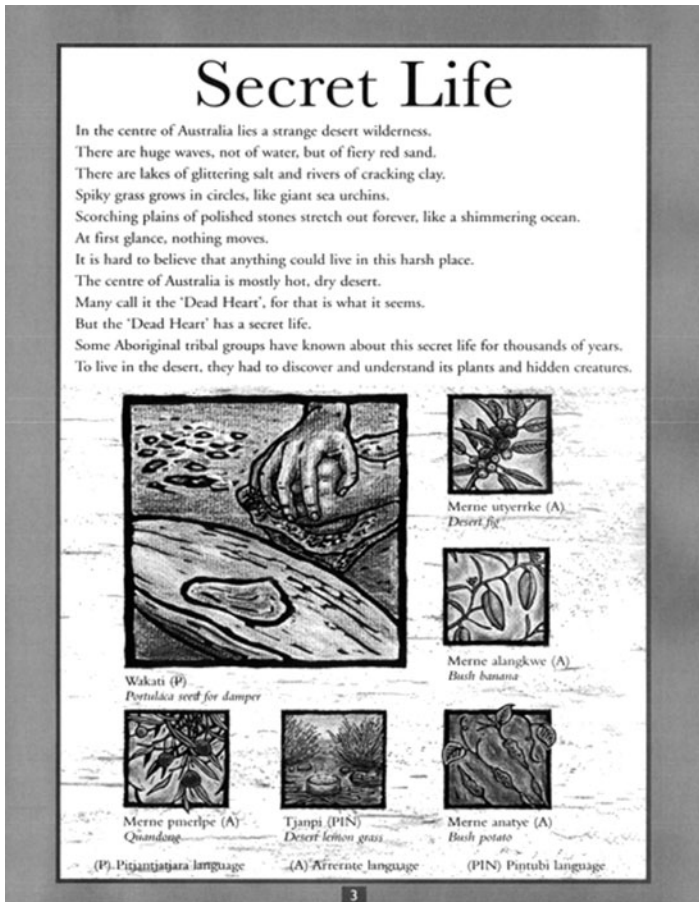


FIGURE 7.4 Exemplification in *Secret Life*. © Reproduced with permission from *Sand Swimmers* by Narelle Oliver, Lothian Children's Books, 1999, an imprint of Hachette Livre Australia

Ideational Complementarity: Image-text Relations of Extension

The concept of ideational *complementarity* has been used to describe intermodal relations where the meanings in image and text are different but complementary – meanings additional to those in one mode are represented in the other mode in a relationship of extension (Unsworth 2006, 2008). Where text and image complement each other in this way, we found three sub-types of extension in the data where image provides ideational elements (i.e. participants, processes and circumstances) additional to those in the text or vice versa: augmentation, distribution and divergence.

Augmentation involves an image extending or adding new meanings to the text or the text extending the image by providing (an) additional ideational element(s). In the current framework for analysis, the new ideational element in augmentation is realized by participants or circumstances represented in the complementary semiotic mode. For example, in the text 'Puddles' (adapted from *The Puddlemán*, Briggs 2004), the comic strip depicts two characters (Figure 7.5), a boy (FRAMES 4–7) and his grandfather [FRAMES 1–7], while the



FIGURE 7.5 Augmentation in *Puddles*. From *The Puddlemán* by Raymond Briggs published by Jonathan Cape/Red Fox. © The Random House Group, 2004. Used by permission of The Random House Group Limited <http://www.randomhouse.co.uk>

words shown in speech bubbles come from three speakers. The third character is the grandma, who is represented indirectly by her projected speech [FRAMES 3 & 8]. In this way the image and text augment each other in representing the human participants in the story.

A second type of extension, *distribution*, refers to juxtaposed images and text jointly constructing activity sequences. 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 might depict the end result of a process described in the verbal text. This occurs in the extract from 'Mr Archimedes' Bath' (Allen 1980), where the text states 'the water rose', while the accompanying image shows water overflowing from the bath (NSW DET 2007a:6).

Inter-process distribution occurs when images fill a gap in the meaning in the text; image and text complement each other in that activities or processes are distributed across the two modes. For example, in the Year 5 stimulus (NSW DET 2005b:6), 'Two Summers' (an extract from Heffernan & Blackwood 2003), the text and images are juxtaposed to jointly construct the events from one summer to the next (Figure 7.6). The activities of opening text (in italics), 'Rick *is coming to stay* again. It *takes* him seven hours on the train from the city. He's *staying* for a whole week . . . 'are represented in the words alone [CLAUSES 1–3]. This introduction is followed by the first image, depicting a scene and

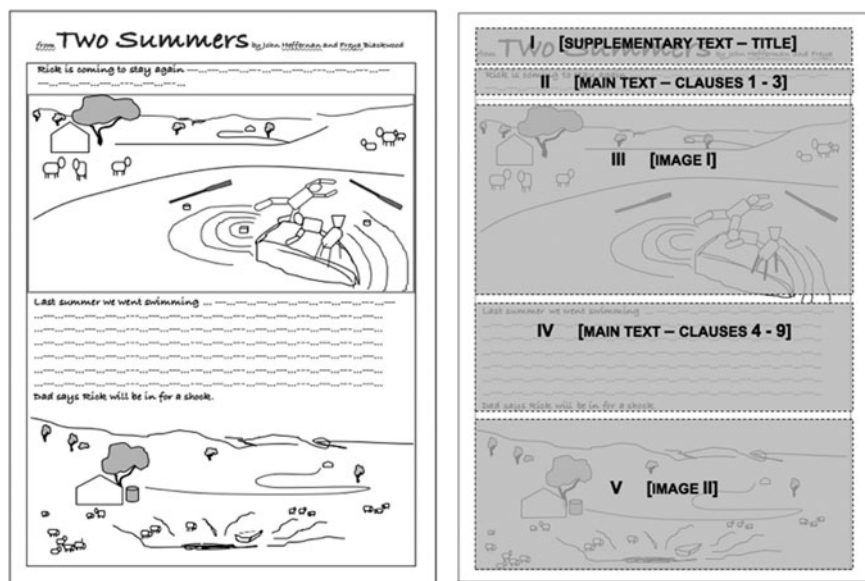


FIGURE 7.6 A schematic representation of inter-process distribution in *The Two Summers*

some of the activities from last summer's visit [IMAGE I]. The main text below this image introduces the contrast portrayed in the images between the last summer – green with plenty of water in the river and dam, and the following drought-stricken summer in the second image [IMAGE II]. While the first image is elaborated upon by the text that follows it [CLAUSES 4–6], the second image conveys through visual representation alone, the effects of the drought on the landscape. The changes of the second summer may be inferred from an integrated reading of the text [CLAUSES 7–9] and image.

Divergence was used to describe the third type of extending relation, where the ideational content of the text is opposed or at variance to that of the image, or vice versa. This term was also applied to instances where the meanings in the text and image contradicted each other. An example of divergence can be found in the extract from Anthony Browne's (1992) *Zoo* (NSW DET 2007b:2–3), where the family's dialogue about the chocolate on the first page of the extract is at variance with the pictures depicting the father and the giraffes.

A quick check for freshness is to pop a raw egg in its shell in a glass of water. If it sinks to a completely horizontal position it's very fresh; if it tilts slightly it's probably around a week old and if it floats it's not very fresh.

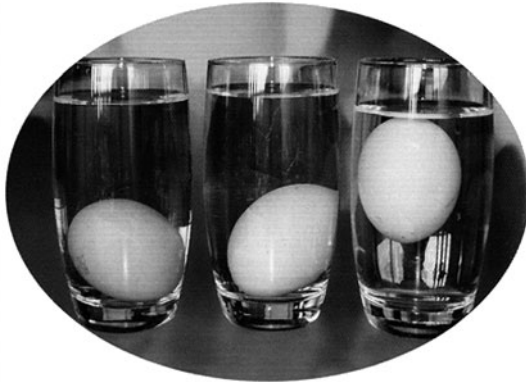


IMAGE 7.1 Enhancement in *Eggs*. 'Eggs'. Article from *Choice Magazine* (Jan/Feb 2001:23). Copyright © Australian Consumers' Association. Reproduced with permission from CHOICE Australian Consumers' Association

Ideational Complementarity: Enhancement and Projection

Language and image complement each other through enhancement when one mode provides meanings which expand another spatially, temporally or causally. While the test items did not target these relations, instances did occur in the stimulus material. For example, in the text ‘Eggs’ (NSW DET 2007b:5) both image and language provide the conditions under which the freshness of eggs may be determined but only the language provides the information about how the position of the eggs may be interpreted (Image 7.1).

Image and text were also found to complement each other through projection, the most congruent instance of this being illustrated texts with speech bubbles such as ‘Puddles’ (Figure 7.5). In this text, the human participants (as sayers and/or sensors) are represented pictorially while their projected ideas and locutions are represented linguistically.

The scheme in Figure 7.7 summarizes the framework applied to the description of image-text relations in this study. The image-text relations in the stimulus materials associated with test items were analysed using this scheme. A total of 64 visual items were identified in the 2005 and 2007 BST and the 2007 ELLA. Forty of those items (62.5%) involved relations of elaboration between image and text, and 24 items (37.5%) involved extension. There were no items which targeted enhancing relations.

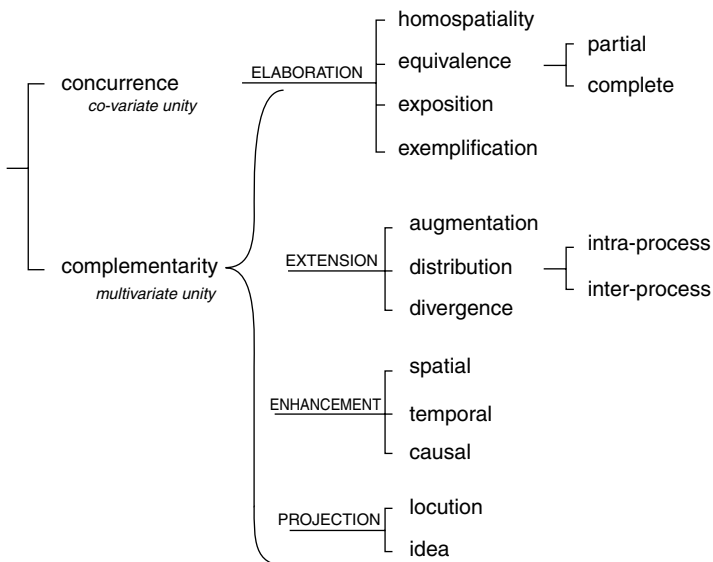


FIGURE 7.7 A summary of the framework for describing intermodal relations

Recovering Meaning from the Visual-verbal Interface

A representative sample of students who completed the 2005 Basic Skills Tests (BST) for Year 3 ($N = 70$) and Year 5 ($N = 55$) with results in low (L = lower 25% of test cohort), medium (M = middle 50% of test cohort) and high (H = upper 25% of test cohort) performance bands were interviewed about their understandings of the stimulus texts and asked to explain their strategies for answering the test questions. A structured 'think-aloud' protocol was used to elicit student verbalizations of: (a) their understandings of the texts and images; (b) whether they thought the information in the words and pictures were similar and/or different, and in what ways they were similar and/or different; and (c) the strategies they used to arrive at their answers to the test questions. The same students were interviewed again following their completion of the 2007 BST for Year 5 students ($N = 55$) and the 2007 English Language and Literacy Assessment (ELLA) for Year 7 students ($N = 41$).

In the excerpts from the interviews below, two female students explain what they think the test stimulus 'Zoo' (NSW DET 2007b:2–3) is about. Student Mf1 (medium result band) comments on the first picture and infers that the father in the story is mean and reads the second picture in the context of the overall activity constructed across the words and pictures of the text. She adopts the view of the family visiting the zoo to make sense of the text as a whole and successfully integrates the visual and verbal elements into a cohesive whole:

Mf1: Well, it's really about a family going to the zoo and what . . . looking at the animals . . . it shows two giraffes, but it doesn't really say anything about this picture. But this one it . . . I reckon it's like when they say he was in one of his moods, and they have two horns, it makes him look like the devil.

This response contrasts with that of student Lf2 (low result band), who describes the text quite literally and does not go beyond a close paraphrase of the literal meanings explicitly stated in the text or represented in the pictures, and so misses the more implicit relationships constructed across the modes:

Lf2: Okay, it's about like a mum brought a chocolate and the two kids, they want to eat the chocolate but the dad's saying no, you can't have it now. And there's tigers walking at the zoo and in the first picture like on the clouds dad had horns and the chocolate has been eaten by dad.

For the stimulus text, 'Puddles' (NSW DET 2007b:4), student responses showed further differences between participants in the high, medium, and low result

groups. In the example below, a low scoring male student (Lm1) gives a literal description of the text:

Lm1: I think it's about a man and he wants peace and quiet. He wants to sit down and drink his tea and read his newspaper. But there's a little boy who wants to play with him and take him for a walk. And . . . oh yeah, he runs away, 'cause he runs to his mum's or his house – grandma's house. And so then he sits down there and reads his newspaper and drinks his tea, but then the boy comes and he attaches a lead to him and wants to take him for a walk.

By contrast, a high performing female student (Hf3) comments more on the expressiveness of the pictures in displaying the feelings of the characters. She remarks specifically on their facial expressions and how these match the projected thoughts and speech:

Hf3: The pictures show like they're a big part of the story, and they show the actions that the boy and the grandfather do. They . . . in the first three pictures, they're showing that he's relaxed, and like the picture . . . the way he's standing, and things, and . . .

Four test items with a spread in terms of item difficulty were associated with the 'Puddles' text (correct answers underlined). Again, some differences can be noted in the responses from students with results in the different performance bands with respect to the difficulties they had in answering the questions and the strategies they used to obtain their answers. To answer question 7 correctly, (*Who says 'Oh well. He's such a little dear?'* (a) the boy; (b) children; (c) the man; (d) *Grandma*), students needed to read both image and text, which complemented each other through augmentation. For the whole test population, this item was one of the most difficult in the test (42nd out of 46; 54% students in the state answered correctly).

When asked whether she found the question difficult, a low performing student responded with the following reason:

Lf2: Kind of. Because there's no person, like they're not showing you the person, it's like there . . . they don't know who's saying it.

Participants were also asked to explain how they obtained their answers. Prompts were used to elicit more elaborated responses where necessary. For example, in response to question 8, *What is the man trying to do?*, participant Lm2 was able to obtain the right answer from the images alone ((a) go for a walk; (b) watch television; (c) make a cup of tea; (d) *read the newspaper*).

This was one of the easiest questions in the test (5/46; 94% of all students answered correctly):

I: How did you get your answer?

Lm2: Cause he escapes and he wants peace and quiet. And he sits down and he wants to read . . . and he reads the newspaper. Well, it doesn't say that he wants to read the newspaper, but . . . it looks like he wants to read it in the pictures.

In question 9 (*The speech bubble is drawn like this to show the speaker is* (a) thinking; (b) whispering; (c) *feeling pain*; (d) feeling excited), students needed to read both image and text to answer correctly. In this instance, text and image displayed equivalence in meaning. While this item required the integrative reading of text and image, most students (86%) answered correctly, for example,

Hf3: Because that . . . usually that's the speech bubble, . . . it shows that it's something expressive, and if he was thinking or whispering, it would be that sort of graphic. And he says 'ow' with it, so it's like he's feeling pain.

In contrast, question 11, which also required an integrated reading text and images, was answered correctly by only 46 per cent students in the state. This was the second most difficult question in the test. Student Lf3 failed to obtain the right answer from the images alone (*How many characters are in this text?* (c) *three*):

Lf3: *Two.*

I: Who are they?

Lf3: The kid and grandpa.

I: How did you get your answer?

Lf3: Looking at the pictures.

Student Hf3, however, utilized a combined strategy and answered correctly:

I: How did you get your answer?

Hf3: Well from the speech a lot, 'cause you can see that there's three people speaking, and yeah . . . the pictures show that there's two of them, so that can mislead you a little bit, from both.

The findings from the first stage of the study (summarized in Table 7.3) indicate that the students in the high reading performance bands effectively integrated meaning across the visual and verbal modes; used a range of test-taking strategies in addition to reading comprehension strategies to arrive at

Table 7.3 How do students recover meaning from the visual-verbal interface?

POOR READERS	GOOD READERS
Use a single strategy (E.g. text or image or guess)	Use multiple strategies (E.g. text and image in combination, as well as a range of test-taking strategies)
Rely on explicit, literal meanings	Infer meanings and implicit relations; draw on prior knowledge
Focus more on discrete, word-level meanings and visual elements	Bring together representational, interactional and compositional meaning
Rely on correspondence between image and text to aid comprehension	Successfully integrate meanings distributed across modes

their answers; and could read beyond the literal representations in the text/images. They also appeared to be more attuned to interactional meanings as well as compositional meanings, although this was not a focus of the analysis.

For the low performing students in the study, concurrence (equivalence) between verbal and visual meaning facilitated comprehension. The reinforcement of linguistic meaning through visual representation appeared to assist the poorer readers, providing additional cues for making sense of the material. Where decoding linguistic meaning was unsuccessful or only partially successful, students would rely on the images to support their interpretation. When decoding visual meaning, students mostly would scan the text to find clues for interpreting images. However, this strategy was seldom successful for struggling readers, particularly where an unfamiliar, abstract visual representation was accompanied by language that was also unfamiliar or grammatically complex. Text features which appeared to cause difficulty for students with results in the low to medium performance bands include: technicality and grammatical abstraction in language, abstraction in images, and image-text relations of extension (and enhancement).

Image-text Relations and Item Difficulty

The qualitative findings described above were further supported by a statistical analysis of the patterns that emerged in the data. A one-way analysis of variance (ANOVA) revealed significant differences at the 0.05 level in the mean item difficulties, measured in logits² (δ), associated with the different types of image-text: in decreasing order of difficulty, 'augmentation', 'distribution', 'exposition' and 'equivalence'. (For more detailed reporting of the preliminary results, see Unsworth & Chan 2008.) Items involving inter-semiotic relations where visual and verbal meanings were complementary (augmentation, distribution) were found to be more difficult than those where there was concurrence of meaning across modes. Of the two types of elaborative relationships occurring in the

data, exposition was more difficult than equivalence. These findings have clear implications for reading comprehension as is indicated by student test results.

Where there is *equivalence* between text and image, there is maximal correspondence of meaning across modes, each mutually reinforcing the meanings afforded by the multimodal text. It could be expected then, that image-text relations of this type are the easiest to comprehend. The performance of students across the state on items targeting this kind of information supports this expectation. For example, 81–91% of all students answered items 1 to 4 correctly in the 2005 BST3; items 3 (91%) and 1 (98%) in the 2005 BST5; items 16 (89%) and 13 (90%) in the 2007 BST3; and items 2 (96%) and 17 (97%) in the 2007 BST5. (The percentages in brackets indicate the proportion of the test cohort who answered the questions correctly.)

In stimulus material where the meanings in the text and image/s extend or complement each other, it could be expected that comprehension of the material would make greater demands on a students' ability to access and integrate meanings from across the modes. The questions involving images that were most difficult in the 2005 and 2007 tests, according to state-wide student performance, were items: 30 (32%) and 31 (51%) in the 2005 BST3; 28 (44%), and 37 (56%) in the 2005 BST5; 35 (29%), and 30 (47%) in the 2007 BST3; and, 32 (59%) and 30 (65%) in the 2007 BST5; all of these items targeted relationships of *augmentation* and *distribution* in comprehending the stimulus materials. This would suggest that the greater the difference in the meanings represented across the modes, the greater the level of cognitive demand on the reader in synthesizing these meanings into a coherent understanding of the material as a multi-semiotic whole.

The Nature of Image-text Relations

In the context of this study, I have restricted the account of image-text relations to the ideational meanings represented in printed test stimulus materials, focusing only on the affordances targeted by the test items. Even so, the analysis of this set of data has brought to light some of the complexities encountered in attempting to model inter-semiotic relations. One of the difficulties in modelling image-text relations is that we are looking at the interface between *typological* meanings in language, which are very often discrete realizations of meaning, and meanings which are more typically continuous or *topological* (Lemke 1998). Where there may be correspondences in ideational material at certain points (what we have termed *concurrence*), there are also continuities in visual meaning that cannot be captured in language. In that sense, any description of image-text relations at best can connect a generalization (word) with a specific instance (image) which stands in a relationship of elaboration to that word and vice versa. When discrete categories are assigned for the description or analysis of continuous meanings, typologies are inevitably imposed. Furthermore, in

language, relationships of expansion are linear and sequential – single clauses, groups or words expand upon other single clauses, groups or words as a text unfolds. Image, which is not confined by a linear semiotic, affords multiple qualification of a represented object. (For further discussion, see Unsworth & Cl  irigh 2009.) Relationships which are often made explicit when meaning is represented in a single mode (e.g. through conjunction in language) are often more implicitly realized when meaning is distributed across modes, for example, through the layout of visual and verbal elements on a page. In this way, the specialized affordances of different semiotic systems imbue multiple levels of relationship across the modes and this has implications for how a reader synthesizes those complementary meanings. As Lemke states:

the contribution of each modality contextualizes and specifies or alters the meaning we make with the contribution from each of the others. The image provides a context for interpreting the words differently, the words lead us to hear the music differently, the music integrates sequences of images, and so forth . . . (Lemke 2006:3)

Despite its theoretical limitations, the framework described in this chapter nonetheless has been shown to have descriptive value in enabling: (a) a systematic description of image-text relations in the test stimulus materials, and, (b) the identification of the types of inter-semiotic semantic relations central to understanding the text, in particular, those at stake in answering a test question correctly. The model, as applied in this research, has also been shown to have some predictive value for identifying the types of image-text interactions which could contribute to a greater level of reading difficulty for some students (Figure 7.8). Whether this finding may be generalized to other types of texts and reading environments has yet to be established.

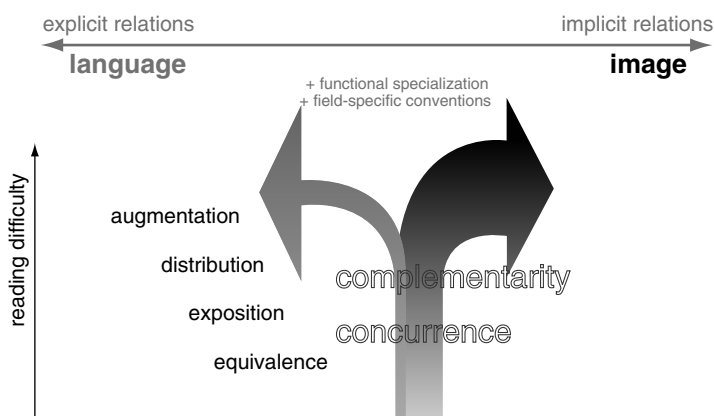


FIGURE 7.8 Image-text interaction and reading difficulty

Notions such as how texts wholly or partially relate to an image are useful as an initial foray into the way meanings are constructed across modes, but in defining test constructs that may yield useful diagnostic and pedagogical information to assist students struggling with reading complex multimodal texts, a more finely tuned model is required. For such purposes, an operational model of image-text relations first needs to be able to systematically describe the different kinds of meaning relations that are constructed intermodally, taking into account the multiple levels on which different semiotic modes connect. It also needs to specify how test items might target particular aspects of these relations that are significant for different levels of text comprehension. For example, in tests of reading, to what degree are we testing students for their visual decoding skills, or their understanding of literal meanings as compared with their skills in synthesizing meaning across different modes or critically reading the representations presented to them? Does the range of items reflect the range of educational goals we set for students growing up in a digitally enhanced, visually rich culture, where information is abundant if not superfluous but at the same time largely under-evaluated? Finally, to be of value for these purposes, any framework for description must be accessible and comprehensible to teachers and test-writers to be workable. From this initial exploration of the impact that different types of image-text relations may have on test item difficulty and by implication reading comprehension, assessment practices need to be carefully considered alongside the educational goals and contexts in which multimodal texts are engaged.

Acknowledgements

This research was supported under the Australian Research Council's Linkage Projects funding scheme (LP0561658).

The author acknowledges the copyright holders for their kind permission to include the following material in this book:

Figure 7.2 'Mapping Islands' (map and key). From *The Earth: Oceans and Sea* by Wendy Blaxland. Copyright © Macmillan Education Australia, 2000. Reproduced with permission from Macmillan Education Australia.

Figure 7.3 'Ten Years of Recycling – The Good, the Bad and the Ugly'. © Reproduced with permission from NSW Department of Education, Educational Measurement and School Accountability Directorate.

Figure 7.4 'Secret Life'. © Reproduced with permission from *Sand Swimmers* by Narelle Oliver, Lothian Children's Books, 1999, an imprint of Hachette Livre Australia.

Figure 7.5 'Puddles'. From *The Puddleman* by Raymond Briggs published by Jonathan Cape/Red Fox. Copyright © The Random House Group, 2004.

Used by permission of The Random House Group Limited www.random-house.co.uk

Image 7.1 'Eggs'. Article from *Choice Magazine* (Jan/Feb 2001, p. 23). Copyright © Australian Consumers' Association. Reproduced with permission from CHOICE Australian Consumers' Association.

Notes

¹ In Martinec and Salway (2005:50) 'image and text are of the same level of generality' in exposition in contrast to those that represent a different level of generality in exemplification. The coupling of these definitions via contrast between categories was found to be unworkable when applied to our data as many instances emerged where image provided more specificity than text or vice versa, such as in Figure 7.3, but not necessarily in an exemplifying relationship.

² Measures of item difficulty applied to the BST and ELLA by the NSW Department of Education and Training were measured in logits (δ) using Rasch item response modelling. This model locates student ability and item difficulty on the same scale, allowing the interpretation of student ability scores in terms of task demands. Item difficulty is defined probabilistically as the level of ability at which the probability of success on the item is 0.5 for a student of average ability.

References

- ACARA (Australian Curriculum, Assessment and Reporting Authority) (2009). *Shape of the Australian Curriculum: English*. Canberra: Commonwealth of Australia. Available from http://www.acara.edu.au/verve/_resources/Australian_Curriculum_English.pdf.
- Allen, P. (1980). *Mr. Archimedes' bath*. New York: Angus & Robertson/Harper Collins Publishers.
- Briggs, R. (2004). *The Puddlemans*. London: Random House Group.
- Browne, A. (1992). *Zoo*. London: Red Fox.
- Djonov, E. (2005). *Analysing the organisation of information in websites: From hypermedia design to systemic functional hypermedia discourse analysis*. PhD Thesis, University of New South Wales, Sydney.
- Gill, T. (2002). *Visual and verbal playmates: An exploration of visual and verbal modalities in children's picture books*. Unpublished BA (Hons) Thesis, University of Sydney.
- Halliday, M.A.K. (1994). *An Introduction to functional grammar* (2nd edn). London: Edward Arnold.
- Halliday, M.A.K. (2004). *An Introduction to functional grammar* (3rd edn). Revised by C.M.I.M. Matthiessen. London: Edward Arnold.
- Halliday, M.A.K. & Hasan, R. (1985). *Language, context, and text: Aspects of language in a social-semiotic perspective*. Geelong: Deakin University Press.
- Heffernan, J. & Blackwood, F. (2003). *Two summers*. Lindfield: Scholastic.

- Kress, G. & van Leeuwen, T. (2006). *Reading images: A grammar of visual design* (2nd edn). London: Routledge.
- Lemke, J. (1998). 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. (2006). Towards critical multimedia literacy: Technology, research, and politics. In M. McKenna, D. Reinking, L. Labbo & R. Kieffer (Eds), *International Handbook of Literacy & Technology*, volume 2.0. (pp. 3–14). Mahwah, NJ: Lawrence Erlbaum Associates.
- 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.
- Martinec, R. & Salway, A. (2005). A system for image-text relations in new (and old) media. *Visual Communication*, 4(3), 337–371.
- Matthiessen, C.M.I.M. (2007). The multimodal page: A systemic functional exploration. In T.D. Royce & W.L. Bowcher (Eds), *New directions in the analysis of multimodal discourse*. Mahwah, NJ: Lawrence Erlbaum Associates.
- McCloud, S. (1994). *Understanding comics: The invisible art*. New York: Harper Collins.
- MCEETYA (Ministerial Council on Education, Employment, Training and Youth Affairs) (2007). National Assessment Program Literacy and Numeracy (NAPLAN). Reading sample questions. Accessed on 18 Feb 2008 from www.naplan.edu.au/test_samples/test_samples.html.
- NSW DET (New South Wales Department of Education and Training) (2005a). Basic Skills Tests. Water, Year 3 BST 2005 (Stimulus material).
- NSW DET (New South Wales Department of Education and Training) (2005b). Basic Skills Tests. Water, Year 5 BST 2005 (Stimulus material).
- NSW DET (New South Wales Department of Education and Training) (2007a). Basic Skills Tests. Puzzles and problems, Year 3 BST 2007 (Stimulus material).
- NSW DET (New South Wales Department of Education and Training) (2007b). Basic Skills Tests. Puzzles and problems, Year 5 BST 2007 (Stimulus material).
- NSW DET (New South Wales Department of Education and Training) (2007c). English Language and Literacy Assessment. Places and possibilities, ELLA 2007 (Stimulus material).
- OECD (Organisation for Economic Co-operation and Development) (2006). Assessing scientific, reading and mathematical literacy: A framework for PISA [Electronic Version]. Retrieved 02.03.2007 from www.oecd.org.
- O'Halloran, K.L. (Ed.). (2004). *Multimodal discourse analysis: Systemic-functional perspectives*. London and New York: Continuum.
- Royce, T. (2007). Intersemiotic complementarity: A framework for multimodal discourse analysis. In T. Royce & W. Bowcher (Eds), *New directions in the analysis of multimodal discourse* (pp. 63–109). Mahwah, NJ and London: Lawrence Erlbaum Associates.
- Royce, T.D. & Bowcher, W.L. (Eds). (2007). *New directions in the analysis of multimodal discourse*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Unsworth, L. (2006). Towards a metalanguage for multiliteracies education: Describing the meaning-making resources of language-image interaction. *English Teaching: Practice and Critique*, 5(1), 55–76.

- Unsworth, L. (2008). Multiliteracies and metalanguage: Describing image/text relations as a resource for negotiating multimodal texts. In D. Leu, J. Corio, M. Knobel & C. Lankshear (Eds), *Handbook of research on new literacies*. NJ: Lawrence Erlbaum Associates.
- Unsworth, L., Barnes, G. & O'Donnell, K. (2006–2008). New dimensions of group literacy tests for schools: Multimodal reading comprehension in conventional and computer-based formats. A linkage project funded by the Australian Research Council conducted by the University of New England and the New South Wales Department of Education and Training.
- Unsworth, L. & Chan, E. (2008). Assessing integrative reading of images and text in group reading comprehension tests. *Curriculum Perspectives*, 28(3), 71–76.
- Unsworth, L. & Cléirigh, C. (2009). Multimodality and reading: The Construction of meaning through image-text interaction. In C. Jewitt (Ed.), *The Routledge handbook of multimodal analysis* (pp. 151–164). London: Routledge.
- Unsworth, L., Thomas, A. & Bush, R. (2004). The role of images and image-text relations in group 'basic skills tests' of literacy for children in the primary school years. *Australian Journal of Language and Literacy*, 27(1), 46–65.
- van Leeuwen, T. (2005). *Introducing social semiotics*. Abingdon and New York: Routledge.

Chapter 8

Rhythm and Multimodal Semiosis

Theo van Leeuwen
University of Technology, Sydney

I once heard the jazz bassist and composer Marcus Miller explain how he composed the score for the film *Siesta*, in 1987, laying a bass line first, then using a synthesizer to build up the percussion, layer by layer. At the end of that process, he realized there was something missing. The rhythm was all too mechanical. So he engaged a drummer to play a single drum in the studio, on top of the tracks he had already laid. What next, he then asked himself. I like Herbie Hancock's chords, I'll put some of those in. It was at this point that I had a revelation. I had always seen harmony as the language of Western music, and harmonic structure as its basic source of textual development, whether in Beethoven, Broadway or the Beatles. But to Marcus Miller chords were just some added spicing, some added colour. It dawned on me that in multimodal texts any semiotic mode can in principle either provide the basic structure or remain incidental, fragmented, providing, here and there, some added colour.

Language is no exception. In transcriptions of intonation and conversation (and today also in email messages), (spoken) language provides the basic structure and other elements are added as diacritics or indications-in-the-margin, providing salience, or emotive overtone, or a deictic connection, as can be seen in Figures 8.1 and 8.2.

' <i>accelerando</i> '	... 'I be lieve that "↑MÔST UNIVÉRSITIES "will be able
' <i>high</i> '	to con↑tribute . 'at any 'rate . in' re↑SÉARCH –
' <i>allegro</i> '	am. tô – the 'im MÉdiate [PRÔblem], after aî .
' <i>monotone</i> '	the ↑money that's re↑QCÍRED 'it seems to MÉ is
	not going to be ↑forth' coming ↑all at "↑ONCE – in the
' <i>low piano</i> '	MÉAntime surely 'something can be "DÔNE – ə ' on a
' <i>narrow</i> '	more ↑MÔdest SCÁLE . bȳ. those UNIVÉRSITIES
' <i>forte</i> '	whiĉh →are EQUIPPED to do "↑SÔMething . and ' here
	I ↑THINK . it is the 'interdisci'plinary – ə ↑BASIS .
	which will providE . the first be↑GINNINGS of
	reSÉARCH ²

FIGURE 8.1 Intonation transcription (Crystal 1969:179). Loudness and tempo are indicated in the margin. Pitch is indicated in the margin as well as by arrows and grave and acute accents. Stress marks have different levels







- 1 Prosecutor: So uh would you.
 2 again consider this to be:
 3 a nonaggressive, movement by Mr. King?
 4 Sgt. Duke: At this time no I wouldn't. (1.1)
 5 Prosecutor: IL is aggressive.
 6 Sgt. Duke: Yes. It's starting to be. (0.9)
 7  This foot, is laying flat, (0.8)
 8  There's starting to be a *bend*. in uh (0.6)
 9  this leg (0.4)
 10  in his butt (0.4)
 11  The buttocks area has started to rise. (0.7)
 12 which would put us,
 13 at the beginning of our *spectrum* again.
 *indicates that Sgt. Duke is pointing on the screen
 at the body part described in his talk.*

FIGURE 8.2 Transcript of excerpt from the Rodney King trial (from Goodwin 2001:176). The pointing finger indicates that Sgt Duke is pointing at the relevant body part on the screen

In transcriptions of this kind, language provides the thread of discourse, other elements are 'para'(sitical) or 'non-verbal'. But negatives exist only in the imagination. White is not non-black, it is white. Gestures are not non-language, they are gestures. And what is marginal and what is central will depend on the cultural and situational context. Perhaps we have not taken the trouble to analyse enough 'language in action' type interactions where action provides the structure and language becomes a more incidental accompaniment. Perhaps if we had done so, we would see this point more clearly.

Neither language, nor action, nor music, is indispensable in the structuring of multimodal texts that unfold over time. What is indispensable is an element all three have in common, rhythm (van Leeuwen 2005). Rhythm provides cohesion, segments the speech, or the action, or the music, into the communicative moves that propel the semiotic event forward. And rhythm is also the physical substratum, the *sine qua non* of all human action. Everything we do has to be rhythmical and in all our interactions we synchronize with others as finely as musical instruments in an orchestra. Without rhythm we fall over and trip each other up.

Analysing multimodality in films brings out how it is now the rhythm of speech, now the rhythm of action, now the rhythm of music, which provides the framework with which the signs of other semiotic modes are aligned. Figure 8.3 analyses a short excerpt from Hitchcock's *North by Northwest*. The rhythmic accents are in italics. The rhythmic phrases are enclosed in brackets and the nuclear accent of each phrase is capitalized as well as italicized. Double brackets enclose larger rhythmic units which are also, and at the same time,

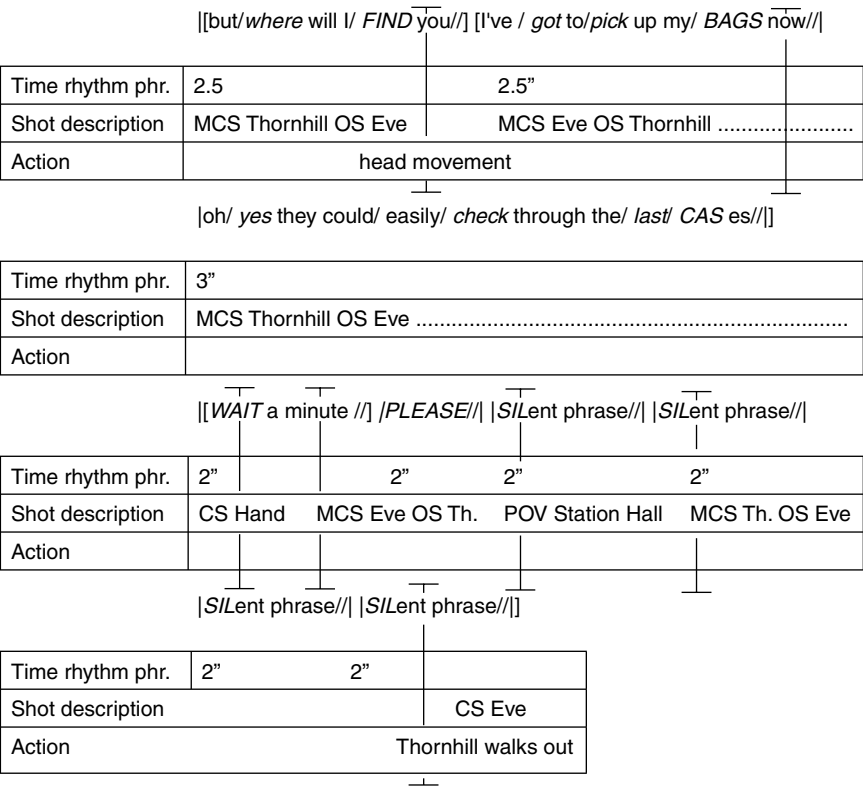


FIGURE 8.3 Rhythmic analysis of an excerpt from *North by Northwest* (Alfred Hitchcock 1959)

larger narrative moves. Note the increase in tempo and tension at the start of the second of these units, where Eve says ‘Wait a minute’. Elements other than speech – the edits of the film, the gestures of Thornhill and Eve – find their place within the temporal order of the speech rhythm. The cuts (indicated by a vertical line across all the rows) coincide with stressed syllables, the gestures with the boundaries between rhythmic phrases. Even when there is no speech, towards the end of the excerpt, the timing of the cuts still follows the rhythm initiated by the preceding speech.

Rhythm frames and delineates the communicative moves of the unfolding text, here the moves of the narrative. The excerpt immediately precedes the famous scene in which Thornhill (Cary Grant) is attacked by a cropduster plane. Eve Kendall (Eva Marie Saint) has just told Thornhill when and where to meet a mysterious man called Kaplan. What Eva knows, and what Thornhill does not know, is that the meeting is a trap and that Thornhill will be attacked.

After some perfunctory lines of dialogue, during which the audience is left to wonder whether Eve will intervene, there is a change of pace. Tension rises. At the last minute Eve seems to have second thoughts. 'Wait a minute', she says, 'Please'. A tense silence hangs between them. But the moment passes, and Thornhill leaves to board his train.

Figure 8.4 analyses a scene from Marcel Carné's *Hotel du Nord*. Here the structure is carried by the rhythm of the actor's movements. Jean (Jean-Pierre Aumont) and his girlfriend Renée (Annabella) have made a suicide pact and locked themselves in a hotel room. Jean has shot Renée but as he points the gun at himself there is a knock on the door. He escapes the hotel room via the balcony and is then seen walking along badly lit, gloomy streets, in deep despair. He stops on a railway bridge, obviously intending to commit suicide by throwing himself in front of a train. Just as he has climbed over the railing, and as an approaching train has nearly reached the bridge, a cart drawn by a white horse passes through frame, close to the camera, obscuring Jean from view. When the steam from the locomotive has cleared, we discover that Jean has not jumped. He climbs over the railing and walks back in the direction he came from to give himself up.

In this excerpt the rhythm is carried, not by speech, but by Jean's actions. The first rhythmic phrase leaves the audience in uncertainty as to what he will do next and ends when a prostitute grabs his arm, speaking the only line of dialogue in the scene. At this point the audience will wonder whether the prostitute is going to play a role in the subsequent events. But no, Jean walks on, and as he stops on the bridge, with a train approaching, the possibility of suicide can be envisaged. The next larger rhythm unit is carried by the rhythm of Jean's deliberate movements as he is getting ready to jump. As the horse-drawn cart passes the rhythm of his movements, now no longer visible, can still be felt. The clock continues to tick. At the tenth measure, well after we might have expected something new to happen, we hear the train's whistle, and exactly at the moment of the twelfth measure, we cut to a frontal view, revealing that Jean has not jumped.

In the scene from *North by Northwest* the edits and gestures were coordinated with the rhythm of the speech. Here the camera movements, the edits and the sounds, including the line of dialogue, are aligned to the rhythm of Jean's actions. And just as tempo and tension increase in the middle of the *North by Northwest* excerpt, so here, too, the tempo becomes tighter and tenser as Jean begins to climb over the railing of the bridge.

Figure 8.5, finally, shows a brief scene from an anonymous travel documentary called *Latin American Rhapsody*. The shots of mothers and babies have neither continuity of action, nor continuity of commentary or dialogue. It is the musical rhythm, which provides cohesion here – edits and gestures are aligned to the musical accents and the boundaries of musical phrases, underlining the expository structure of the short scene, which forms a mini catalogue of ethnic variety in Latin America.



FIGURE 8.4 Rhythmic analysis of a scene from *Hotel du Nord* (Marcel Carné 1938)

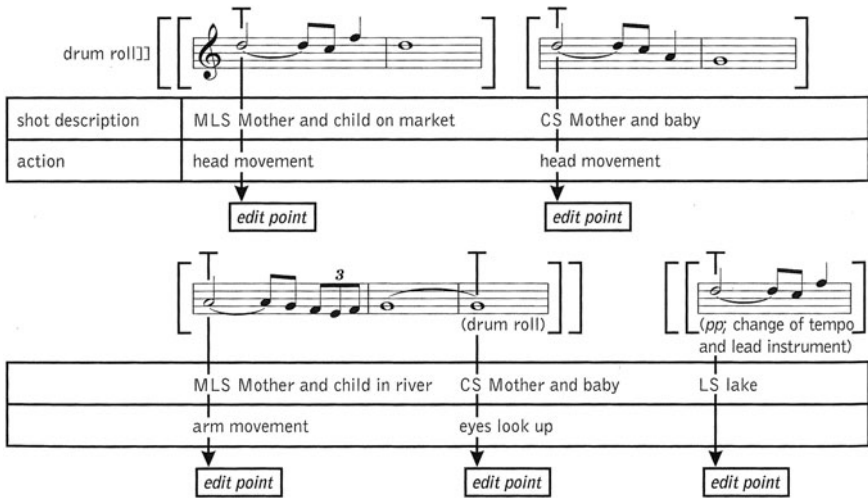


FIGURE 8.5 Rhythmic analysis of an excerpt from *Latin American Rhapsody*

In sum, either music or speech or action can provide the rhythm that carries the narrative and expository development of texts of this kind. Of course, it may be that two semiotic modes join in carrying the rhythm, as in dance, or that two rhythms are in some kind of polyrhythmic relation (cf. van Leeuwen 1999), but the general point stands: language, action and music can all be either ‘para’, ‘marginal’, or central. It would be worthwhile to study such crossmodal rhythmic relationships not just in film (although film provides convenient examples), but also in everyday interactions, a promising strand of research (see, for example, Hall 1983) which was abandoned when the tape recorder replaced the 16 mm film camera as the primary research tool in the late 1960s.

In spatially ordered texts, too, cohesion, structure and identity do not just come from language. Densely printed pages are normally read from left to right and from top to bottom, but so are many comic strips. In comic strips, language may be ‘para-visual’, consisting of little more than occasional verbal gestures (AKA Comics 2004:18):

- Just a few steps and . . . AAH!!
KRICK
- Help me!
- Hold on. I’m coming.
- Don’t let go . . . Don’t let go!
- I can’t
- Nooo!

In other cases, overall structure is carried by visual composition, whether in the form of diagrammatic structures or less conventionalized compositions.

Here, too, the language may be restricted to nouns and nominal groups which, on their own, without the visual structure would make no sense. Here is a home page, without the boxes, the colours, the columns, the colours, the fonts, the bullet points (from Lupton 2004:161):

Font Merchandise LetterSetter Free Catalog News Licensing Tech Support
Contact Custom Work Free Fonts Search Find it Jump to font kit Try fonts
before you buy with LetterSetter Strike! House-a-Rama Font Kit \$ 100 Three
Fonts 54 Dingbats 14 Illustrations Four patterns Buy it Now! House-A-Rama
\$ 100 View fonts View Font Specimens View Illustrations View Patterns &
Dingbats Try Fonts with LetterSetter House-a-Rama Buy it now!

Documents which only yesterday would have taken the form of discursive reports are now often prepared on Excel sheets originally designed for figures. For ‘personal action plans’, for instance, a template may be provided with columns for ‘action’, ‘person responsible’, ‘purpose’, ‘timeline’ and so on. Elements of this kind used to be connected through the grammar of clauses. If I am the person responsible, and writing is my action, I write ‘I write’, and not:

Action	Person responsible
Writing	Theo van Leeuwen

Now such elements are more and more often connected by the grammar of the diagram, the grid or the network. Martinec and I (Martinec & van Leeuwen 2008) have described a number of such diagrammatic ways of arranging information, showing how they underlie the structure of contemporary multimodal texts and websites. In such contexts, words and pictures become interchangeable. I could also ‘write’ this multimodal ‘clause’ as in Figure 8.6.

Here is another one of my favourite examples. A single page magazine advertisement for Sheba catfood, which has just four words ‘Spoilt, spoilt, spoilt, spoilt’ (see Image 8.1). Analysing its language only makes little sense. But together with the pictures these four words begin to make sense as an almost rebus-like sentence – something like ‘This fluffy kitten is spoilt four times over, once by each variety of Sheba cat food’. And the cohesion between the disparate elements of this multimodal ‘clause’ is predominantly visual – cohesion of colour (the yellow of the cat’s eyes is repeated in the tins of cat food and the grey of the text coheres with the grey of the cat’s fur) and cohesion of line and texture (both the outline of the kitten and the outline of the letters are soft and flowing).

The question to ask is not, or no longer: What is the relation between language and action, language and image, image and music, language and music, and so on, as if they could adequately communicate on their own, or as if some generalized statement about their central or marginal role in multimodal texts could be made. Yes, in the past, image and caption, text and

Action	Person responsible
Writing	Theo van Leeuwen

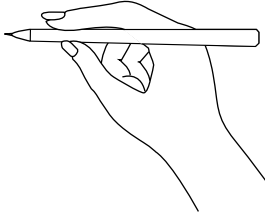

Action	Person responsible
	

FIGURE 8.6 Replacing words with pictures

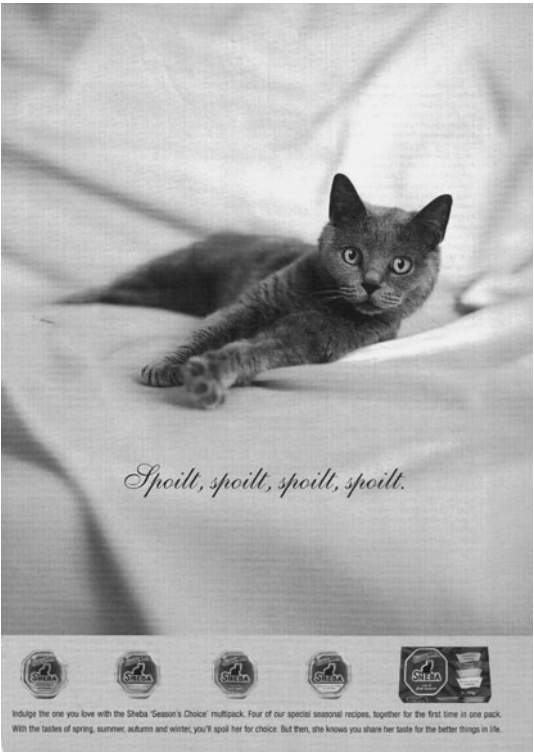


IMAGE 8.1 Sheba catfood advertisement

illustration, were relatively distinct, and the performance of spoken words did not count for as much as the words themselves. Today this is changing. Modes can become so utterly intertwined with one another that they no longer make sense on their own. Scholars exploring these issues, like the contributors to this volume, may at present still feel they are in the semiotic margins, but they will not be so for long, and their work deserves a place in the centre.

References

- AKA Comics (2004). *Sword of Majido*. Cairo: AKA Comics.
- Crystal, D. (1969). *Prosodic systems and intonation in English*. Cambridge: Cambridge University Press.
- Goodwin, C. (2001). Practices of seeing visual analysis: An ethnomethodological approach. In C. Jewitt & T. van Leeuwen (Eds), *Handbook of visual analysis* (pp. 175–182). London: Sage.
- Hall, E.T. (1983). *The dance of life – The other dimension of time*. New York: Anchor Press.
- Lupton, E. (2004). *Thinking with type – A critical guide for designers, writers, editors & students*. New York: Princeton Architectural Press.
- Martinec, R. & van Leeuwen, T. (2008). *The language of new media design*. London: Routledge.
- van Leeuwen, T. (1999). *Speech, music, sound*. London: Macmillan.
- van Leeuwen, T. (2005). *Introducing social semiotics*. London: Routledge.

Chapter 9

Meaning beyond the Margins: Learning to Interact with Books

David Rose
University of Sydney

Introduction

It's a strange form of communication, reading and writing. Like speaking, it involves at least two interactants, a reader and a writer, but sunders them utterly in time and space. It could hardly be more remote from Trevarthen's portrayal of dialogic interaction:

In a dialogue, face-to-face, two persons fill the space between with expressions of emotion. They are linked by many threads of contact between senses and movements. Each emotion is a test or judgement in that space between selves in the eyes of each other, a vibration in the threads. Eyes make a reciprocal link, each person's regard both signalling interest, or disinterest . . . But the voice carries a more intimate message of rhythms and tones, and the hands are active in gesturing the impulses of intention and memory, often referring in explicit mimetic ways to absent places and events, and to hopes and fears of protagonists in the spoken narrative . . . By the way all these parts of the body move in concert, the traffic of thoughts and feelings in one's mind are offered to, and crave response from, the sensibility of the other. (2005:104)

For most of us, this is the lived experience of everyday social interaction, but for a few of us there is an additional parallel universe of interaction mediated by written texts, disembodied from the direct relationship between speaking people, and the actual times and places in which we speak. Yet as intangible as the written world may be, it can be as real and meaningful for writers and readers as the spoken world of interacting people, things and events. I am not thinking here merely of losing oneself in the plot of an absorbing novel, but of scholars exploring new fields of knowledge, or excavating old ones, making discoveries and recharting the borders of their disciplines, all through the virtual world of the written word.

Writing makes available the realms of knowledge that have accumulated over the centuries, as our power to control the natural and social worlds has expanded. But these 'vertical discourses', as Bernstein (1996) calls them, are still only available to the small minority of citizens who have learnt how to read them, to enter their imaginary pathways and interact virtually with their writers, primarily those of us with a tertiary education. The hierarchy of educational opportunity created by such disparities in reading skills has large-scale long-term consequences for individuals and communities. In a wealthy nation such as Australia, only 10–20 per cent of citizens are afforded access to higher education, another 30 per cent access vocational training, but the majority receive no further education after school, including 10–20 per cent who will spend their lives in poverty (Saunders et al. 2007). Perhaps more disturbing is the fact that it is overwhelmingly the children of tertiary educated parents who acquire the reading and writing skills in school that are needed to matriculate to university; other children are much less likely to do so. A meagre 10 per cent improvement in university access in Australia in three decades (Rose 2004, 2007, Australian Bureau of Statistics 1994, 2004) indicates that whereas up to 90 per cent of children from tertiary educated families may matriculate, perhaps 10 per cent of children from other families do so.

It is widely recognized that children's experience of reading in the home is a major factor in their success in school. Children in literate middle class families reportedly spend up to 1000 hours in parent-child reading before starting school (Adams 1990), and large-scale research has found significant differences between tertiary educated and other parents, in the way that they read with their children (Torr 2004, Williams 1995). Beyond reading, ethnographic studies have found general semantic differences in parent-children interactions in families with different educational backgrounds (e.g. Cloran 1999, Hasan & Cloran 1990, Painter 1996). As interest in multimodal discourse analysis has grown, there have also been several studies of semantic patterns in children's picture books (e.g. Unsworth & Wheeler 2002). The focus of these studies is particularly on what is being learnt in the home, that is, variations in the grammatico-semantic resources that children are acquiring. What is less well understood is how children learn to engage with books as a mode of communication, and how this provides a foundation for learning from reading in school. Yet this is an essential further step if we are to design teaching practices that can provide these resources to all children equally. The aim of this chapter is to outline some aspects of how children learn to engage with reading in the home, and how these insights can be used to design interactional practices in school that enable all students to succeed.

Interaction in Speaking and Writing

To trace the movement from spoken to written modes of interaction, the first step is to sketch relations between the modes themselves, that is, variations in

the roles of language in communication. The analysis here assumes the stratified model of language in social context described by Martin and Rose (2007a, 2008). At the level of register, the tenor of relations between speakers may be equal or unequal, close or distant, and fields of activity may be everyday, specialized, technical or institutional. The roles of language are to simultaneously enact the tenor of relationships and construe these fields of experience. These dimensions of register are coordinated at the higher level of genre, the types of text-in-context that are recognizable in a culture, from stories to arguments to casual conversation, each of which may vary in its tenor, field and mode.

As language both enacts relations and construes fields, its mode varies in two dimensions: in terms of field, from texts that accompany activity (language-in-action) to texts that constitute their own field (language-as-reflection), and in terms of tenor, from spoken dialogue to written monologue. Values along these two dimensions are independently variable, for instance one can talk like a book, or write speech down. But taking both together, at the most dialogic end of language-in-action are direct interactions between people, the mode in which children first acquire language. Further along the continua are oral stories, in which speakers reconstruct past experience in face-to-face contact with listeners. More remote again are written texts that construct new fields, from literary fiction to academic theory. A focus of this chapter is on what happens to the relationship between interactants in this progression, from interacting directly with people, through interacting with oral stories, to interacting with books. These variations in mode are modelled in Figure 9.1, and illustrated in Tables 9.1–9.3.

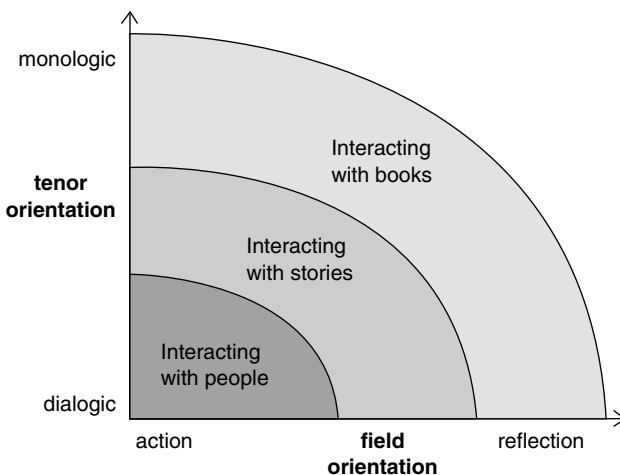


FIGURE 9.1 Mode variations

Interacting with People

Dialogic language-in-action is illustrated here with an exchange between two people engaged in an activity. It is translated from the Australian language Pitjantjatjara, in order to illustrate these pan-cultural patterns of spoken discourse (original in Rose 2001). The activity is digging for the desert delicacy, honey ants, which store nectar in their distended abdomens, in an underground chamber. The person doing the digging is learning how to identify and follow the tiny tunnel leading to the honey ants' chamber. The other person is directing the activity, teaching her how to find it. The genre is thus a pedagogic interaction.

The exchange is analysed as follows (after Martin & Rose 2007a): As the learner is doing the work, she is the 'primary actor' in the exchange, labelled here as A1. The teacher is telling the learner what to do, and so is 'secondary actor' or A2. The teacher also provides information, as the 'primary knower' or K1, while the learner asks for information, and so is 'secondary knower' or K2. The purpose of each move is also labelled, as the learner checks with the teacher, who evaluates, instructs what to do, and directs attention. Evaluations are also marked as positive + or negative –.

Table 9.1 Digging for honey ants

	Move		Purpose
Learner	K2	<i>What's this?</i> [points to tiny hole]	check
Teacher	K1	<i>No that's no good.</i>	evaluate –
	A2	<i>Throw more soil over here.</i> [points to other side of excavation]	instruct
Learner	K2	<i>This?</i> [pointing]	check
Teacher	K1	<i>Yes exactly, that hole there.</i>	evaluate +
Learner	A1	[starts to dig]	comply
Teacher	K1	<i>No, that's become no good.</i>	evaluate –
	A2	<i>Look.</i> [pointing to other side]	direct attention
	K1	<i>This is good.</i>	evaluate +
	A2	<i>Look.</i> [pointing]	direct attention
	K1	<i>It's over yonder.</i>	direct attention
	A2	<i>Dig away on the other side.</i>	instruct
Learner	K2	<i>This?</i> [pointing]	check
Teacher	K1	<i>Yes, that's it!</i>	evaluate +
	A2	<i>Try that there.</i>	instruct
Learner	A1	[starts to dig]	comply
Teacher	K1	<i>That's it!</i>	evaluate +
Learner	K2	<i>Aha!</i>	recognize

The teacher-learner relation is directly enacted move-by-move in the exchange here, as the learner asks for information and the teacher provides it, while the teacher directs activity and the learner performs it. However, a key role of the teacher, as primary knower, is to provide evaluations to the learner's questions and actions. The evaluations are positive or negative, and steadily guide the learner towards the pedagogic goal: competence in the task of recognizing the honey ants' tunnel. In addition, the field of the text is entirely dependent on the context, accompanying the activity of digging and looking for the tunnel. Each thing or place is referred to by exophoric pronouns (in bold), directing the learner's attention around the context. And these verbal directions often accompany manual pointing.

Interacting with Stories

The intermediate values along the two mode dimensions, of spoken monologue that reconstructs experience, are illustrated in Table 9.2, a traditional Pitjantjatjara story about the origin of fire. The story genre is a narrative: its Complication is that only the plains bustard *Kipara* possessed fire, and the people, who are likened to crows, could not snatch it from him, as he travelled across the country and submerged it in the ocean; the Resolution is its rescue by the black falcon *Warutjulyalpai* (literally 'snatches-fire'), who distributes it to the people. This story, among many others, would be heard often by Pitjantjatjara children, particularly around the evening fire. The Complication comprises a series of worsening problems, and the Resolution includes a solution followed by the people's reaction. These phases are labelled to the right (for discussion of story phases see Martin & Rose 2008, Rose 2006).

Table 9.2 *Kipara*

Orientation

This is a Dreaming story (*tjukurpa*), it is said. The people were living in this land. In all the land, it's said, lived the people. setting

Complication

And they, those people, had useless fire, with black firesticks (i.e. useless for igniting a fire). With black firesticks it's said they were living. problem 1

Look, they were unable it's said to obtain fire. It was like perpetual night, like living in darkness, in the dark night, and those people were living in ignorance. problem 2

And it's said one man, *Kipara* (plains bustard), was living with fire with good firesticks. So in numerous places men were thinking of this one man, of getting that fire from him. problem 3

And they were unable to get it, as they followed him and followed him continuously, snatching at the fire. All those men were unable to snatch the fire from him. problem 4

(Continued)

Table 9.2 (Cont'd)

And this journey became the <i>tjilka</i> (the annual pilgrimage for male initiation ceremonies). It was the <i>tjilka</i> host itself that was carried along in this journey.	comment
And they were unable to snatch it, as they followed him continuously, snatching at the fire. And he kept going continuously, travelling through yonder country, travelling and travelling across the land.	problem 4
At another place, at the sea he arrived, at the great ocean, and those men also he carried along with him. Into the sea it's said <i>Kipara</i> submerged, into the ocean.	problem 5
Resolution	
And <i>Warutjulyalpai</i> , the man, the bird <i>Warutjulyalpai</i> (black falcon, literally 'fire-snatches'), soared through the sky, as <i>Kipara</i> it's said submerged. Here on his head the fire was burning. And it's said <i>Warutjulyalpai</i> , flying swiftly, snatched the fire.	solution
He brought it back this way. To <i>Watar</i> he brought it, and he cast out firesticks to various places. And <i>Watar</i> is now the place of 'fire burning', the sacred well of fire. The sacred well of fire is <i>Watar</i> , Mt Lindsay. And from there he cast out firesticks to many different places.	
And those crows who lacked fire (i.e. the people) saw it and said, 'Hey, fire is burning towards us!' and they snatched up firesticks. Then they jumped up and danced, singing 'Waii!' Joyously, it's said, those crows who lacked fire, who had been crouching miserably, it's said, jumped up at that, and they saw 'There is fire over there with firesticks.' It was burning. And they danced with great joy. That's how it was.	reaction
And that is all the fire, the fire that we now have. It is ignited by rubbing sticks. And fire is a good thing. That's how it was.	comment

In oral story telling, the personal relation between speakers in a dialogue is displaced, at least in part, to a personal relation that listeners engage in with protagonists in the story. The affective core of this relation is empathy or antipathy. Empathy may be enacted by presenting protagonists with whom the listener can identify; protagonists encounter problems that create feelings of apprehension or commiseration in the listener, which are eventually relieved by resolving the protagonists' problems. Empathetic feelings may be intensified by introducing antagonists who threaten the protagonists, by building problems in a crescendo, and by the protagonists expressing their own reactions to the problems and solutions.

In the *Kipara* story, the protagonists are initially the people, whose problem of lack of fire is compounded by the darkness in which they must live, contrasting with the light and warmth of the family fire, around which children listen to the story. The people's misery, *Kipara's* villainy in denying them fire, and *Warutjulyalpai's* heroism in rescuing it, are not stated explicitly in the words

of the story, but are all expressed by the sensory contacts between storyteller and listeners that Trevarthen describes – eyes widening and narrowing, hands gesturing intention and direction, the voice intimating pity, frustration, fear, relief, joy. Furthermore, the children listening are familiar with the protagonists, crows who crouch in darkness and snatch at scraps, the bustard who walks in long strides with his beak in the air, and the black falcon who hovers high above grassfires and dives into them after prey.

Where attention was directed in Table 9.1 by exophoric references to the context, in the story listeners' attention is directed by anaphoric references to previous mentions in the text (underlined), and by the textual organization of its clauses. For example, each shift from one phase to the next is signalled to the listener by marked starting points in a clause. The first problem is signalled by iterating identities *And they, those people*, the next problem by a series of circumstances *Like perpetual night, in darkness, in the dark night*, the third problem by iterating an identity *And it's said one man, Kipara*, the fourth problem by a series of circumstances *At another place, at the sea*, and the solution by again iterating an identity *And Warutjulyalpai, that bird Warutjulyalpai*.

In sum, the **resources** that storytellers draw on to engage listeners include (at least):

- **global structures such as a Complication and Resolution,**
- **more local phases that build and release tension** and intensify feelings,
- empathy for protagonists and antipathy for antagonists,
- bodily expressions that give visceral values to empathy and antipathy,
- textual prominences that direct listeners' attention to salient events.

In each of these respects, patterns of interacting with stories are recognizable in one culture after another (see Rose 2005 for a survey).

Interacting with Books

In contrast, written texts cannot call at all on the resources of sensory contact, and far less on familiarity with their field to engage listeners. In their place, lexical and grammatical resources become more developed in written modes. This trend is illustrated in an extract from the novel *Follow the Rabbit-Proof Fence* (Pilkington 1996), in which the Aboriginal girls at the heart of the story are taken from their families by a white policeman. In Table 9.3, the stages of this anecdote are labelled – Complication and Reaction¹ – along with the story phases within each stage, including worsening problems, a description, characters' reactions, and a final comment by the author. In addition, words that enrich description (lexis) are marked in italics, and words that express attitudes (appraisal) are in bold type.

Table 9.3 *Follow the Rabbit-Proof Fence* (extract)

Orientation	
Molly and Gracie finished their breakfast and decided to take all their dirty clothes and wash them in the soak further down the river. They returned to the camp <i>looking clean and refreshed</i> and joined the rest of the family in the shade for lunch of <i>tinned corned beef damper and tea</i> .	setting
Complication	
The family had just finished eating when all the camp dogs began barking, <i>making a terrible din</i> . ‘Shut up,’ <i>yelled</i> their owners, throwing stones at them. The dogs <i>whined and skulked</i> away.	problem 1
Then <i>all eyes turned to the cause of the commotion</i> . A <i>tall, rugged</i> white man stood on the bank above them. He could easily have been mistaken for a <i>pastoralist or a grazier with his tanned complexion</i> except that he was <i>wearing khaki clothing</i> .	description
Fear and anxiety swept over them when they realized that the fateful day they had been dreading had come at last . They always knew that it would only be a matter of time before the government would track them down.	reaction
When Constable Riggs, <i>Protector of Aborigines</i> , finally spoke his voice was full of authority and purpose .	problem 2
They knew without a doubt that he was the one who took children <i>in broad daylight</i> – not like the evil spirits who came into their camps at night.	
‘I’ve come to take Molly, Gracie and Daisy, the three half-caste girls, with me to Moore Rive Native Settlement,’ he <i>informed</i> the family.	
Reaction	
The old man nodded to show that he understood what Riggs was saying. The rest of the family just hung their heads , <i>refusing to face</i> the man who was taking their daughters away from them. Silent tears welled in their eyes and trickled down their cheeks .	reaction
Molly and Gracie sat <i>silently</i> on the horse, tears streaming down their cheeks as Constable Riggs turned the <i>big bay stallion</i> and led the way back to the depot.	reaction
A high-pitched wail broke out . The cries of agonized mothers and the women, and the deep sobs of grandfathers, uncles and cousins filled the air . Molly and Gracie looked back just once before they <i>disappeared through the river gums</i> . Behind them, those remaining in the camp found sharp objects and <i>gashed</i> themselves and <i>inflicted deep wounds</i> to their heads and bodies as an <i>expression of their sorrow</i> .	reaction
The two frightened and miserable girls began to cry, silently at first, then uncontrollably ; their grief made worse by the lamentations of their loved ones and the <i>visions</i> of them sitting on the ground in their camp letting their tears mix with the red blood that flowed from the cuts on their heads.	reaction
This <i>reaction</i> to their children’s <i>abduction</i> showed that the family were now in mourning . They were grieving for their <i>abducted</i> children and their <i>relief</i> would come only when the tears ceased to fall , and that will be a long time yet.	comment

After presenting the protagonists, the author introduces tension with the dogs barking, pauses to describe the antagonist who caused it, and intensifies it with the family's feelings towards him, then worsens the problem with his brutal announcement, followed by a series of climaxing reactions, which are then explained to the reader. As in the oral Pitjantjatjara story, shifts from phase to phase are signalled by their starting points, the first description with *Then all eyes turned*, the reaction by *Fear and anxiety*, the next problem by an iterated identity *When Constable Riggs, Protector of Aborigines*, and the series of reactions by shifts from one identity to another *The old man, Molly and Gracy, A high pitched wail, The two frightened and miserable girls*.

So written stories can deploy the same resources of generic stages and phases as oral stories do for enacting empathy and antipathy, apprehension and commiseration, tension and relief. But in the absence of sensory contact with storytellers and familiarity with the field of a story, the events are expanded instead with far more diverse descriptive lexis and appraisals (including metaphors), as well as with grammatical expansions. In this extract, descriptive lexis and appraisals comprise a full third of the total words. The immediate sensory exchange between speakers in a dialogue, and between storyteller and listeners in an oral story, has been replaced by words alone. Instead of a living, feeling, speaking, gesturing person, the reader now interacts with words on the pages of a book.

Learning and Interaction

Follow the Rabbit-Proof Fence is a novel written for adult readers, but the capacity for being absorbed by its events, characters, scenes, feelings and judgments begins for most readers in early childhood, particularly with parent-child reading in the home. How do young children learn to do without the direct expressions of interpersonal relations in spoken interactions, and instead engage on their own with emotions expressed by written words? The answer, of course, is that reading most often begins not as a solitary activity, but as a medium for the sharing of emotion and attention between adult and child.

In this respect learning to read is no different from learning to speak. Careful observers consistently foreground the sharing of emotion and attention in early childhood learning. Painter (2003) shows how language begins in infancy, not with experiential categorizations, but with affective appraisals of perceptions that are shared with caregivers. Halliday (1993) describes how each new breakthrough in language learning occurs in the context of emotionally charged events. Trevarthen (2005) describes how communication between child and adult begins immediately after birth with the exchange of emotion.

Work on the neurophysiology of learning shows the central role of emotion in focusing attention:

Critical effects of emotional arousal relate to modulation of attention. First, attention appears to be focused on emotionally arousing stimuli, increasing the likelihood that emotional aspects of experiences are perceived. Second, emotionally arousing items appear privy to prioritized or facilitated processing, such that emotional items can be processed even when attention is limited . . . Because the ability to attend to and to perceive stimuli is a necessary requirement for remembering information, these effects of emotional arousal on attention influence the frequency with which emotional information is remembered. (Kensinger 2004:242)

What needs emphasizing here is the intimate relation between emotion, attention and interaction. From about 9 months, children are able to attend to both the adult and external phenomena simultaneously, a peculiarly human behaviour known as joint attention (Tomasello 2000). Joint attention then becomes the medium for cultural learning, as adults direct children's attention, or follow their attention to things and activities, then name them, evaluate, demonstrate, explain and so on. Again, shared emotion is critical as adult and child exchange evaluations of things and actions. These processes are illustrated in Image 9.1, in which an adult demonstrates an activity (drawing), directs the child's attention, and evaluates it with positive emotion, indicated by the smiles of both adult and child. This positively evaluated demonstration then prepares and motivates the child to attempt the activity, watched by the adult who will warmly praise his efforts.



(a)



(b)

IMAGE 9.1 Directing attention and emotion

Source: Rose 2006

Such cycles of learning interactions can be observed in all manner of pedagogic contexts. In classroom interactions they have been dubbed ‘IRF’, or initiate-response-feedback cycles, in which the teacher’s initiating move is typically a question. In other contexts, such as the parent-child interaction in Image 9.1, the initiating move typically prepares the learner to perform a task – here demonstrating and directing attention. Such preparations by demonstration are common in learning domestic activities, manual trades and crafts, sports and technical professions, from engineering to medicine. Preparations are designed by teachers to enable learners to successfully do a task, or steps in a complex task, so that the feedback can be affirming. The affirmation provides the ‘emotional arousal’ for attending to a further step, in which the task may be elaborated on in some way, extending the learning. For example, a child’s drawing such as in Image 9.1 may be interpreted, by the adult identifying elements in it, or by asking the child to say what she had drawn.

Reading in the Home

Parent-child reading works with this same repertoire of emotion and attention, to engage young children in the act of reading as a meaningful activity, that is, to learn to interact with a book as a partner in communication. How this engagement with books develops is illustrated in the following interaction between a mother and her 18-month old child (from McGee 1998:163), around *The Three Little Pigs* (Kellogg 1997), with relevant pages shown in Images 9.2 and 9.3. The extract includes three cycles of interaction, over four pages of the book.

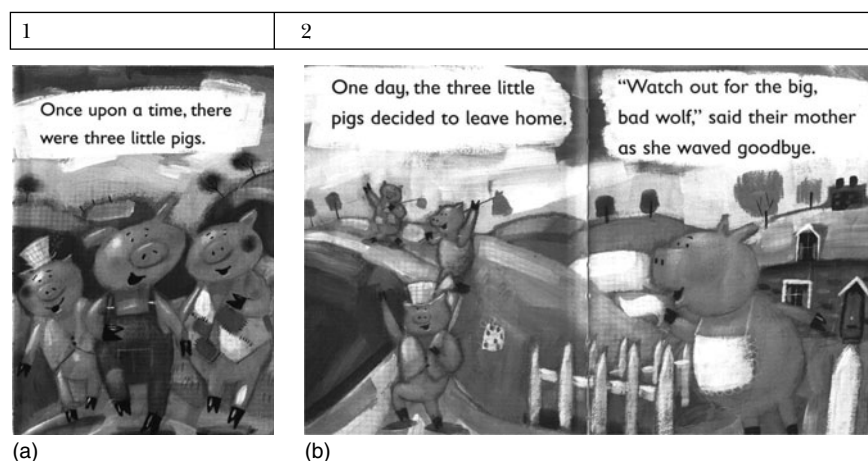


IMAGE 9.2 Images in Table 9.4 (Kellogg 1997)

As with Table 9.1, each move is labelled as K1, K2, A1 or A2. Non-verbal moves are further distinguished as ‘nv’. For example, the first move in the exchange is the child bringing and opening the book. This is labelled A2nv, as she is implicitly demanding her mother read it.

In addition, the purpose of each move is labelled in two steps. First, each interaction cycle consists of four types of phases. In one phase, the mother **prepares** the child to recognize a feature of the text; in the second the child **identifies** a text feature; in the third the mother **evaluates** her response; in the fourth she may **elaborate** with more information.

Within each of these phases, the purpose of each move is further specified. For example, in the first interaction cycle, the mother draws the child’s attention to an image in the book by pointing at it. This move is labelled as A2nv, as the mother is implicitly demanding the child pay attention to the image. She then names the image, and this move is labelled as K1, as she is giving information.

Table 9.4 First cycle

Move			Purpose	
1	Child	A2nv	[Brings the book, sits on her mother’s lap, opens the book]	
	Mother	A2nv	[points to each of the pigs on page 1]	Prepare image
		K1	<i>The three little pigs.</i>	name
	Child	A1	[points to picture of a tree]	Identify image
		K2	<i>Tee</i>	name
		K2nv	[looks up at mother]	Evaluate expect affirm
	Mother	K1	<i>Yes</i>	affirm
		K1	<i>It’s a tree.</i>	Elaborate wording
	Child	A1	[turns to page 2 and points to another tree]	Identify image
		K2	<i>Tee</i>	name
		K2nv	[looks up at mother again]	Evaluate expect affirm
	Mother	K1	<i>Um, um.</i>	affirm

In the Prepare phase, the mother draws attention to the story’s main characters by pointing (A2nv) and names them (K1). The child is too young to recognize the significance of the characters, but interprets the mother’s move as preparing her to likewise point and name. The Identify phase thus involves her pointing at the background images of trees (A1) and naming them (K2).

Significantly the child does not simply imitate her mother, but responds with her own innovation on pointing and naming. Her motivation for doing so is apparent as she looks to her mother to affirm her effort. This move is labelled K2nv as she is asking for evaluation, so that the mother's affirmation is K1. The Evaluation thus involves both these moves, apparently initiated by the child.

The positive emotion induced by success and affirmation expands the child's potential for learning something more. Elaboration phases capitalize on this positive emotion, and on the learner's attention to what has just been identified. Although the child has not recognized the significance of the characters here, the mother capitalizes on her attention, by repeating what she had said, with correct pronunciation in a full sentence. The child has thus received a micro-lesson in grammar and articulation, at the moment when she is affectively and cognitively most likely to retain it.

The child then innovates again by turning the page and identifying another tree, and asks and receives another affirmation. However, the mother does not elaborate this time, but takes advantage of her attention to initiate a second cycle, drawing her attention to the characters, and elaborating on their actions.

Table 9.4 (cont) Second cycle

2	Mother	A2nv	[Points to the little pigs on page 2]	Prepare	image
		K1	<i>Here are the little pigs.</i>		name
		K1	<i>Bye bye mama.</i>	Elaborate	image
		K1nv	[waves her hand]		activity
		K1	<i>We're going to build a house.</i>		expectancy
	Child	K2nv	[laughs]	Identify	affect
			[waves at the mama pig in the illustration]		image
		A2nv	[turns to page 3]		expectancy

In this second cycle, the learning goal progresses from identifying characters to engaging the child's empathy with their activities, and expectancy of events to come. Again the mother prepares by pointing and naming the characters, but then elaborates their activities in words and a gesture 'Bye bye mama [waves her hand]. We're going to build a house'. These are not the words in the text, rather the images are re-interpreted in terms she knows the child will recognize from her own experience.

The child can thus see herself reflected in the characters, in their activities and their relationship with their mother. This identification with the protagonists is the seed of empathy. Accordingly, the child laughs in recognition, repeating the waving gesture. Her identification also engages her interest in the characters' intentions, and so in the events to come, so that she turns the page to see what happens next.

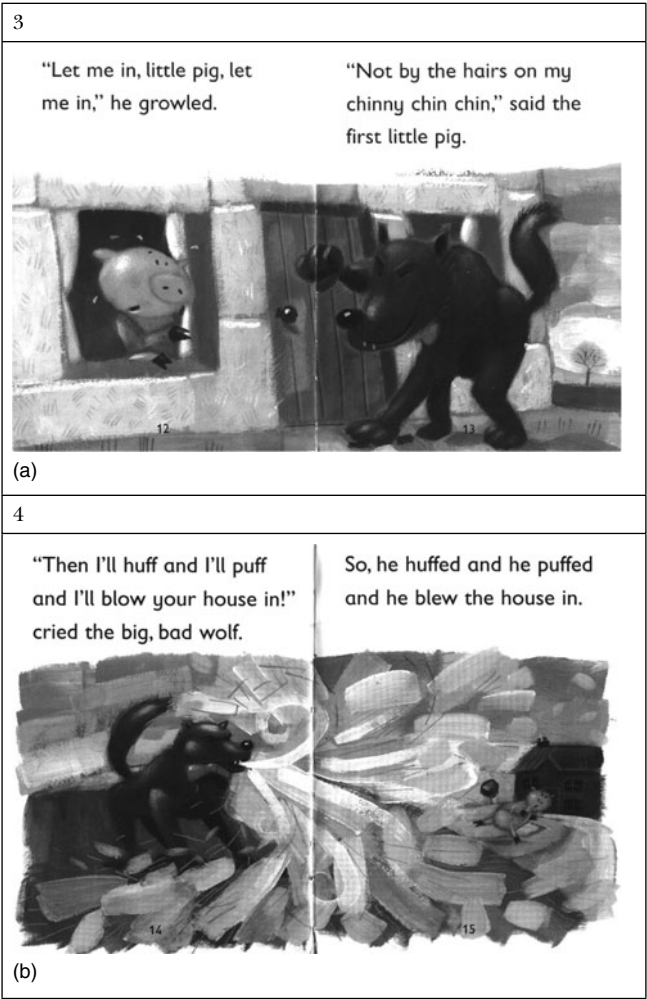


IMAGE 9.3 Images in Table 9.4 continued (Kellogg 1997)

Table 9.4 (cont) Third and fourth cycle

3	Mother	A2nv	[points to the wolf]	Prepare	image
		K1	<i>Oh oh,</i>		affect
		K1	<i>I see that wolf.</i>		self
		K1nv	[eyes get larger as if in fright]		affect
Child		A1	[turns to page 4 and points to wolf]	Identify	image
		K2	<i>Oh oh.</i>		affect
Mother		K1	<i>Oh oh.</i>	Evaluate	affirm

(Continued)

4	Mother	K1	<i>He huffed and puffed</i>	Prepare	wording
		K1nv	[blowing on child]		activity
		K1	<i>and he blewww that pig away.</i>	Elaborate	wording
		K1	<i>Very bad, isn't he?</i> [in different tone directed toward child as an aside].		judgement

In the third cycle the learning focus progresses explicitly to feelings of empathy and antipathy. This time the mother directs attention to both the image by pointing, and her own facial expression with *I see that wolf*. She evaluates the image with the apprehensive *Oh oh*, interpreting the pig's facial expression with her own, modelling the reader's empathy with the protagonist, and the antipathy to the antagonist. The child thus recognizes both the emotion and expectancy inherent in the apprehension, and responds by turning the page, and pointing to the next picture of the wolf and repeating *Oh oh*, which the mother affirms by repeating *Oh oh* herself.

In the fourth cycle the mother reads the words on the page for the first time. She prepares the child to recognize their relation to the image by blowing on her, imitating the wolf in the image. Recognizing the wolf's behaviour in both words and image then provides a context for elaborating with a moral judgement *Very bad, isn't he?*

Here are the core elements to be found in any learning interaction: the teacher directs attention, or follows the learner's attention, and models a behaviour, the learner applies the model, the teacher evaluates, and may then capitalize on the learner's success and positive feelings, by elaborating with more information. The teacher is almost always the primary knower, with the authority to evaluate the learner's responses, as well as providing information, as we saw in Table 9.1. The learner is by definition the secondary knower, the beneficiary of the information provided, whose own offerings are evaluated by the teacher/parent.

We have described these patterns as scaffolding interaction cycles (Rose 2004, 2007). In the parent-child reading genre they appear to consistently include the four phases, Prepare, Identify, Evaluate and Elaborate, diagrammed in Figure 9.2.

The Pedagogic Genre

In this brief excerpt, the child's attention has been drawn to features that identify main characters, engage readers in their activities, expect sequences of events, enact emotional reactions, and judge their behaviour. The continual affirmations serve to engage the child in the activity of story reading, rewarding

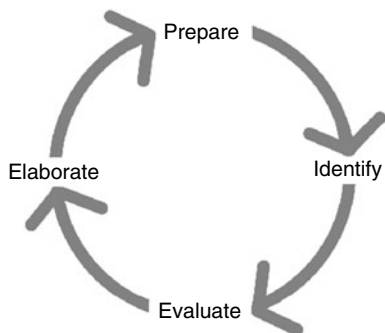


FIGURE 9.2 scaffolding interaction cycles in parent-child reading

her for responding to the mother's preparing moves. But the affirmations also function to give intense positive value to the meanings that the mother presents and the child repeats. Each exchange of value-laden meanings then enhances the child's capacity for understanding a further elaboration, which the mother usually takes advantage of.

The mother carefully and deliberately interprets the meanings in the book for the child. She adjusts, translates and reduces the meanings expressed by words and images in the book, down to the level of spoken language she knows the child will understand. This includes making implicit meanings explicit, which must be inferred by readers from the co-text, or interpreted from their own experience and values. So in order to make the text's field accessible to the child, the mother commits fewer wordings than are presented in the text, but commits more meanings that are implicit in the text. In Bernstein's theory of pedagogic classification (1990, 1996), the boundary between the child's oral experience and the written discourse of the book is weakened in each preparation move. But once the child understands each meaning in her own terms, the boundary is then strengthened in elaboration phases, to extend her understanding of the esoteric field of the book.

Over weeks, this book will be read again and again. Each time the book is read, the new meanings presented in elaborations become shared meanings; these then become the basis for preparing more new meanings until the child is thoroughly familiar with both the book's words and its semantic patterns. These patterns will then be identified and further elaborated in the next book. Over months and years the complexity of reading books increases, that is, their mode becomes more highly written. The long-term instructional sequence, through which the child's repertoire is steadily expanded, is thus shaped by the system of written language, the reservoir of meanings she encounters in children's literature. At the same time, the child will tacitly acquire a general orientation towards recognizing, interrogating and interpreting patterns of meaning in written texts. This is the semantic orientation that generates and is

fed by the play of layered meanings in literature, the literary ‘gaze’ that distinguishes members of the middle class’ inner circles. Furthermore, the child is building an orientation to interacting about these meanings with her parents, or talk-around-text. When she gets to school, the child will be ready to apply these orientations to texts and talk-around-text, and so display an aptitude for school learning that will win her constant praise from her teacher, which will in turn enhance her capacity for further learning, and so on, into the bright future of a successful student.

The elements of learning that we have identified to this point constitute what we shall call the pedagogic genre, including 4 dimensions:

1. learning activities (doing or studying)
2. modalities of learning (visual, manual, spoken, written)
3. social relations and identities (inclusive/exclusive, successful/failing)
4. instructional field (skills/knowledge).

The instructional field is projected or brought into being by the pedagogic activities, modalities and relations, as the act of saying ‘projects’ what is said (after Halliday’s 1994/2004 description of the grammar of saying). The projecting relation between the instructional field and pedagogic activities, modalities and relations is modelled in Figure 9.3.²

Re-interpreted in these terms, the instructional field in parent-child reading includes both the story of the particular book, and general patterns of meaning to be found in written stories, such as characters, expectancies, feelings and judgements. The child is acquiring an orientation towards recognizing, interrogating and interpreting such semantic patterns, a discourse about discourse, or metadiscourse. The field of the story is made explicit for the child, but

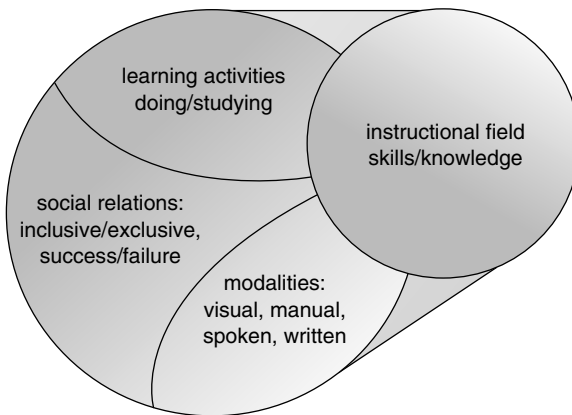


FIGURE 9.3 The pedagogic genre

the metadiscourse is necessarily implicit, as the mother cannot name the categories of discourse she is drawing the child's attention to. This two-level acquisition is made possible by the cycles of preparation and elaboration in the pedagogic activity of parent-child reading, that shunts between spoken, written, visual and manual modalities, and provides continual affirmation to enhance the potential for understanding, well beyond the child's independent competence.

Reading to Learn

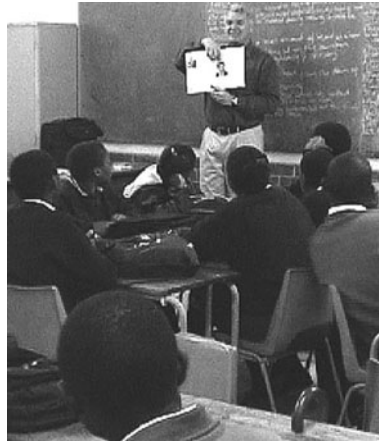
These lessons from parent-child reading are applied in the literacy pedagogy, *Reading to Learn* (Martin 2006, Martin & Rose 2005, 2007b, Rose 2004, 2007, 2008, www.readingtolearn.com.au), together with an explicit metalanguage designed from genre and register theory and discourse analysis (Martin & Rose 2007a, 2008). The sequence of the pedagogy is informed by this model of language-in-context, ordering the complex task of reading and writing in manageable steps, from patterns in the context, to the text, to its sentences and words, enabling all learners to succeed with each component in turn.

The first step prepares learners for following a text as it is read aloud, using spoken, visual and manual modalities to explore the text's field, depending on the nature of the text and the needs of students. In early years classes, for example, the teacher may talk through a picture book with children, using discussion around the pictures, similar to that above in Table 9.4. As with parent-child reading, the text may be read again and again until the children are thoroughly familiar with the field and can say and understand all the words of the text. With older students, visual images may be used to explore the field, including illustrations in books, video or other images. The sequence of the text will then be orally paraphrased or summarized by the teacher in terms familiar to the students, providing a framework for them to follow with general understanding as it is read aloud, as illustrated in Image 9.4.

Once students are familiar with the sequence of meanings in the text, they are supported to read it themselves, sentence by sentence, in an activity known as Detailed Reading. With young children beginning to read, the teacher first writes sentences from the reading story on cardboard strips. The children are then shown how to point at each word in the familiar sentence as they say them, and then to cut up words and word groups, put them back in the sentence and read it again, until they can read the sentence accurately. This practice is a powerful catalyst for children to make the semiotic journey from the spoken to the written medium, via visual and manual modalities. Older students are orally guided to identify each group of words in each sentence from the reading text, using cues for their meaning and position in the sentence. The students then mark the words with highlighters or underlining, and their meaning may be elaborated. These techniques are shown in Image 9.5.



(a)



(b)

IMAGE 9.4 Preparing before reading: Spoken–visual

(a)



(b)

IMAGE 9.5 Reading manually with sentence strips and highlighters

The manual practice of manipulating and marking wordings powerfully reinforces the movement between aural and visual modalities, leading to reading with understanding and fluency. These activities materialize the semiotic relations between meaning and wording and between spoken and written expression. When a child is first learning to read, they must consciously recognize the Token = Value relation between the spoken wordings they know and the written wordings on the page. The acts of pointing and naming, and cutting up and manipulating word groups and words in a sentence, focus the child's attention on these functional segments, as they are both physically and semiotically in control of these objects-as-meanings.

Once this control has been mastered, the Token = Value relation of spoken and written expression evaporates, as graphology replaces phonology as the medium of expression. That is, experienced readers do not translate from written to spoken expression in order to recognize meanings. Martin (2006) describes this as a shift in the child's understanding of reading from 'book tells us meaning' to 'writing realizing meaning'; the semiotic relation shifts from projection (a says 'b') to identification ($a = b$). The automaticity of written expression then allows the reader to focus their conscious attention wholly on semantic patterns in the content plane. This is what Vygotsky observes in the development of 'higher psychological functions':

At the centre of development during the school age is the transition from the lower functions of attention and memory to higher functions of voluntary attention and logical memory . . . the intellectualisation of functions and their mastery represent two moments of one and the same process – the transition to higher psychological functions. We master a function to the extent that it is intellectualised. The voluntariness in the activity is always the other side of its conscious realization. (Wertsch 1985:26, cited in Hasan 2004)

Detailed Reading aims to make 'the intellectualization of functions and their mastery' explicit. The combination of spoken, visual and manual modalities enhances learners' voluntary control, supporting them to distinguish patterns in both expression and content planes. This is a complex activity that supports all students in a class to read and interpret at both instructional levels, of the meanings within the text and the general semantic patterns they instantiate, and so requires careful planning. It is applied at all levels of education, in all curriculum fields, to enable all students to read texts with detailed critical understanding, and to identify patterns of meaning that they can then recognize in other texts, and apply in their writing.

The carefully planned interaction cycles in Detailed Reading are illustrated here with a Year 6 class (Table 9.5). The school is in a low socioeconomic area, and most students are from non-English speaking backgrounds. The text that they are reading is from a novel about an earthquake, that would normally be well beyond the independent reading capacities of most. The teacher has prepared and read the first chapter to the class, and has selected a short passage from it for Detailed Reading.

In the first step, the teacher prepares the class to follow the first sentence of the passage as she reads it to them. She begins with the visual mode, directing students' attention to its position in the text. She then generalizes the experiential meanings in the sentence as 'the sound' and 'where the sound came from'. Then she instructs them to look at the sentence as she reads it aloud.

Table 9.5 Read sentence

	Move		Purpose
Teacher	A2	So if we look at that very first sentence,	position
	K1	the writer begins by describing the sound to us, OK, and just where the sound came from .	grammatical meanings
	A2	So if we have a look at it . . .	read along
Students	A1	[look at sentence]	
Teacher	K1	. . . it says, <i>It started with a long low roar that seemed to be approaching from the north of the city.</i>	

In terms of modalities, the movement here is from directing students' visual attention to the text, to a spoken preparation that directs attention to segments of meaning in the sentence, to visual attention to the written sentence and aural attention to the words of the sentence, as the students follow the written words as the teacher reads aloud.

In some respects, teaching reading here displays similarities with teaching digging in Table 9.1, with the instructions to 'look' and the directions to positions 'very first' and 'begins'. A key difference is the use of meta-language to direct attention to semiotic things, including the grammatical segment 'sentence', the lexical category 'sound', the grammatical functions 'describing the sound' (Epithet+Thing³) and 'where the sound came from' (Medium+Place), and the expression-content relation 'it says'. Meta-language is the semiotic equivalent of gesturing manually and referring exophorically in material activities, illustrated in Tables 9.1 and 9.4. But it is far more diverse as the semiotic phenomena it directs attention to are more complex.

In addition, the metadiscourse is concerned not just with the text, but with the pedagogic activity and social relations. The students are expected to recognize the interpersonal metaphor, 'if we have a look', as commanding their attention, at the same time as including them with the teacher. Secondly, 'the writer begins by describing the sound to us' makes the communicative act of writing explicit in the reading activity.

In the next cycle, the teacher prepares the class to identify the first element in the sentence. She begins by directing attention to its position in the sentence, then generalizes the meaning as 'what the earthquake did', then asks a particular student to say the wording. When the student says the words, the teacher affirms her, and then instructs the class in exactly what words to highlight.

Table 9.5 (cont) Identify wording

	Move		Purpose	
Teacher	K1	So in that very first sentence, right at the beginning . . .	Prepare	position
Students	A1	[look at sentence]		attend
Teacher	K1	. . . it tells us what the earthquake did.		grammatical meaning
	dK1	What did it do? Chanila?	Focus	Meaning & position
Student	K2	<i>It started with a long low roar.</i>	Identify	wording
Teacher	K1	That's great, fantastic. So <i>It started.</i>	Evaluate	affirm
	A2	So let's highlight <i>It started.</i>	Highlight	instruct
Students	A1	[highlight wording]		highlight

This cycle shares many similarities with the parent-child interaction in Table 9.4. The students' task is to identify text elements. The teacher prepares by directing attention and interpreting meanings, and evaluates with affirmation. But in addition she uses a 'Focus' question to elicit a response from one student, and a 'Highlight' instruction to ensure that all students mark the same words (Martin 2006). Here the direction of attention is from the position in the text and sentence, to the grammatical function 'what the earthquake did' (Medium+Process), to the grammatical structure *It started*. Instead of manually pointing, the teacher explicitly states the position (K1), which implicitly demands the students look at the position (A1).

The meaning cue is then restated as a Focus question, directed to a particular student. This question is labelled dK1, for 'delayed primary knower', as the teacher already knows the answer. The purpose of dK1 questions, which are pervasive in classroom discourse, is to get students to attend to and repeat information. They function to hand control over to students to do a task themselves, rather than simply listening to the teacher, and then allow the teacher to evaluate and elaborate on students' responses.

As one student says the wording aloud (K2), all the others are also seeing it and reading it silently, interpreting it in terms of the semantic category given by the teacher. The teacher's affirmation and repetition of the wording intensifies the affective value of the identifying activity, then the manual activity of highlighting the wording (A1) cements its value for each learner.

As they repeatedly do the task of identifying word groups from such cues, all students rapidly come to consciously recognize relations between grammatical functions, denoted by the natural metalanguage of *who* or *what*, *what did/happened*, *where*, *when*, *how*, and so on, and the written grammatical structures that realize these functions. (At this stage a more technical metalanguage is not yet required for students to identify such function structures.)

Next, the teacher capitalizes on the shared foundation of successful activity, positive feeling and understanding of the wording's semantic value, to add another layer of meaning to it. Here she moves from the meaning within the sentence (Medium+Process) to its meaning beyond the sentence (reference to previous mentions of the earthquake). In elaborating phases such as this, student responses shift from identifying wordings in the text to selecting meanings from their memories.

Table 9.5 (cont) Elaborate discourse function

1	Teacher	K1	Now I used the word earthquake, because we know it's an earthquake.	Prepare	precedingcycle preceding text
	Teacher	dK1	What have they used instead of earthquake? What's the word they've used, there to begin that paragraph? Bonita?	Focus	reference item wording position
	Student	K2	<i>It.</i>	Identify	wording
2	Teacher	K1	<i>It.</i>	Evaluate	affirm
	Teacher	K1	And we can use <i>It</i> because we already know what <i>It</i> is.	Prepare	reference item
	Teacher	dK1	<i>It</i> is . . . ?	Focus	position
	Students	K2	The earthquake.	Select	referent
	Teacher	K1	OK, fantastic.	Evaluate	affirm
3	Teacher	dK1	Now what do we call little words like 'it' that refer to other words?	Prepare	metalanguage
	Students	K2	Pronoun.	Select	metalanguage
	Teacher	K1	Exactly, <i>it</i> is a pronoun.	Evaluate	affirm

This elaboration includes three cycles. In cycle 1, the teacher first directs attention to remembering the preceding preparation 'I used the word earthquake', then to remembering the preceding mentions in the text 'we know it's an earthquake', then to the discourse function 'what have they used instead of earthquake' (anaphoric reference), then the wording 'what's the word they've used' (a pronoun), then the position in the text 'there to begin that paragraph'.

In cycle 2, she uses affirmation and repetition to intensify students' attention to the discourse function, getting them to repeat the referent back to her, and strongly affirming them. This creates a firm semantic basis in cycle 3 for asking the class to remember a linguistic term that denotes a word class and its discourse function, 'pronoun'. Repetition and affirmation of terms like this, within

elaboration phases, will eventually enable all students in the class to remember and use such metalanguage appropriately. In this way, the class builds an explicit, systematic and consistent metalanguage, through experiencing instances in actual texts.

The next element to be identified is a grammatical metaphor *long low roar*, which the teacher prepares by glossing as a ‘sort of sound’. As ‘roaring’ is actually a process, and the qualities *long low* are normally associated with concrete objects, many children may not recognize this lexical item without such support.

Table 9.5 (cont) Prepare wording

Teacher	K1	Now, so the earthquake started, now when <i>it started</i>	Prepare	position
	dK1	what sort of sound did it make?		lexical meaning
	K1	It tells us it started <i>with</i> something.		position
Teacher	dK1	What was it that <i>it started with</i> ? Chanila?	Focus	position
Student	K2	<i>Long low roar.</i>	Identify	wording
Teacher	K1	Fantastic.	Evaluate	affirm
	A2	So let's highlight <i>long low roar</i> .	Highlight	instruct

Again the cycle of attention begins here with the position in the sentence ‘when it started’, then the lexical category ‘what sort of sound’, then the position of the grammatical structure ‘it started *with* something’, so the students know that the wording follows *with*, making it easier to identify. And again one student says the words, is affirmed, and the class is directed to highlight the words.

Next the students are guided to interpret the conceptual image evoked by *long low roar*, by reference to their previous experience.

Table 9.5 (cont) Elaborate field

1	Teacher	dK1	Now can you think of something else?	Prepare	memory
		dK1	What else do we associate with that roar sound?		everydayfield
		dK1	What do you think?		memory
	Student	K2	A lion roar.	Select	field
	Teacher	K1	OK, a lion roars.	Evaluate	affirm

(Continued)

2	Teacher	dK1	What else do we associate with a roar? Another thing?	Prepare	field
	Student	K2	The sea can roar.	Select	field
	Teacher	K1	The <i>sea</i> ,	Evaluate	affirm
		K1	on a really stormy day.	Elaborate	field
3		K1	Yes it does give a bit of a roar.	Evaluate	affirm
	Teacher	dK1	Justin?	Focus	memory
	Student	K2	A tornado?	Select	field
	Teacher	K1	Yes.	Evaluate	affirm
		K1	Those other natural disaster types of sounds.	Elaborate	curriculumfield
		K1	Yes.	Evaluate	affirm

Here the teacher begins by directing students' attention to their memories of other fields (dK1). As each student proposes a field associated with roaring sounds (K2), the teacher affirms and may elaborate the field (K1). Finally she takes advantage of one response, to associate 'a tornado' with the curriculum field of 'natural disasters' that the class is studying.

The aim of this exchange is to reveal the Token-Value relation between an elaborated lexical item or metaphor in the text, and readers' material experience, that is, its role in evoking imagery. Even though the students select responses from their own memories, the primary knower remains the teacher, as she affirms and elaborates. In so doing she includes their personal experience within the field of school knowledge. This repeated affirmation and inclusion intensifies students' attention to the Token-Value relation between written wordings and readers' imaginations.

On this basis, the teacher next guides their attention to remembering a more specific field, the roar of a jet engine. This excites the memory of one student, which she takes advantage of to relate the jet roar to a specific feature of the story's field 'ground starts to shake'.

Table 9.5 (cont) Elaborate discourse function

Teacher	dK1	Ever heard a jet?	Prepare	everydayfield
	K1	Oh, you've all been to the airport.		memory
	dK1	The roar of the engine?		field
Student	K2	An airshow.	Select	field
Teacher	K1	The airshow, exactly.	Evaluate	affirm
	K1	The whole ground starts to shake.	Elaborate	field
	K1	Exactly.	Evaluate	affirm

(Continued)

Table 9.5 (Cont'd)

K1	So that sound vibration even makes the ground move, doesn't it?	Elaborate	field
K1	Yes, fantastic.	Evaluate	affirm
K1	And it starts off low, and builds up, doesn't it?	Elaborate	field
K1	So we have this roaring sound, but it starts off long . . . low.		discourse

Again the teacher uses repeated affirmation here, to intensify the class' attention to the next elaboration, which focuses on two features, the qualities 'long low' and process 'starts'. The goal of this sequence is to direct students' attention to the function of these elements in the discourse structure of the text 'it starts off long . . . low'. That is, tension builds through the text passage as the earthquake approaches. A key technique the author uses to build tension is to start *low* and uncertain *seemed to be approaching*.

The students need to understand both the meaning of each of these elements within the sentence, and their discourse function in the text. The teacher's strategy is to relate the local meaning to their own experience, drawing their attention to aspects that are relevant to the discourse function. As the Detailed Reading of the passage continues, she will point out the global discourse patterns of mounting tension, and remind them of the aspects of each wording that contribute to this pattern.

Options for Pedagogic Interactions

From analyses of pedagogic interactions that are exemplified in Tables 9.1, 9.4 and 9.5, the following systems of options emerge, for each phase in an exchange, including initiating phases, responses and feedback. First, initiating moves (Figure 9.4) either instruct students to perform an action, or elicit a verbal response. Instructions and elicitations may be directed to an individual student, or to the class as a whole.

Instructing moves may demand an action, that may be associated directly with the pedagogic activity, for example, 'throw more soil over here', 'let's highlight *It started*', or the demanded action may be behavioural control, such as admonishments. Alternatively, instructing moves may direct learners' attention, either to an object or text, 'if we look at that very first sentence', or to the learner's memory, 'what else do we associate with that roar sound'. Instructing moves may also be non-verbal, where the learner's responding action can be assumed, for example, Mother: [points to the wolf].

Eliciting moves may deliberately **prepare** students for a successful response, by providing specific criteria. Tables 9.4 and 9.5 give many examples of such

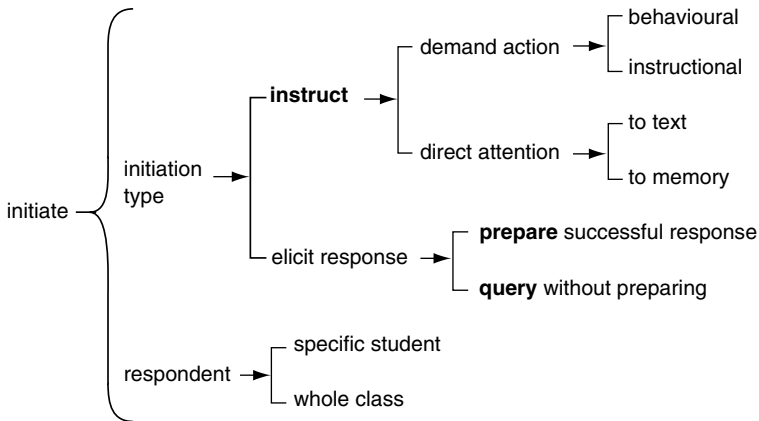


FIGURE 9.4 Options for initiating

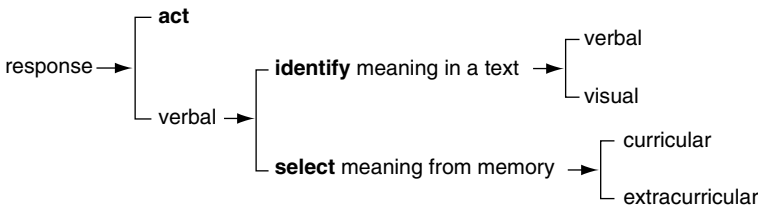


FIGURE 9.5 Options for responding

preparations. In Table 9.4, these may include a preparation directed to the whole class, and a focus question inviting one student to respond. However, teachers more often than not elicit with a **query** that does not provide such criteria. When queries are directed to the whole class, only those students who can infer the desired response are able to respond successfully. There are no examples of such queries in the parent-child reading and designed interaction in Tables 9.4 and 9.5, but see Rose (2004) for analysed instances, and almost any transcript of classroom interactions will include copious examples.

Secondly, responses (Figure 9.5) are either to **act** on an instruction, for example, [starts to dig], [points to picture of a tree], [highlight wording], or to give a verbal response to an elicitation (preparation or query). Verbal responses either **identify** features in texts, including wordings (verbal) and images (visual), or **select** information from the learners' memory. Remembered information may have been previously taught (curricular), for example, 'what do we call little words like "it"? – pronoun'. Or the response may be from personal experience (extracurricular), 'What else do we associate with that roar sound? – a lion roars . . . the sea can roar . . . the airshow'.

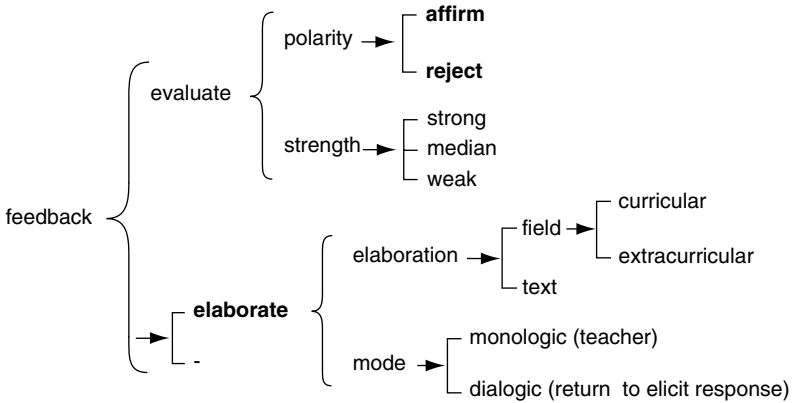


FIGURE 9.6 Options for feedback

Thirdly, feedback moves (Figure 9.6) always involve evaluations that either **affirm** or **reject** the response, with more or less strength. For example, affirmations may range from ‘yep’ to ‘fantastic’ and are often intensified by repetition; rejections range between qualifying responses, ignoring, negating or even admonishing. Where affirmations function to enhance learning capacity and engagement, rejections may have the opposite effect, particularly for students with weak learner identities. In the stratified context of the typical classroom, affirmations and rejections can thus serve to differentiate students. On the other hand, where differentiation is not an issue, an interplay of affirmation and rejection can serve to guide learners towards a goal, as in Table 9.1.

In addition feedback may **elaborate** on the response, providing more information about either the text or the field. Again the field of elaboration may be curricular or extracurricular. The mode of elaboration may be a teacher monologue, or a dialogue with students. If the elaboration is dialogic, the cycle begins again with eliciting a response (usually elicited by the teacher but students may also ask questions that demand elaborations). Elaborations are optional (shown by the minus option in Figure 9.6), but teachers typically use students’ responses as stepping stones in a lesson, expanding them with more technicality or detail, either strengthening the boundaries between everyday and esoteric knowledge, or traversing back and forth between them, as illustrated in Tables 9.4 and 9.5.⁴

Conclusion: Tools for Redesigning Pedagogy

One thing that stands out in our analyses is the central function of evaluation in pedagogic interactions, realized as the obligatory KI move in an exchange. For Bernstein, ‘the key to pedagogic practice is continuous evaluation . . .

evaluation condenses the meaning of the whole [pedagogic] device' (1996:50). In Tables 9.1, 9.4 and 9.5, evaluations are used to guide learners towards a goal and to enhance their learning capacity and engagement. But in the standard initiation-response-feedback cycles that pervade classroom practice across the world (see Alexander 2000 for variations), evaluations also rank students on their capacity to respond successfully to teacher queries. In a stratified socio-economic order, the broadest social function of the education system is to reproduce unequal outcomes. So the meaning condensed in each evaluation of student responses is one of inequality.

From their first day in school, children start to learn, not only that some responses are more successful than others, but that some students are more successful at responding than others. Naturally, the more successful responders are those who have been well prepared by extensive talk-around-text in parent-child reading. They will not only be evaluated as more 'able' learners, but will consistently receive the lion's share of teacher affirmation, as feedback to their responses. By these means, the continuous micro-interactions of classroom discourse serve to relentlessly construct differential learner identities, as more or less 'able', naturalizing the different experiences that children arrive at school with. These identities are internalized by children and cemented over the years of schooling by classroom evaluations, by formal assessments and by 'streaming' of students into different classes and different activities within classes. Bernstein portrays this process baldly:

The school must disconnect its own internal hierarchy of success and failure from ineffectiveness of teaching within the school and the external hierarchy of power relations between social groups outside the school. How do schools individualize failure and legitimize inequalities? The answer is clear: failure is attributed to inborn facilities (cognitive, affective) or to the cultural deficits relayed by the family which come to have the force of inborn facilities. (1996/2000:5)

The *Reading to Learn* methodology subverts this universal inequity by redesigning the classroom pedagogic genre, in its four dimensions. First, reading is recognized as the primary mode of learning in school, and spoken, written, visual and manual modalities are systematically deployed to teach the skills required for reading at each stage of schooling. Secondly, the activity of classroom learning is carefully redesigned to consistently provide students with the preparation they need to succeed in each task, and then to use their success as a basis for extending their understanding. Thirdly, the social interaction of learning is redesigned to ensure that all students are continually successful at the same task level, and are continually affirmed. The redesign of these three dimensions in the pedagogic register enables the instructional field to focus explicitly on patterns of discourse, at the same time as teaching the curriculum topics that these patterns realize. The payoff for all this design work is that students' results,

for teachers trained in *Reading to Learn*, are consistently twice to four times beyond expected rates of growth (Culican 2006, Rose et al. 2008), accelerating the learning of all students, while rapidly closing the gap in their levels of achievement.

Notes

¹ Anecdotes are not resolved like narratives, but conclude with a Reaction (Martin & Rose 2008).

² The model of pedagogic genre is derived from Bernstein's model of 'pedagogic discourse', including an instructional discourse 'which creates specialized skills and their relationship to each other', but is embedded in and dominated by a regulative discourse 'which creates order, relations and identity' (1996/2000:46). Extending Martin (1999), Bernstein's regulative discourse is re-interpreted as the pedagogic register, including the field of learning activities, the tenor of pedagogic relations, and the mode of learning. These three variables in pedagogic register project the instructional field of skills and knowledge to be acquired.

³ Grammatical functions, such as Epithet, Thing, Medium, Place, are described in Halliday 1994/2004 and Martin and Rose 2007a.

⁴ Some of the points made in this analysis have been identified by neo-Vygotskian activity theorists such as Mercer (2000) or Wells (1999). Key differences here include:

- detailed analysis of the functions of each exchange move, informing interpretations of learning interactions,
- emphasis on the Prepare (or 'initiate') phase, in addition to the Elaborate (or 'feedback') phase that neo-Vygotskians value for developing 'higher order thinking',
- the role of the teacher/parent in preparing and elaborating, to enable and extend learning, far beyond what is possible in peer-peer interactions,
- the focus on teaching reading as the grounding for elaborating meanings, in contrast to privileging 'talk'.

These analytic developments are crucial for designing pedagogic interactions that enable all students to read successfully.

References

- Australian Bureau of Statistics (1994, 2004). *Australian Social Trends 1994 & 2004: Education - National summary tables*. Canberra: Australian Bureau of Statistics, www.abs.gov.au/ausstats.
- Adams, M.J. (1990). *Beginning to read: Thinking and learning about print: A summary*. Urbana-Champaign: University of Illinois.
- Alexander, R. (2000). *Culture and pedagogy: International comparisons in primary education*. London: Blackwell.
- Bernstein, B. (1990). *The structuring of pedagogic discourse*. London: Routledge.

- Bernstein, B. (1996). *Pedagogy, symbolic control and identity: Theory, research, critique*. London: Taylor & Francis. [Revised Edition 2000].
- Cloran, C. (1999). Contexts for learning. In Christie (Ed.), *Pedagogy and the shaping of consciousness: Linguistic and social processes* (pp. 31–65). London: Cassell.
- Culican, S. (2006). Learning to read: Reading to learn, a middle years literacy intervention research project, final report 2003–4. Catholic Education Office: Melbourne. [www.cecv.melb.catholic.edu.au/Research and Seminar Papers](http://www.cecv.melb.catholic.edu.au/Research%20and%20Seminar%20Papers).
- Halliday, M.A.K. (1993). Towards a language-based theory of learning. *Linguistics and Education*, 5(2), 93–116.
- Halliday, M.A.K. (1994/2004). *An introduction to functional grammar* (2nd edn). London: Edward Arnold.
- Hasan, R. (2004). Semiotic mediation and three exotropic theories: Vygotsky, Halliday and Bernstein. In J. Muller, B. Davies & A. Morais (Eds), *Reading Bernstein, researching Bernstein* (pp. 30–43). London: RoutledgeFalmer.
- Hasan, R. & Cloran, C. (1990). A sociolinguistic interpretation of everyday talk between mothers and children. In M.A.K. Halliday, J. Gibbons & H. Nicholas (Eds), *Learning, keeping and using language. Vol. 1: Selected papers from the 8th World Congress of Applied Linguistics*. Amsterdam: John Benjamins.
- Kellogg, S. (1997). *The three little pigs*. New York: Morrow Junior Books.
- Kensinger, E.A. (2004). Remembering emotional experiences: The contribution of valence and arousal. *Reviews in the Neurosciences*, 15(4): 241–251.
- Martin, J.R. (1999). Mentoring semogenesis: ‘Genre-based’ literacy pedagogy. In F. Christie (Ed.), *Pedagogy and the shaping of consciousness: Linguistic and social processes* (pp. 123–155). London: Cassell.
- Martin, J.R. (2006). Metadiscourse: Designing interaction in genre-based literacy programs. In R. Whittaker, M. O’Donnell & A. McCabe (Eds), *Language and literacy: Functional approaches* (pp. 95–122). London: Continuum.
- Martin, J.R. & Rose, D. (2005). Designing literacy pedagogy: Scaffolding asymmetries. In R. Hasan, C.M.I.M. Matthiessen & J. Webster (Eds), *Continuing discourse on language* (pp. 251–280). London: Equinox.
- Martin, J.R. & Rose, D. (2007a). *Working with discourse: Meaning beyond the clause* (2nd edn). London: Continuum.
- Martin, J.R. & Rose, D. (2007b). Interacting with text: The role of dialogue in learning to read and write. *Foreign Languages in China*, 4(5): 66–80.
- Martin, J.R. & Rose, D. (2008). *Genre relations: Mapping culture*. London: Equinox.
- McGee, L.M. (1998). How do we teach literature to young children? In S.B. Neuman & K.A. Roskos (Eds), *Children achieving: Best practices in early literacy* (pp. 162–179). New Jersey: Newark International Reading Association.
- Mercer, N. (2000). *Words & minds: How we use language to work together*. London: Routledge.
- Painter, C. (1996). The development of language as a resource for thinking: A linguistic view of learning. In R. Hasan and G. Williams (Eds), *Literacy in Society* (pp. 50–85). London: Longman.
- Painter, C. (2003). The ‘interpersonal first’ principle in child language development. In G. Williams & A. Lukin (Eds), *Language development: Functional perspectives on species and individuals* (pp. 133–153). London: Continuum.
- Pilkington, D. (1996). *Follow the rabbit-proof fence*. St Lucia: University of Queensland Press.

- Rose, D. (2001). *The western desert code: An Australian cryptogrammar*. Canberra: Pacific Linguistics.
- Rose, D. (2004). Sequencing and pacing of the hidden curriculum: How indigenous children are left out of the chain. In J. Muller, A. Morais & B. Davies (Eds), *Reading Bernstein, researching Bernstein* (pp. 91–107). London: Routledge Falmer.
- Rose, D. (2005). Narrative and the origins of discourse: Construing experience in stories around the world. *Australian Review of Applied Linguistics Series S19*, 151–173.
- Rose, D. (2006). Reading genre: A new wave of analysis. *Linguistics and the Human Sciences*, 2(2), 185–204.
- Rose, D. (2007). A reading based model of schooling. *Pesquisas em Discurso Pedagógico* 4: 2, www.maxwell.lambda.ele.puc-rio.br.
- Rose, D. (2008). Writing as linguistic mastery: The development of genre-based literacy pedagogy. In D. Myhill, D. Beard, M. Nystrand & J. Riley (Eds), *Handbook of writing development* (pp. 151–166). London: Sage.
- Rose, D., Rose, M., Farrington, S. & Page, S. (2008). Scaffolding literacy for indigenous health sciences students. *Journal of English for Academic Purposes*, 7(3), 166–180.
- Saunders, P., Hill, T. & Bradbury, B. (2007). *Poverty in Australia: Sensitivity analysis and recent trends*. Sydney: Social Policy Research Centre, University of New South Wales www.sprc.unsw.edu.au/nl/NL102/ASPC-2009.pdf.
- Tomasello, M. (2000). *The cultural origins of human cognition*. Cambridge, MA: Harvard University Press.
- Torr, J. (2004). Talking about picture books: The influence of maternal education on four-year-old children's talk with mothers and pre-school teachers. *Journal of Early Childhood Literacy*, 4(2), 181–210.
- Trevarthen, C. (2005). First things first: Infants make good use of the sympathetic rhythm of imitation, without reason or language. *Journal of Child Psychotherapy*, 31(1), 91–113.
- Unsworth, L. & Wheeler, J. (2002). *Re-valuing the role of images in reviewing picture books. Reading literacy and language*. London: Blackwell Synergy.
- Wells, G. (1999). *Dialogic inquiry: Toward a sociocultural practice and theory of education*. Cambridge: Cambridge University Press.
- Wertsch, James V. (1985). *Vygotsky and the social formation of mind*. Cambridge, MA: Harvard University Press.
- Williams, G. (1995). Joint book reading and literacy pedagogy: A socio-semantic examination. unpublished PhD Thesis, Macquarie University.

Part Four

Imaging Representations of Meaning

This page intentionally left blank

Chapter 10

Visualizing Logogenesis: Preserving the Dynamics of Meaning

Michele Zappavigna
University of Sydney

Introduction

Discourse analysts are concerned with understanding patterns in language. These patterns are often highly complex involving relationships between many variables and across many dimensions. Data is said to have high dimensionality when there are multiple values that uniquely identify a data point. For example, we might model a point in two dimensions using X and Y axes, in three dimensions using X, Y and Z axes and in N dimensions using N number of axes. As the degree of dimensionality increases, our capacity to perceive the data as meaningful diminishes. This is in part because metaphors that we can intuitively understand such as three-dimensional space cannot be applied to visualize the data directly. The problem is pertinent for linguistics because language has high dimensionality. So for a discourse analyst wanting to understand patterns in language, some kind of technological support is required.

Since discourse analysts are concerned both with understanding relationships between variables and with understanding how patterns of variables unfold in a text, the kind of support that they need should be dynamic. In other words it needs to be able to account for logogenesis: the unfolding of meaning in a text. Fortunately, advances in computer technology now afford us the possibility of annotating, managing and visualizing highly complex data. We can now track multiple relationships between variables unfolding in time or along other dimensions. As a result we have the potential to model logogenesis and understand how meanings work together as they unfold in real-time and across semiotic modes. This chapter explores three enabling text visualization techniques that use three different methods to represent the temporal sequencing of a text: text arcs (Wattenberg 2002), streamgraphs (Byron & Wattenberg 2008, Havre et al. 2002) and animated networks (Fry 2000b).

I begin by considering the complementarity of dynamic and synoptic perspectives on texts and introduce the concept of logogenesis as it is theorized in Systemic Functional Linguistics (SFL). I then introduce the field of Text Visualization and describe some of its principles and methods with a view to

suggesting how text arcs, streamgraphs and animated networks might be used by functional discourse analysts as tools for exploring text.

Preserving Logogenesis

The toolkit available to the Systemic Functional linguist is currently largely composed of strategies suited to a synoptic gaze. SFL has focused much effort over the last half-century on modelling the meaning potential of language as a semiotic system from a paradigmatic perspective. This effort has centred upon using system networks (resembling, visually, tree diagrams turned on their side), as a means of displaying choices in systems of possible meanings. Complementary to the paradigmatic view of ‘what could go instead of what’ (Halliday & Matthiessen 2004:22) that these system networks afford, is a syntagmatic perspective that considers ‘what goes where when’.

The preoccupation with paradigms has directly impacted how strata are modelled in SFL as levels of abstraction. For Matthiessen (2007), strata are subsystems at particular orders of abstraction that are held together in a *realization* relationship whereby patterns at a lower level are realized by those at higher levels. The co-tangential circles representation in Figure 10.1 presupposes a paradigmatic gaze in which features are (metaphorically) distributed across the two-dimensional plane rather than sequenced. This is in keeping with the fact that realization affords minimal information about sequencing beyond the small window of structure specified in function structures.

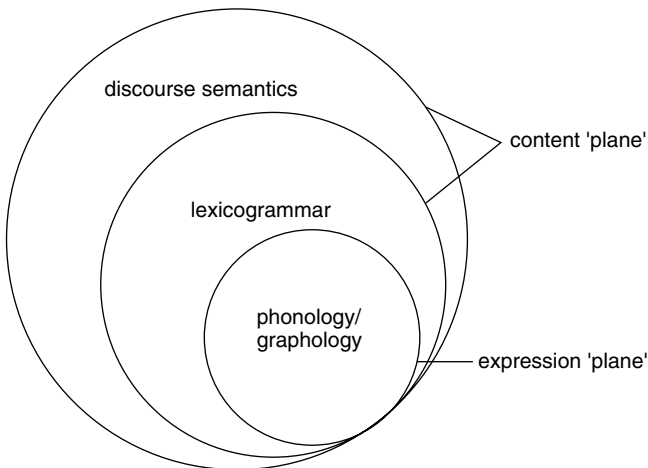


FIGURE 10.1 Stratification conceived as levels of abstraction (Halliday & Martin 1993)

However, a syntagmatic perspective considers an additional dimension when modelling meaning-making, which is change. Halliday recognized that once contiguity relations are added to a paradigmatic modelling strategy that the linguist is 'taking on a dynamic commitment' (Halliday 1991:40), that they are involved in modelling change. Halliday (1993) described three kinds of semiotic change or 'semogenesis': *logogenesis*, the unfolding of text, *phylogenesis*, the evolution of culture, and *ontogenesis*, the development of the meaning-making potential of a human over time. In this chapter I am concerned with logogenesis and thinking about meanings in unfolding in texts 'dynamically as currents flowing through a stratified semiotic system' (Halliday 1991:40).

In modelling change we come up against the problem of the limitations of the information that systemic probability represented using system networks can provide (Zappavigna et al. 2008). Lemke (1991) suggests the emergent complexity of language as a dynamic open system. In contrast to the position that argues that interdependency can be modelled at the stratum of lexicogrammar, Lemke suggests that 'relations of interdependency' are dynamic semantic relations (Lemke 1991:24):

If we imagine the description of dynamic systems to be mainly a matter of the *dynamic weightings of selection probabilities*, then we wish to know how the selections 'up to now' condition the probabilities for selections 'now'. (Lemke 1991:26)

In other words, unless we have a dynamic model of register, we are unable to reset the 'probability weightings in just the right way just in time for each pass through the network' (Lemke 1991:241). He notes that these kinds of sets or probabilities can be described mathematically 'and amount, in fact precisely to the re-weightings of dynamic systems needed for text production to produce texts of recognizable social formations' (Lemke 1991:31).

In accord with Lemke's position outlined above, Martin (2004:341–342) foregrounds the importance of maintaining the logogenetic integrity of texts when he argues that 'as social discourse analysts we need to guard against studies that submerge unfolding texture in processes of counting and averaging that look for trends across texts rather than contingencies within them'. While, in accounting for the unfolding of a text, it may be clear that we wish to avoid approaches that characterize the text as a 'bag of words', we also want to avoid the position where the text is reduced to a collection of clauses:

It doesn't matter how many clauses we analyse, it's only once we analyse meaning beyond the clause that we'll be analysing discourse. And we need to analyse discourse right along the cline of instantiation if we want to make sense of the semiotic weather we experience in the ecosocial climate of our times. (Martin & Rose 2003:272)

In order to make such a jump out of the clause, we need means of communicating the kinds of patterning that we will find. Static forms of representation such as bar charts will not meet our needs because they reduce the complexity to a single value visualized in two dimensions. Instead we perhaps require techniques that assist linguists in exploring the patterning of annotations that they have made to a text across as many dimensions as are necessary.

Rather than reducing the annotated text to a table of statistical values we might employ various kinds of text visualization to achieve a dynamic lens on the data. For example, consider the description, provided by a developer of a text visualization system that presents texts in a three-dimensional network:

Instead of focusing on numeric specifics . . . the piece provides a qualitative feel for perturbations in the data, in this case the different types of words and language used throughout the book. This provides a qualitative slice into how the information is structured. On its own the raw data might not be particularly useful. But when relationships between data points can be established, and these relationships expressed through movement and structural changes in the on-screen visuals, a more useful perspective is established. (Fry 2000a:67)

Such a 'qualitative slice' may be of great use to the discourse analyst because it emphasizes relationships between linguistic features in texts as they are interwoven to create particular meanings. What is presented here is not an argument against 'counting' these features, but a suggestion that should not toss out information about their sequencing.

Patterning and Co-occurrence in Systemic Functional Linguistics

When we turn from modelling potential paradigmatically to considering the unfolding of meanings realized in texts, different patterns of coordination need to be foregrounded. Logogenesis is clearly more than the text unfolding in a simple linear progression. The orchestration of a text might involve different kinds of semiotic crescendo and decrescendo as different meanings emerge and fade prosodically. This potential in discourse, a kind of 'snowballing' (Martin, personal communication, 16 July 2008) of meaning, is apparent when manually analysing, for example, evaluative language using Appraisal theory (Martin & White 2005). However, we currently have an impoverished repertoire for talking synoptically about this kind of patterning. A naive representation of a text accumulating particular meanings while shedding others is presented in Figure 10.2. Current strategies for annotating patterning include colour-coding, tabular organization, and, in some cases, annotations in formats such as XML. While any kind of annotation is a useful first step, the problem of

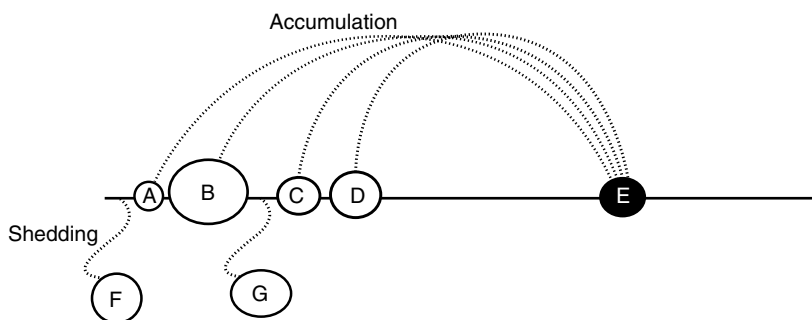


FIGURE 10.2 Representing accumulation and shedding of meaning in a text as it unfolds

making visible patterns that emerge in sequences that exceed a single page or screen is significant. This is a problem of tractability. Until we have a way of representing extended patterning we are limited to probing small co-patterning of meaning deemed qualitatively relevant to the particular questions the analyst is asking about the text.

To be true to the unfolding of a genuine, multimodal text, however, we need to find ways of analysing and representing the unfolding of two kinds of co-occurrence in actual texts: co-occurrence across unfolding modes, for example, simultaneous use of a particular intonation and a particular gesture, and co-occurrence within the same text sequence, for example, use of *AFFECT* together with *GRADUATION*¹ in a clause in the unfolding text. Time, in the second type of co-occurrence is not clock time but instead a form of 'text time' dependent on the dimension of meaning that the discourse analyst is interested in exploring. The latter type of co-occurrence has begun to be explored in the notions of coupling (Martin 2000) and syndromes (Zappavigna et al. 2008) introduced earlier. Coupling refers to meanings that are co-related in a text, for example, relationships between evaluative and ideational meaning integral to construing shared values in a community (Knight 2008). Syndromes are larger-scale configurations involving multiple associations between different meanings involved in the overall rhetoric being developed as the text unfolds. I will now suggest how the domain of text visualization may offer assistance to linguists trying to analyse unfolding textual patterns.

A Brief Introduction to Text Visualization

Text Visualization is an emergent field, a cousin of Scientific Visualization, and often classified as part of Information or Knowledge Visualization. Those interested in visualizing text often have a background in both computer science and digital art, bringing both technical and aesthetic skills to the endeavour

(see, for example, Martin Wattenberg at www.bewitched.com, Ben Fry at www.benfry.com and Lee Byron at www.leebyron.com). The field tends to naturalize the encoding of language as a product, typically written 'raw text' repositioned as 'data'. The 'raw text' is a string of characters with lexical items as the focus of inquiry. This attention to lexis is partly pragmatic, resulting from the difficulty of training a computer to identify linguistic features of greater complexity (e.g. clauses). Thus, most visualization techniques are 'word'-based, often excluding apparently irrelevant 'stop words' (often function words). They are also constituency-oriented, chunking the text into units. **These two limitations mean that, to date, visualization techniques have not been used to explore meaning beyond the clause in all its prosodic complexity.** In short, the area has inherited some of the bad habits of generative computational linguistics.

However, it is entirely possible that we may move beyond lexis and out of the clause via text visualization. Visualization offers us an important opportunity to gain synoptic views of the text that concurrently preserve a dynamic view of logogenesis. This is because many visualization techniques allow the text to be manipulated along multiple dimensions while allowing us to track multiple kinds of relationships between features.

Visualization, in general, is concerned with finding methods of representation that best leverage the characteristics of human visual perception to make complex data meaningful. A synoptic view of a text or corpus should assist the viewer by lessening the cognitive burden of perceiving patterns in texts:

For any reader, the rather slow serial process of mentally encoding a text document is the motivation for providing a way for them to instead use their primarily preattentive, parallel processing powers of visual perception . . . The goal of text visualization, then, is to spatially transform text information into a new visual representation that reveals thematic patterns and relationships between documents in a manner similar if not identical to the way the natural world is perceived. (Wise et al. 1995:51–52)

Thus, a visualization will only be effective to the extent that it can profitably make use of preattentive perceptual capabilities. In addition, as with all forms of computing, 'bad data in equals bad data out'. Careful attention needs to be paid to which visualization strategies best accommodate the kinds of linguistic relationships that we want to explore. We risk creating a representation that resemiotizes our data in misleading ways.

The following sections present three visualization techniques that may be useful in resolving the tension between gaining a synoptic perspective on the text (the paradigmatic perspective) and capturing its unfolding (the syntagmatic perspective). The overview of these three techniques is intended as an invitation to the reader to think about how we might begin the task of exploring the emergent complexity of logogenesis.

Text Arcs: Visualizing Repetition

Text arcs are a technique for summarizing repetition in long strings. They have been used to visualize text, code (Wattenberg 2001), DNA (Spell & Brady 2003) and music (Wattenberg 2001). Text arcs are a development of the dotplot technique, a form of recurrence plot used in, for example, bioinformatics, to graphically compare repetition in genomic sequences (Figure 10.3). Dotplots represent repetitions in a similarity matrix by shading identical cells. The text arc layout, on the other hand, creates links between repeated units using translucent arcs (Figure 10.4) and is thus able to preserve a view of the time sequence.

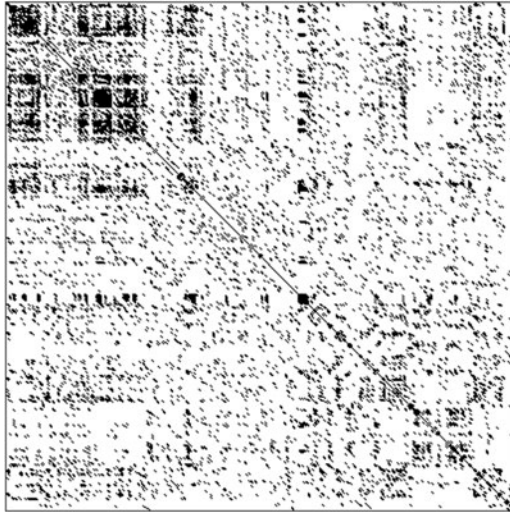


FIGURE 10.3 A DNA dotplot of a human zinc finger transcription factor (GenBank ID NM_002383), showing regional self-similarity

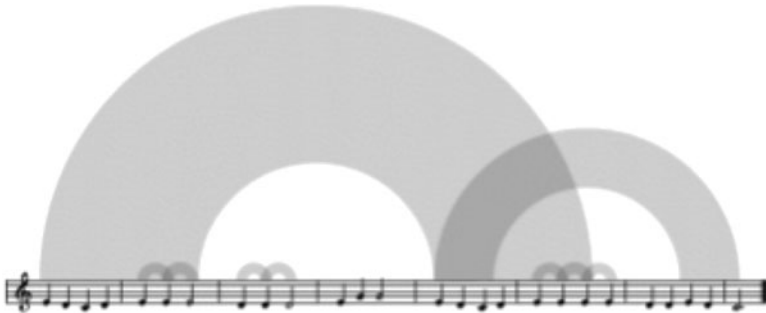


FIGURE 10.4 A text arc representation of music (Wattenberg 2002:5)

Translucency allows for differing levels of aggregation to be represented on the one diagram. The arcs overcome the problem of scalability, meaning that a long text sequence can fit onto a single page or screen with the time series represented along the horizontal axis. In essence, text arcs make a text tractable by providing a representation strategy that makes a time series manageable.

An example of the text arc technique used dynamically is ‘animated text arcs’ such as that developed by Byron (2007). Byron’s system dynamically renders text arcs while an audio-text unfolds to assist children with learning about rhyming in poetry (see an example at www.vimeo.com/734478):

A simplified text to speech engine is used to break down the poem into individual phonemes, so that ‘Once upon a time’ becomes ‘w-ah-n-s ax-p-aan-ey t-ay-m’ these phonemes can then be identified in patterns representative of alliteration, rhyme and rhythm. (Byron 2007)

The steps beneath the arcs represent rhythm, while the link repeated rhyme represents alliteration and homophones (Figure 10.5). The rhyming engine has also been used to create an interactive limerick writing assistance application with which a child can begin to type a line and be prompted with information about how many syllables remain to be used in that line. As you exhaust ‘remaining syllables the words become shorter, if you begin to type a word, words that begin with what letters you have typed so far are presented’ (Byron 2007).

Text arcs have also been used to ‘represent visually different types of multi-modal prosody so that a single text can be explored or comparisons can be made between different texts’ (Zappavigna & Caldwell 2008). Caldwell and

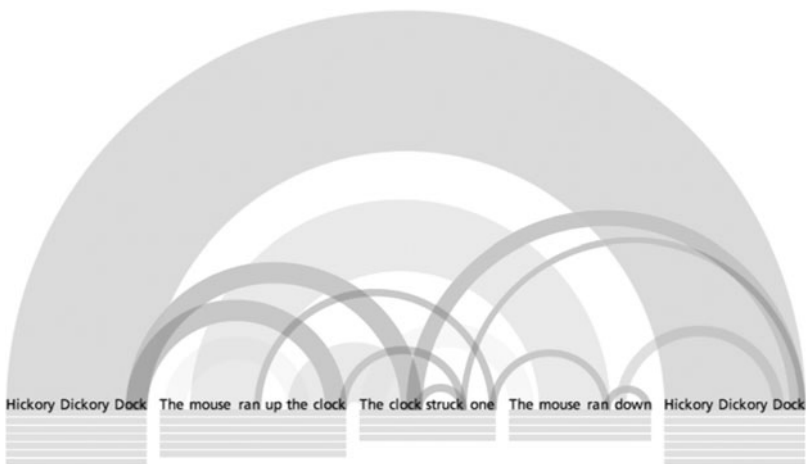


FIGURE 10.5 Dynamic text arc visualization of ‘Hickory Dickory Dock’ (Byron 2007)

Zappavigna (Chapter 11) explored how text arcs could be used in visualizing the patterning of end-rhymes in rap music. They also showed how end-rhymes unfold in popular rap music, providing a logogenetic view that allows the rhyming style of rap artists to be compared in terms of how they unfold with the text. In general, the text arc technique may be useful to discourse analysts investigating how repeated patterns differ across texts of the same or different genres.

Streamgraphs: Visualizing Multiple Data Series

Discourse analysts are usually interested in tracking the unfolding of more than one linguistic feature as it varies over a text or across a corpus. A visualization technique able to represent multiple features on the same diagram is the streamgraph. Streamgraphs are a form of stacked graph, a display where multiple data series are positioned one on the top of the other, offering a way of fulfilling this requirement. Streamgraphs visualize multiple variables as coloured ‘streams’ flowing with the time series on a single graph. Smooth curves are generated by interpolating between points to produce the ‘flowing’ river of data. The technique has been used to visualize box office revenues changing over time (Byron & Wattenberg 2008), changes in music listening habits (Byron 2008), shifts in lexical themes in corpora with time (Havre et al. 2002) and changes in word association in Twitter status messages (Clark 2008b). For example, Figure 10.6 is a streamgraph depicting a user’s ‘listening history’, which is the variation in artists that a user listens to over time. In this graph

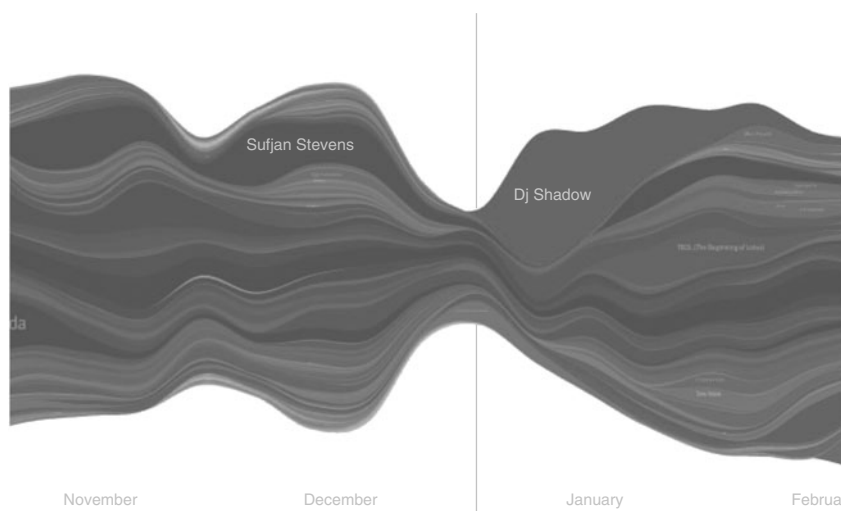


FIGURE 10.6 Extract from a streamgraph depicting a user’s ‘listening history’ (Byron 2008)

each layer or ‘stream’ represents a different artist and the width of the layer represents the frequency of the listening. Time is the movement from left to right over an 18 month span. The developer describes the graph as ‘a sort of virtual mirror, reflecting very personally significant events made visible by the changes in listening trends’ (Byron 2008). The colour scheme, represented here only in greyscale, was also used to indicate the level of interest a user had in each artist:

I ultimately decided on a color scheme that highlighted both the point of discovery of a musician as well as the user’s overall interest in them. Cool colors represent a core’ musician who the user is familiar with, while warmer colors represent a more recent discovery. The most saturated the color, the more interest the user has in that musician. (Byron 2008)

This kind of representation is potentially useful to linguists because it is a technique that allows multiple types of instances to be displayed as unfolding in time. The graph layout also provides a mechanism for representing information about the ‘weighting’ of those instances visualized as the width of the stream. Such weighting may be a simple frequency count of annotated items in a text, or based on a more complex metric in accord with a particular linguistic theory. If applied to an annotated text, the technique could be used to show how different types of linguistic phenomena co-vary over time. So, for example, we might export annotation series made in the video annotation software, ELAN, that have been encoded against a time series, and after some post-processing, visualize how modes such a gesture and facial expression are working together in the video.

An example of the streamgraph technique applied to textual data is *Themeriver* (Havre et al. 2002), a system that uses the river metaphor to visualize ‘themes’ varying over time in a collection of documents. ‘Themes’ are represented by colour-coded horizontal streams of varying width (Figure 10.7). Variance in width indicates the ‘strength’ of a theme, defined as the frequency of particular lexical items or the frequency of texts containing particular lexical items, depending on the customization selected. When the former frequency metric is adopted the system offers a topic-centered approach to visualizing a corpus in contrast to document-centered approaches (Berry 2003). *Themeriver* has been used to visualize the shifting of themes in a collection of texts by Fidel Castro from 1959 until 1961 (Figure 10.7). In this visualization particular points in time can be selected along the time series and annotated with labels (e.g. ‘Cuba and Soviet relations resume’ in Figure 10.7). This feature supports hypothesizing about context in particular domains, for example, political discourse, in various ways:

Providing such context allows users to evaluate content in relation to issues beyond those contained within the documents themselves. Continuing with

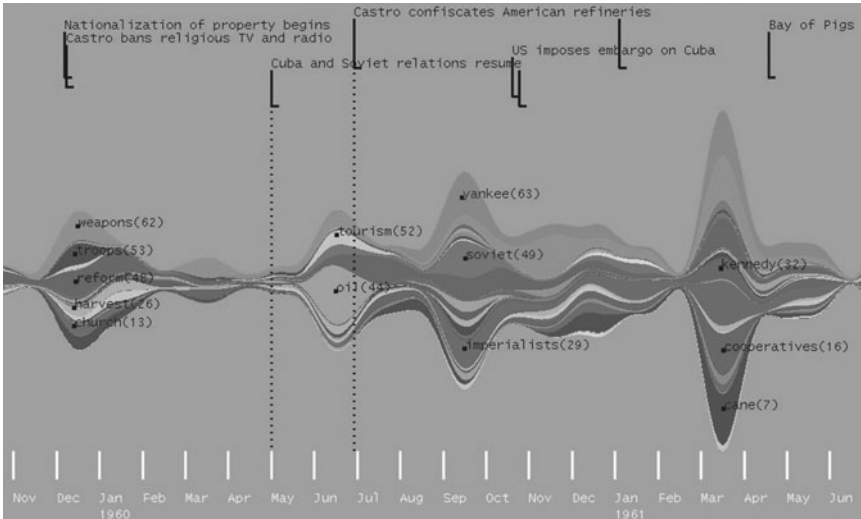


FIGURE 10.7 ‘Theme River uses a river metaphor to represent themes in a collection of Fidel Castro’s speeches, interviews and articles from the end of 1959 to mid-1961.’ (Havre et al. 2002:11)

the earlier example of candidates running for election, we might ask how the candidates’ themes change in response to news events. Do their speeches appear to trigger news events? Does a candidate’s opinion have any apparent impact on the stock market? (Havre et al. 2002:11)

However, as always, the old adage that ‘correlation does not equal cause’ needs to be kept in mind.

Streamgraphs have been used to visualize Twitter feeds (Clark 2008b). Twitter is a micro-blogging service that allows users to post status messages in text of up to 140 characters. Other users can subscribe to an individual’s twitter feed to receive these updates automatically. For example, Figure 10.8 shows a ‘Twitter Topic Stream’ for the top 100 twitter users (twits), which uses a variation of the Streamgraph technique to represent the distribution of the most ‘interesting’ capitalized words that occur in a database of twitter messages for the top 100 users. The developer employed a particular operationalization of ‘interestingness’:

The interestingness of a word was quantified by a function of the total references as well as the burstiness of the word distribution.

The most ‘interesting’ words in this data are primarily product, technology, or technology event names with the exceptions of ‘Scoble’ and ‘Obama’. This isn’t surprising since the top twitter users are early-adopters interested in technology. I was a bit surprised at the large volume for Seismic but

Animated Networks: The Text as ‘Becoming’

The metaphor often used in SFL of the text unfolding (logogenesis) invokes ideas about linear progression that might not be optimal for modelling a text’s complexity. An alternative metaphor that might be invoked is that of an animated network. This type of visualization seems more in accord with viewing the text as a complex adaptive system in which changes, particularly in initial conditions, have repercussions throughout the system. These types of systems are common in nature. An animated representation also invokes a metaphor of ‘becoming’ or propagation. Indeed it is through propagation that systems such as evaluative language swarm in a text, forming prosodic rather than constituent structures (Zappavigna et al. 2008). Fry’s (2000a:19) concept of ‘organic information visualization’ deploys related ideas, conceiving visualization as functioning to employ ‘simulated organic properties in an interactive, visually refined environment to glean qualitative facts from large bodies of quantitative data generated by dynamic information sources’. His system, Valence (Fry 2000b), will be reviewed in this section. A simplified example of Valence reading another of Mark Twain’s works, *The Innocents Abroad*, is available at www.benfry.com/valence (screen capture provided in Figure 10.10).

Valence (Fry 2000b) is a system that visualizes word usage as a network unfolding in a three-dimensional globe. The system renders words as ‘nodes’ in the network and connects words with branches if they are adjacent in the text so that ‘each time these words are found adjacent to each other, the connecting



FIGURE 10.10 A simplified version of Valence reading ‘The Innocents Abroad’ by Mark Twain (Fry 2000b)

line shortens, pulling the two words closer together in space' (Fry 2000a:67). An important aspect of the value of the system is this foregrounding of the relationality of language:

The premise is that the best way to understand a large body of information . . . is to provide a feel for general trends and anomalies in the data, by providing a qualitative slice into how the information is structured. The most important information comes from providing context and setting up the interrelationships between elements of the data. If needed, one can later dig deeper to find out specifics, or further tweak the system to look at other types of parameters. (Fry 2000b)

While the **system only models one kind of relationship, lexical adjacency**, a logical extension appears to replace the input data, currently 'raw text' (Figure 10.11), with annotated data and to specify different kinds of relationships between annotation series. This would occur at the 'preprocessor engine' stage of the information pipeline that Fry proposes as a software engineering method (Figure 10.11).

Valence 'reads' the text by moving words that are used most frequently to the edges of the globe and less frequent words to its centre (Figure 10.12). Within the system, logogenesis is represented as a proximal–distal relationship rather than movement from left to right across a page. The text 'unfolds' by moving the current lexical item being 'read' to the centre front of the three-dimensional space. In some versions of the system a small page is shown next to the network with lines of the text appearing in sync with the 'reading' provided by the movement of the network. A Quicktime video of Valence in reading Mark Twain's *The Innocents Abroad* is available at www.benfry.com/valence/movie.html.

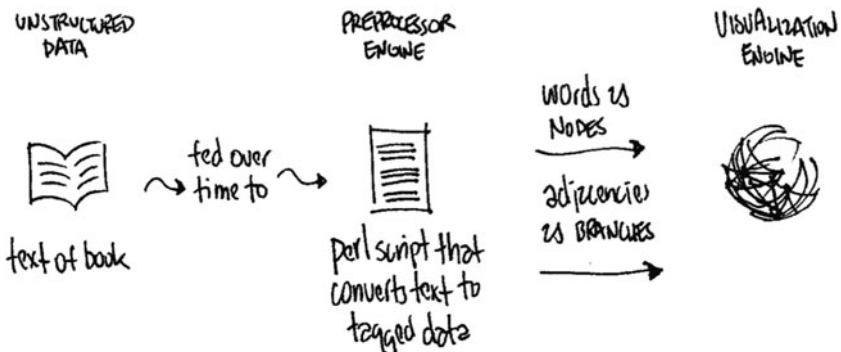


FIGURE 10.11 The 'information pipeline', a software engineering method for the Valence visualization (Fry 2000a:65)

The dynamic network representation is an attempt to overcome the problem of how to make tractable hefty data sources such as texts that contain large quantities of unique elements. As Fry explains:

A bar chart containing this many elements would be nearly worthless. It would be too large to take in at a glance, or if shrunk to one's field of view, too small to understand. A focus + context technique like the *Table Lens* could be used, but due to enormous disparities in word usage . . . less than 25% [in the case of the text 'The Innocents Abroad' by Mark Twain] would be worthwhile at all, with the interesting features not even appearing until the top 5%. (Fry 2000a:66)

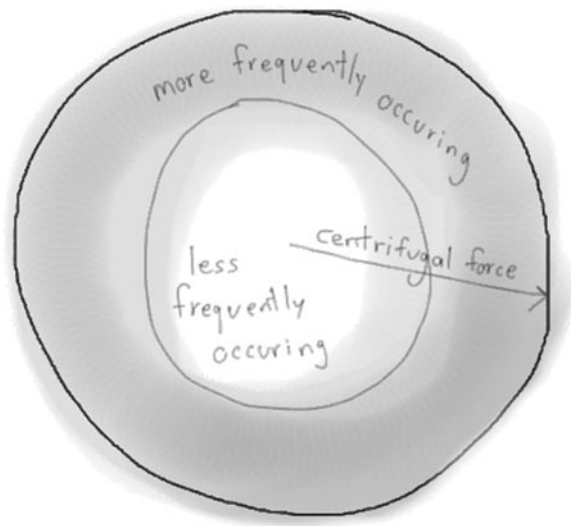


FIGURE 10.12 Metaphors of space used in Valence

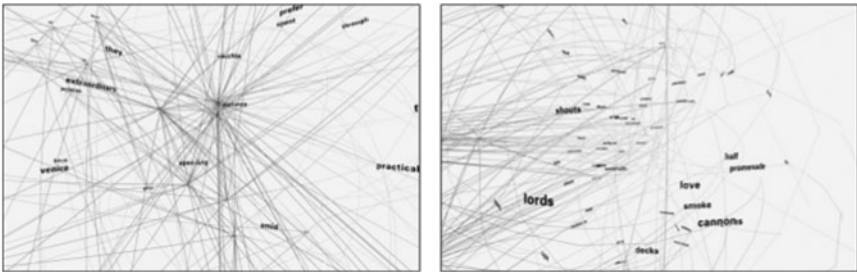


FIGURE 10.13 Two viewpoints on a network (Fry 2000a:68)

The three-dimensional visualization affords a way for the user to move around inside the text and explore relationships between words. The user is able to zoom in or view the network from different viewpoints (Figure 10.13), depending on the relationships that they wish to investigate.

Conclusion

This chapter has presented three text visualization techniques that use particular representation strategies for making logogenesis both visible and tractable: Text Arcs, Streamgraphs and animated networks. The first technique is useful for discourse analysts exploring repeated patterns in texts, the second for representing the unfolding of more than one linguistic feature on the same graph, and the third for achieving a dynamic representation of features unfolding in time. The techniques are examples of moving beyond a ‘bag of entities’ perspective on texts to embrace the complex sequencing of discourse. If we are able to develop these techniques to cope with annotated systemic functional input then we will have a powerful lens on our data. We will also have a useful mechanism for communicating analyses of patterns that will allow us, in turn, to develop functional theory about discourse patterning without factoring out time² (Zhao 2009, forthcoming).

Effective annotation is the first step in visualization of features that cannot be automatically extracted from text with current computational techniques. This means that we require systems that support easy manual annotation of texts by the linguist. Examples of existing text annotation systems developed by Systemic Functional linguists include Systemics (Judd & O’Halloran 2001), UAM Corpus Tool (O’Donnell 2008) and SysAM (Matthiessen & Wu 2001). To date, there has been no work done on how the output of these systems might be visualized. We might think of ourselves as biologists trying to map the genome without a theory of sequencing.

Acknowledgements

I would like to acknowledge the support of the Australian Research Council in funding this research.

Notes

¹ These categories are taken from Appraisal theory (Martin & White 2005) and refer respectively to language about emotional responses and language scaling evaluation in a text.

² By time, I do not refer to physical time but instead to 'text time' in the sense of logogenetic unfolding.

References

- Berry, M. (2003). *Survey of text mining: Clustering, classification, and retrieval*. New York: Springer.
- Byron, L. (2007). *Children's poetry and limerick visualizations*. Retrieved 11 August 2008, from Lee Byron: www.leebyron.com/what/poetry/.
- Byron, L. (2008). *Last.fm listening history – What have I been listening to?* Retrieved 31 July 2008, from Lee Byron : www.leebyron.com/what/lastfm/.
- Byron, L. & Wattenberg, M. (2008). *Stacked graphs – Geometry & aesthetics*. Retrieved 8 July 2008, from Lee Byron: www.leebyron.com/else/streamgraph/.
- Clark, J. (2008a). *Tom Sawyer character streamgraph*. Retrieved 11 August 2008, from Neoformix: Discovering and illustrating patterns in data: www.neoformix.com/2008/TomSawyer.html.
- Clark, J. (2008b). *Twitter topic stream*. Retrieved 31 July 2008, from Neoformix: Discovering and illustrating patterns in data: www.neoformix.com/2008/TwitterTopicStream.html.
- Clark, J. (2008c). *Twitter topic streams for some top users*. Retrieved 11 August 2008, from Neoformix: Discovering and illustrating patterns in data: www.neoformix.com/2008/TwitterTopicStreamsTopUsers.html.
- Fry, B. (2000a). Organic Information Design. Unpublished dissertation. Boston, MA: Massachusetts Institute of Technology.
- Fry, B. (2000b). *Valence*. Retrieved 18 July 2008, from Ben Fry: www.benfry.com/valence/applet/.
- Halliday, M.A.K. (1991). Towards probabilistic interpretations. In E. Ventola (Ed.), *Functional and systemic linguistics: Approaches and uses* (pp. 39–61). Berlin and New York: Walter de Gruyter.
- Halliday, M.A.K. (1993). Language in a Changing World. Occasional Paper Number 13. Toowoomba, Queensland: Applied Linguistics Association of Australia, Centre for Language Learning and Teaching, University of Southern Queensland.
- Halliday, M.A.K. & Martin, J.R. (1993). *Writing science: Literacy and discursive power*. London: Routledge, Taylor & Francis Group.
- Halliday, M.A.K. & Matthiessen, C. (2004). *An Introduction to Functional Grammar*. London: Edward Arnold.
- Havre, S., Hetzler, E., Whitney, P. & Nowell, L. (2002). ThemeRiver: Visualizing thematic changes in large document collections. *IEEE Transactions on visualisation and Computer Graphics*, 8(1), 9–20.
- Judd, K. & O'Halloran, K. (2001). *Systemics*. Singapore: Singapore University Press 2001. (Educational software).
- Knight, N. (2008). 'Still cool . . . and american too!': An SFL analysis of deferred bonds in internet messaging humour. In N. Nørgaard (Ed.), *Systemic Functional Linguistics in Use, Odense Working Papers in Language and Communication* (Vol. 29) (pp. 481–502). Odense: University of Southern Denmark, Institute of Language and Communication.

- Lemke, J.L. (1991). Text production and dynamic text semantics. In E. Ventola (Ed.), *Functional and Systemic Linguistics: Approaches and Uses 23* (pp. 23–38). Berlin and New York: Mouton de Gruyter.
- Martin, J.R. (2000). Beyond exchange: Appraisal systems in English. In J. Martin, S. Hunston & G. Thompson (Eds), *Evaluation in text: Authorial stance and the construction of discourse* (pp. 142–175). Oxford: Oxford University Press.
- Martin, J.R. (2004). Mourning: How we get aligned. *Discourse and Society* 15, (2–3), 321–344.
- Martin, J.R. (2008, July 21–25). Chaser's war on context: Making meaning. *Paper presented at the 35th International Systemic Functional Congress*. Sydney.
- Martin, J.R. & Rose, D. (2003). *Working with discourse: Meaning beyond the clause*. London, New York: Continuum.
- Martin, J.R. & White, P.R.R. (2005). *The language of evaluation: Appraisal in English*. New York: Palgrave Macmillan.
- Matthiessen, C. (2007). The 'architecture' of language according to systemic functional theory: Developments since the 1970s. In R. Hasan, C. Matthiessen & J. Webster (Eds), *Continuing discourse on language: A functional perspective (Volume two)*. London: Equinox.
- Matthiessen, M.I.M. & Wu, C. (2001). SysAm. [Programs for computational Analysis] Available at: www.iminerva.ling.mq.edu.au.
- O'Donnell, M. (2008). Demonstration of the UAM CorpusTool for text and image annotation. *Proceedings of the ACL-08:HLT Demo Session (Companion volume)* (pp. 13–16). Columbus, OH: Association for Computational Linguistics.
- Spell, R. & Brady, R. (2003). BARD: A visualization tool for biological sequence analysis. *Proceedings of the IEEE Symposium on Information Visualization*. Seattle, WA: IEEE.
- Wattenberg, M. (2001). *The shape of song*. Retrieved 31 July 2008, from Turbulence: www.turbulence.org/Works/song.
- Wattenberg, M. (2002). Arc diagrams: Visualizing structure in strings. *Proceedings of the IEEE Symposium on Information Visualization* (pp. 110–116). Boston, MA.
- Wise, J., Thomas, J., Pennock, K., Lantrip, D., Pottier, M., Schur, A., et al. (1995). Visualizing the non-visual: Spatial analysis and interaction with information from text documents. *Proceedings of the IEEE Information Visualization Symposium* (pp. 51–58). Atlanta, GA.
- Zappavigna, M. & Caldwell, D. (2008). Visualising multimodal patterning. Paper presented at Critical Dimensions in Applied Linguistics, July 4–6. Sydney.
- Zappavigna, M., Dwyer, P. & Martin, J. (2008). Syndromes of meaning: Exploring patterned coupling in a NSW Youth Justice Conference. In A. Mahboob & K. Knight (Eds), *Questioning linguistics* (pp. 103–117). Newcastle: Cambridge Scholars Publishing.
- Zhao, S. (2010). Intersemiotic relations as logogenetic patterns: Towards the restoration of the time dimension in hypertext description. In M. Bednarek & J. Martin (Eds), *New discourse on language: Functional perspectives on multimodality, identity, and affiliation* (pp. 195–218). London: Continuum.

Chapter 11

Visualizing Multimodal Patterning

David Caldwell
University of Sydney

Michele Zappavigna
University of Sydney

Background: Text Visualization

Text Visualization is an emergent field closely related to the more general field of information visualization, a field that represents abstract data visually. The objective is to computationally process a text so that it can be represented in ways that leverage the ‘primarily preattentive, parallel processing powers of visual perception’ (Wise et al. 1999:442). In short, visualization aims to make complex data that is encoded by machines meaningful to humans, using tools such as colour, space and animation to produce visual representations.

Text visualization is especially useful for discourse analysts, and linguists more generally, as they are interested in making claims about text patterns. Such patterning is often highly complex, involving different types of linguistic features, depending upon the linguistic theory deployed. For example, patterning of interpersonal meaning has been analogized with musical patterning:

These structures can be likened to the harmonic progressions in a piece of music, which have a distinctive quality in themselves but also enter in relationship with other ‘chord progressions’ in the piece and contribute to the interpersonal structure of the text as a whole. (Macken-Horarik 2003:314)

Because of the high dimensionality of language, many such patterns are not necessarily directly evident through close analysis of individual texts, especially in the case of extended texts or corpora. Ware (2004) notes a number of important advantages afforded by visualization that may assist the linguist in exploring large data sources:

- Understanding of large amounts of data;
- Perception of emergent properties of the data that were not anticipated;
- Detection of errors in the data that otherwise remain hidden;
- Understanding of both local and global features.

Reconstruing ‘data’ as ‘text’ in the above, we may think of visualization as a tool to assist the linguist in exploring text patterns and explaining them to others.

This chapter introduces arc diagrams, a visualization technique that represents repeated patterns in text. We begin by explaining the technique and its application to music and children’s poetry. We then introduce the case study: end-rhymes in a Kanye West rap song. The technique of arc diagrams is applied West’s rap song and the findings are discussed. We conclude by considering how arc diagrams may assist the linguist in making claims about the virtuosity of rap artists and multimodal patterning more generally.

Arc Diagrams: Visualizing Repetition

Methods for representing sequencing in strings have become particularly important with developments for understanding gene sequencing in bioinformatics. Arc diagrams are a novel technique that visualizes repetition. An arc diagram represents repetition in text strings ‘by using a pattern-matching algorithm to find repeated substrings, and then representing them visually as translucent arcs’ (Watternberg 2002:2). For example, the translucent arcs in Figure 11.1 indicate that the string ‘1234’ is repeated three times. The wider the shading of the arc, the longer the sequence that is repeated in a pattern of patterns.

Arc diagrams were initially developed for visualizing music in a project called *The Shape of Song* (Watternberg 2002). For example, Figure 11.2 shows an arc



FIGURE 11.1 Arc diagram visualization [adapted from Watternberg (2002:2)]

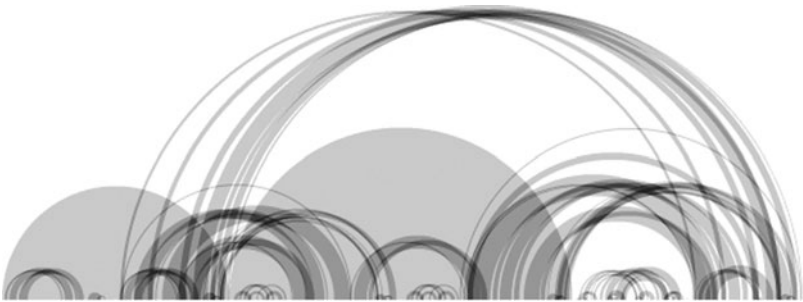
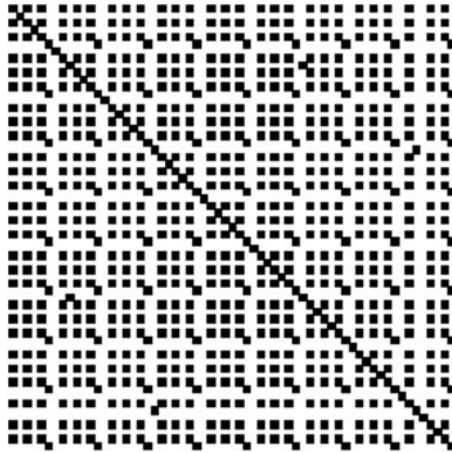


FIGURE 11.2 Arc diagram of Chopin’s *Mazurka in F# Minor* (Watternberg 2002)



FIGURE 11.3 Arc diagram of the folk song *Clementine* (Watternberg 2002)



(a)



(b)

FIGURE 11.4 Comparison of a dotplot and arc diagram for the same string (Watternberg 2002:3)

diagram of Chopin's *Mazurka in F# Minor* produced by the Watternberg project. The elaborate nesting of this piece is very different to the simple repetition of the refrain in the folk song *Clementine* (Figure 11.3).

A technique used in bioinformatics for visualizing repetition in sequential data is the dotplot. This has been employed, for example, to compare genes (left image, Figure 11.4). Repetition in this form of representation is shown through shading identical cells. A diagonal line occurs when there is a common subsequence, that is, a series of items that occur in both sequences. Arc

diagrams may be thought of as an improvement on this technique depending on the kind of conclusions that the analyst is trying to make with the diagram. Wattenberg demonstrates how the arc diagram technique reveals repetitions in substrings that are difficult to appreciate in the ‘visual clutter’ of a dotplot (Wattenberg 2002:3). As Figure 11.4 shows, the fact that two substrings in the sequence are repeated only once is difficult to ascertain from the dotplot but is clearly apparent in the arc diagram.

Arc diagrams offer a synoptic view of the text, that is, a view of the text as a static entity, by making the text tractable on a single page or screen, depending on the level of aggregation selected. However, they simultaneously offer a view of the text unfolding as a sequence in a way that a table of statistics, for example, would not. Providing both these perspectives is an important aim of linguistically motivated text visualization (Zappavigna 2007). Following Martin (2004), it is an attempt to avoid submerging texture when trying to create an overview of the text. By texture, Martin (2004) is referring to the linguistic patterns that are construed dynamically in a text, for example, a generic staging that works towards achieving a particular social purpose.

While the arc diagrams produced in this chapter were drawn manually to exemplify the technique on a small data set, arc diagrams can be produced automatically using a patterning matching algorithm. This algorithm may define two items as matching, based on criteria specified in the algorithm. While this restricts the criteria to features that may be automatically detected in text, more complex features could be employed using an annotated text. This would enable linguists to represent features that are pertinent to their particular projects.

A recent application of arc diagrams of direct relevance to our study is Byron’s (2007) animated arc diagrams. **Animated arc diagrams were developed to assist children in learning about rhythm, repetition and rhyming. Byron’s (2007) system dynamically renders arc diagrams while an audio-text unfolds** (see an example at www.vimeo.com/734478). The system uses a ‘simplified text to speech engine to break down the poem into individual phonemes, so that “Once upon a time” becomes “w-ah-n-s ax-p-aa-n ey t-ay-m”’ (Byron 2007). Once the phonemes are identified patterns of rhyme, rhythm and alliteration can then be shown visually using arcs to link repeated units (Figure 11.5). Rhythm is represented beneath the arcs by grey horizontal bars.

This animated arc system was developed into an interactive limerick writing assistance application. A rhyming engine was used to create an application in which a child would begin to type a line and be prompted with information about how many syllables remain to be used in that line. As one exhausts ‘remaining syllables, the words become shorter, if you begin to type a word, words that begin with what letters you have typed so far are presented’ (Byron 2007). While the authors in this chapter do not have access to the code used to develop this system, the same interactive system could be used to enable young people to explore rhyming and repetition in rap music.

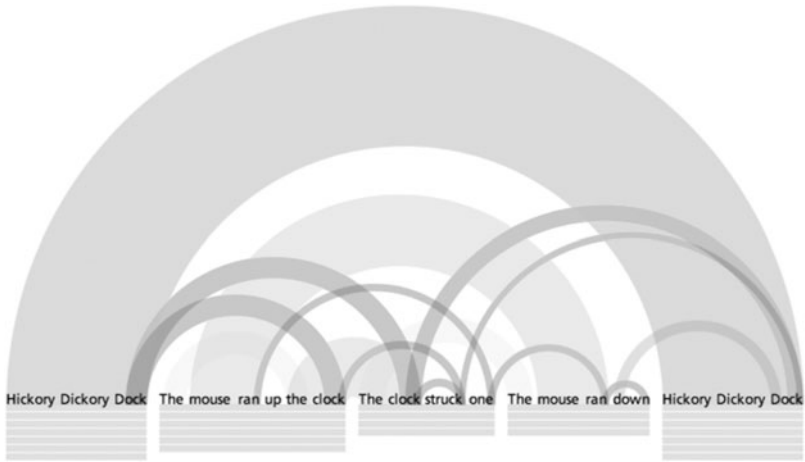


FIGURE 11.5 Dynamic arc diagram visualization of ‘Hickory Dickory Dock’ (Byron 2007)

Irrespective of whether data is analysed manually (Watternberg 2002), or automatically animated (Byron 2007), arc diagrams provide a means by which to visually represent linguistic patterns from both a synoptic and dynamic perspective. The following analysis will apply Watternberg’s arc diagrams with an aim to visually capture the progression and extent of end-rhymes in Kanye West’s rap music. The ultimate aim is to capture the extent to which a rapper consecutively repeats the same end-rhyme.

Data: Rap Music, Rhyme and Kanye West

The data for this chapter is from the contemporary, ‘popular’ North American rap musician Kanye West. The song chosen for analysis is titled *Spaceship* from West’s inaugural album: *The College Dropout* (2004). The lyrics (which for copyright reasons are only reproduced here as individual rhymes) have been accessed online from *The Original Hip-Hop Lyrics Archive* (www.ohhla.com).

Drawing inspiration from Wattenberg (2002) and Byron (2007) and their visualizations of music, we considered rap music an attractive source of data. Generally speaking, the vocal performance of ‘rapping’ requires a performer to match the rhythm of their voice to the beat of music, and this is often unrehearsed and spontaneous. In addition, rapping is articulated in poetic form so it involves rhyme, as well as African-American language practices such as narrativizing, toasting and punning (Richardson 2006:11).

Rap is about virtuosity. It is a means by which one can establish a reputation within the hip-hop community. And in most cases, rap artists are explicitly

judged by that community in terms of their capacity to synchronize their vocals to the beat of the music, as well select lyrics that appeal for both sound and sense. While these are just some of the more general markers of virtuosity, they are integral to the way in which people compare the skills of different rappers. We have chosen to focus exclusively on the repetition of rhyme as a marker of virtuosity. According to rap expert Keyes, 'the ideal rendering of lyrics must be grounded in poetic flow . . . Effective rhyming in rap, as with most poetic forms, requires selecting words for both sound and sense' (Keyes 2002:126–127). There are of course many varied ways in which rappers deploy rhyme in rap music and these are reviewed extensively by Alim (2003). An analysis of every type of rhyme in a set of rap lyrics is not only beyond the scope of this chapter, but would most probably be very difficult to effectively visualize with arc diagrams. Therefore, we have limited our analysis to end-rhymes (see Alim 2003:70).

End-rhymes, in contrast with internal rhymes, occur at the end of a clause, and generally carry the major pitch movement. Internal rhymes do not occur at the end of the clause and do not necessarily have any notable pitch movement. Following the 'rhyme tactics' described by Alim (2003:63), we include *any* type of end-rhyme in our analysis, such as masculine rhymes (one rhyming syllable, for example, 'chain', 'fame' 'game') and feminine rhymes (two rhyming syllables, for example, 'musc-le', 'russ-ell', 'hust-le'), as well as assonance and basic repetition. As illustrated, when coding for rhyme, we will only highlight (using **bold** and underline) the vowel phoneme(s), which we will simply refer to as the rhyming 'sound'. Finally, we will only code a text arc for an end-rhyme that is *consecutive* and the *same* sound. We will discuss the exclusion of strings of *non-consecutive*, *differing* sounds when we introduce the visualizations.

So why Kanye West? We chose West mainly because he is *not* renowned for having the best rapping skills, despite his immense commercial success. In fact, West is much more renowned within the hip-hop community for his editing, sampling and production skills. As co-producer of his rap songs, West will often collaborate with other rap artists. And generally, those artists are more highly skilled, 'well-credentialed' rappers. Accordingly, we are interested in how West's rhyming skills compare with such collaborators, in this case, the rappers GLC and Consequence in the rap song *Spaceship* (West 2004).

Method: Rhyme and Graduation

From a more theoretical perspective, we are also interested in relating rhyme, particularly the kind of consecutive rhyming in rap lyrics, to the Appraisal system of GRADUATION (Martin & White 2005). Appraisal theory, from Systemic Functional Linguistics, is an analytical framework designed to identify interpersonal meanings in language. With respect to the three main Appraisal

systems, ATTITUDE concerns the semantic resources used to negotiate emotions, judgements and valuations, while GRADUATION and ENGAGEMENT concern the resources that amplify and engage with ATTITUDE. We are focused exclusively on GRADUATION and the extent to which it relates to the kinds of consecutive rhyming in rap music. The system of GRADUATION covers a wide range of linguistic resources, all of which are used to grade a speaker's evaluations (see Martin & White 2005:154 for a complete system network). In short, GRADUATION comprises two main systems: FORCE and FOCUS, where FORCE is 'grading according to intensity or amount' and FOCUS is 'grading according to prototypicality and the preciseness by which category boundaries are drawn' (Martin & White 2005:137). Consecutive rhyme is understood here as a kind of FORCE, or more specifically, it is classified as repetition; a sub-type of FORCE: INTENSIFICATION (see Figure 11.6).

However, when we think about rhyme, its semiotic 'force' or INTENSIFICATION (Intensification from here on) is much more than simply lexical repetition. At this point, it is worth acknowledging the growing interest in the application of Appraisal to other modes of meanings:

... work on paralanguage (gesture, facial expression, laughter, voice quality, loudness etc.) and attendant modalities of communication (image, music, movement etc.) are central arenas for further research on the realization of attitude [and graduation] as we move from a functional linguistic to a more encompassing social semiotic perspective. (Martin & White 2005:69)

We could argue then that repetition (as Intensification) is much more like 'paralinguistic' repetition than discourse semantic repetition. First, while the 'sense' or meaning of consecutive rhyming might have some kind of semantic thread between the particular lexemes (see examples in Figure 11.6), this is not

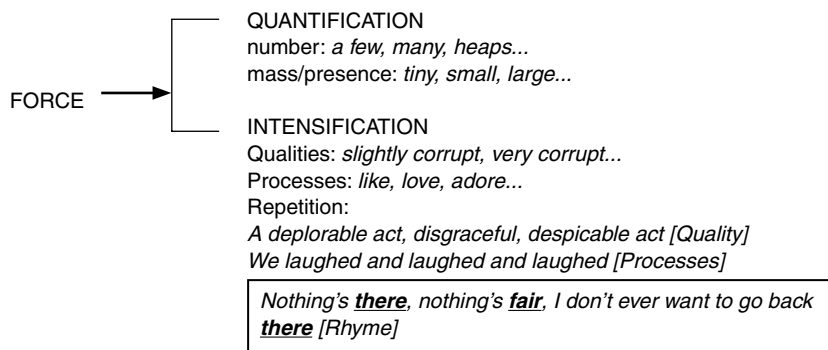


FIGURE 11.6 Some examples of GRADUATION: FORCE, including Repetition (after Martin & White 2005:154)

necessarily the case. Moreover, we would argue that it is the ‘sound’, or ‘sensory force’ of consecutive rhymes, particularly of the same sound, that signifies Intensification or ‘force’. In a way, it can be likened to a gradual increase in loudness (a crescendo), albeit realized through the repetition of sounds that do not necessarily increase in amplitude. In musical terms, a crescendo is a passage of music that gradually increases in force or loudness. So with respect to the system of GRADUATION (Figure 11.6), we include consecutive rhyming (of the same sounds) as part of the Intensification system, and in particular, the sub-system of repetition. However, we do note that this is not the same as repetition of the discourse semantic kind; it is better classified as a kind of paralinguistic or ‘sensory’ intensification (for want of a better term).

Analysis: Arc Diagrams and ‘Virtuosity’

The following set of arc diagrams aim to visualize the ‘virtuosity’ of a rapper in terms of their capacity to produce consecutive end-rhymes using syllables of the same sound. As mentioned, there are three data sets: Kanye West, GLC and Consequence, all of which have been taken from the same song: *Spaceship* (West 2004). A basic generic structure of *Spaceship* is outlined in Figure 11.7.

We have limited our analysis to the three verses of *Spaceship*. Each verse varies slightly in size: West’s verse comprises 41 clauses, GLC’s comprises 50; and Consequence’s comprises 25.

Before we introduce the analysis of the data, it is important to explain the way in which we have used the texts arcs to represent the build-up of consecutive end-rhymes. Figure 11.8 is one segment of analysis of the West data set.

The horizontal axis represents time, or more technically, logogenesis; the text as it unfolds ‘in the world’. Each horizontal axis is segmented into smaller components with a single, vertical line. Each of these segments represents a single line of text, basically equivalent to a clause, tone group and poetic ‘line’.

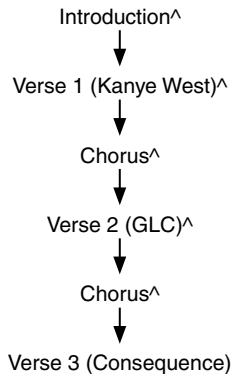


FIGURE 11.7 Generic Structure of *Spaceship* (West 2004)

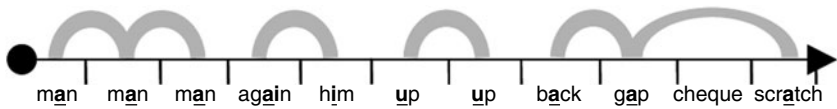


FIGURE 11.8 Arc diagram of consecutive end-rhymes from Kanye West in *Spaceship* (West 2004)

Below the horizontal axis, and within each of these segments, is the end-rhyme. While it would be more accurate to place the end-rhyme to the far right-hand side of each segment, there is simply not enough space. A single, non-translucent arc is used to represent repetition; in this case, a consecutive end-rhyme of the same sound, for example, ‘again’/‘him’ and ‘up’/‘up’ (see Figure 11.8). However, we do not code for any end-rhymes that are not consecutive, or that are not the same sound (with some exceptions discussed below). So, for example, the end-rhymes ‘again’/‘him’ and ‘up’/‘up’ only constitute two arcs. In other words, there is no arc between ‘him’ and the initial ‘up’. Unlike Wattenberg (2002), we are only using non-translucent, lower level arcs because the small quantity of instances does not really offer us the potential to show patterns of patterns, just single patterns.

On several occasions, we do ignore some ‘non-consecutive’ rhymes so as to visualize a lengthy, consecutive end-rhyme coding. As illustrated above, ‘back’, ‘gap’, ‘cheque’ and ‘scratch’ are all coded as part of a consecutive string,

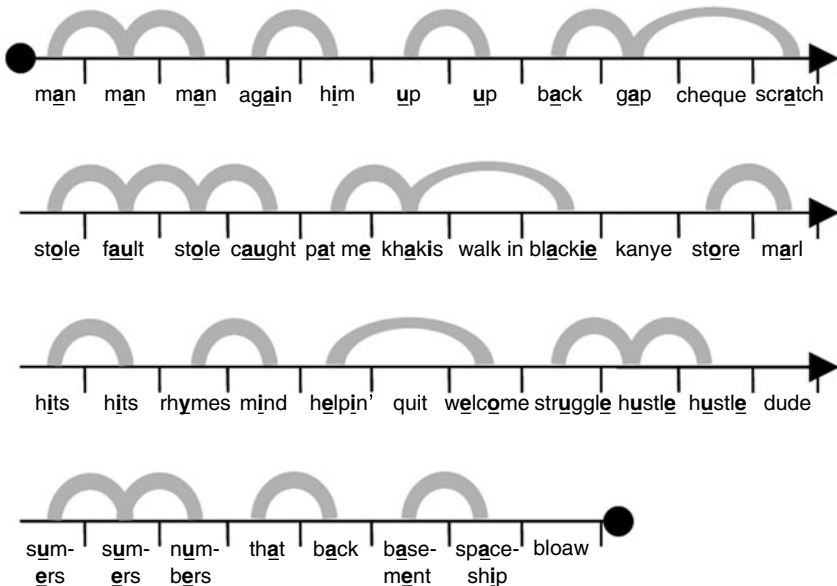


FIGURE 11.9 Arc diagrams of end-rhymes from West in *Spaceship* (West 2004)

indicating a wave of repetition beyond a basic rhyming couplet. In this case though, 'cheque' does not actually rhyme with either 'gap' or 'scratch'. The smaller, non-translucent arc extends from 'gap' to 'scratch', missing 'cheque' altogether. In cases where a rapper deviates from consecutive rhyming for maybe one, two or three end-rhymes, we still code it as a larger segment of repetition. The fact that the rapper quickly returns to their initial rhyming sound is usually not coincidental. And importantly, in those cases, the 'force' or sonic Intensification is still maintained, despite the momentary lapse in rhyming repetition.

The three sets of text arcs begin with West, then GLC and Consequence (see Figures 11.9, 11.10 and 11.11).

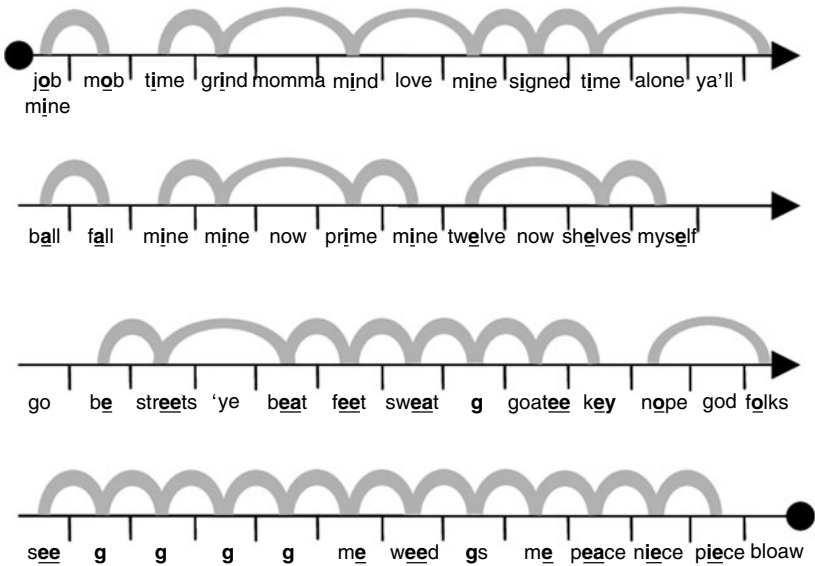


FIGURE 11.10 Arc diagrams of end-rhymes from GLC in *Spaceship* (West 2004)

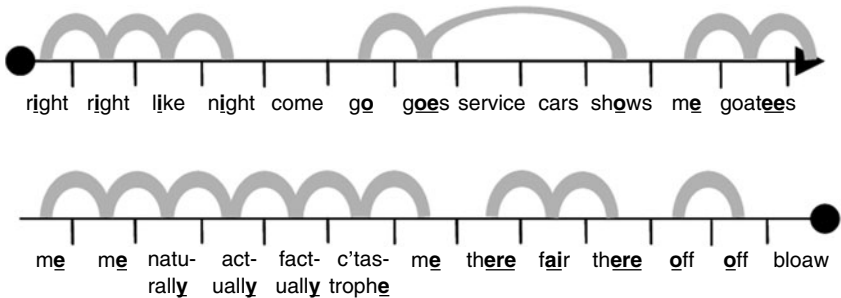


FIGURE 11.11 Arc diagrams of end-rhymes from Consequence in *Spaceship* (West 2004)

We will explain these arc diagrams using an aggregated view in the following section.

Summary of Findings

Before we compare the rhyming capacity of the three rappers, it is worth noting that there are very few instances where the rappers do *not* rhyme. While there are clear differences in terms of the extent to which the rappers do or do not rhyme consecutive syllables of the same sound, all three rappers clearly display some kind of virtuosity in terms of their capacity to rhyme consistently. And while that may not constitute or evoke the same kind of semiotic ‘force’ as consecutive syllables of the same sound, it does however, at the very least, demonstrate a capacity to rap. It would be worth comparing these arc diagrams to other rap songs in which the artists were *not* studio recorded. In ‘freestyle’ rapping, for example, these kinds of findings would show an even greater level of virtuosity given that the performance is spontaneous, and does not afford the luxury of rehearsal.

Most significantly though, the arc diagrams reveal a clear difference between West’s rhymes and those of his two collaborators. Figure 11.12 is an aggregated text arc view, comparing the consecutive rhyming of the three artists.

This aggregated arc diagram view shows that Consequence and particularly GLC rhyme consecutive end-rhymes of the same sound to a much greater extent than West. So what does this then say about West’s virtuosity as a rapper? Quite simply, we could argue that West’s status as an ‘average’ rapper is somewhat justified, at least in terms of his use of rhyme as a means of Intensification.

In addition, it is important to note that the really significant build-up of Intensification through end-rhymes occurs mainly at the end of both GLC’s and Consequence’s verse. This appears to be a very deliberate tactic, especially when we consider that all three rappers finish their verse with ‘blow’, a reference to the metaphorical spaceship ‘taking off’. One could certainly argue that

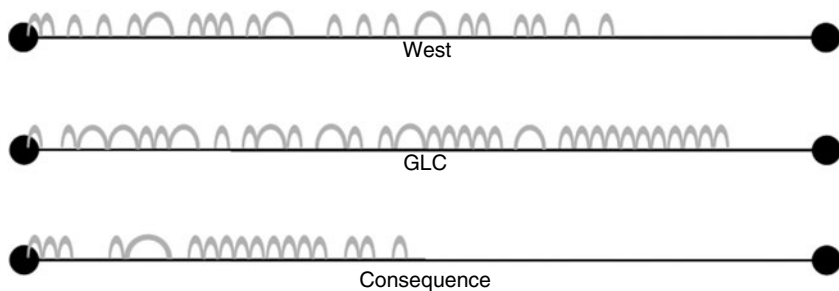


FIGURE 11.12 Aggregated text arc view from *Spaceship* (West 2004)

both GLC and Consequence are very aware of the ‘graduating’ function of consecutive end-rhymes and deploy them accordingly. Moreover, this finding suggests that consecutive rhyming of the same sound does function in a similar way to a musical crescendo.

From a different perspective though, it could be argued that by avoiding consecutive end-rhymes of the same sound, West is actually able to express many more detailed and elaborate meanings in his lyrics given that he is not continually limited by having to find appropriate lexis that matches a particular sound. Compare, for example, West’s rhyming couplets with the consecutive end-rhymes from GLC shown in Table 11.1.

The point here, albeit difficult to recognize without a complete clause, is that GLC’s verse is not as semantically ‘rich’ when compared with West’s. Or perhaps, more technically, it lacks the same level of semantic or ‘ideational’ coherence (see Martin & Rose 2003). In the extract above, GLC lists people he hopes to ‘see’, for example, ‘freddy **g**’ and ‘yousef **g**’. He then explains, without any obvious semantic link, that police watch him (‘**mē**’) smoking marijuana (‘**wēed**’). It is only in the final four clauses where GLC’s rhymes have some kind of semantic coherence. In those clauses he is self-reflective: he recognizes that he has people counting on him (‘**mē**’), that he is trying to find ‘**peace**’, and that, somewhat related, he should have finished school like his ‘**niece**’ instead of using a ‘**piece**’ (a gun).

In contrast, West’s lyrics are much more semantically coherent as they clearly relate to the ‘macro’ theme of the song. *Spaceship* is basically about leaving one’s

Table 11.1 Comparing Rhyme Tactics:
GLC and West in *Spaceship* (West 2004)

GLC:	West:
see	hits/
g	hits
g	rhymes/
g	mind
g	helping/
me	welcome
weed	struggle/
g’s	hustle/
me	hustle
peace	
niece	
piece	

ordinary circumstances and ‘taking off’ to a better place, hence the metaphorical ‘spaceship’ (see Smitherman 2006:99). West’s rhyming couplets provide a really clever juxtaposition of his adverse circumstances and his tenacity. For example, West contrasts the fact that he receives ‘**hits**’, that is, punches (metaphorical or not), but at the same time, writes ‘**hits**’, that is, successful song lyrics. Or, despite his job not ‘**helping**’, he quits with a departing phrase, ‘you’re **welcome**’. And in the final three rhymes West is even more explicit, where he claims that no one knows his ‘**struggles**’, but at the same time, they can’t match his ‘**hustle**’, that is, his tenacity.

Perhaps this hypothesis, which would obviously benefit from an analysis of more data, is best explained in terms of ‘sound versus sense’. When the ‘sound’ or sonic intensification of the consecutive rhyme is foregrounded, as it is here with GLC, then the artist must compromise their lyrical meaning potential. If, however, the artist, like West, foregrounds their lyrical meaning potential, then it is more likely that the artist cannot foreground the sonic force, in this case, through consecutive end-rhymes of the same sound.

Conclusion

At times, in the history of linguistics one ideology of empiricism or another has tended to privilege generalizations across groups of texts over close readings of single ones. It may be that the rise of corpus linguistics heralds a new phase of generalizing privilege of this order. If so, as social discourse analysts, we need to guard against studies that submerge unfolding texture in processes of counting and averaging that look for trends across texts rather than contingencies within them. (Martin 2004:341–342)

This chapter has applied only one type of text visualization to a very small and unique data set. And despite these obvious limitations, it has proved to be a good illustration of the need to complement large-scale corpus analyses with methods of analysis that enables us to visualize large amounts of qualitative data. In the case of Kanye West and his collaborators, some noteworthy findings and hypotheses may never have been considered if we were not able to visualize such a specific linguistic variable like rhyme as it unfolded throughout a complete text.

The arc diagrams revealed a logogenetic patterning of rhyme that would have almost certainly been lost with large-scale, quantitative methodology. It was found that both GLC and Consequence dramatically increased their rhyme as they neared the end of their verse. This logogenetic intensification or ‘crescendo’ is important and should never be lost or submerged. It *means* something. And in this case, those rappers deliberately built-up their rhyme to reach a point of semantic and sensory salience which perfectly coincided with their spaceship ‘taking off’: ‘**blow**’ . . .

References

- Alim, H.S. (2003). *On some serious next millennium rap ish* Pharoahe Monch, hip hop poetics, and the internal rhymes of internal affairs, *Journal of English Linguistics*, 31, 60–84.
- Byron, L. (2007). *Children's poetry and limerick visualizations*. Retrieved 11 August 2008, from Lee Byron: www.leebyron.com/what/poetry/
- Keys, C. (2002). *Rap music and street consciousness*. Chicago, IL: University of Illinois Press.
- Macken-Horarik, M. (2003). Envoi: Intractable issues in appraisal analysis? *Text*, 23(2), 313–319.
- Martin, J. (2004). Mourning: How we get aligned. *Discourse and Society*, 15(2–3), 321–344.
- Martin, J. & Rose, D. (2003). *Working with discourse: Meaning beyond the clause*. London: Continuum.
- Martin, J. & White, P. (2005). *The language of evaluation: Appraisal in English*. London and New York: Palgrave.
- OHHLA.com (2008). – Favorite artists: Kanye West. *The Original Hip-Hop Lyrics Archive*. Retrieved 1 May 2008, from: www.ohhla.com/anonymous/kan_west/college/spaceshp.wst.txt.
- Richardson, E. (2006). *Hip hop literacies*. London: Routledge.
- Smitherman, G. (2006). *Word from the mother: Language and African Americans*. New York and London: Routledge.
- Ware, C. (2004). *Information visualization perception for design*. San Francisco, CA: Morgan Kaufman.
- Wattenberg, M. (2002). Arc diagrams: Visualizing structure in strings. Paper presented at the IEEE Symposium on Information Visualization (InfoVis'02), Boston, MA, 28–29 October 2002, 110–116.
- West, K. (2004). Spaceship. *The college dropout*. Roc-A-Fella/Def Jam, 5:24.
- Wise, J.A., Thomas, J.J., Pennock, K., Lantrip, D., Pottier, M., Schur, A. & Crow, V. (1999). Visualizing the non-visual: Spatial analysis and interaction with information from text documents. In S.K. Card, J.D. Mackinlay & B. Shneiderman (Eds.), *Readings in information viusalization: Using vision to think* (pp. 442–450). San Francisco, CA: Morgan Kaufmann.
- Zappavigna, M. (2007). Visualising instantiation: Text visualisation techniques for preserving logogenesis. Paper presented at Semiotic Margins: Reclaiming Meaning, Sydney, 10–12 December 2007.

Chapter 12

Multimodal Semiotics: Theoretical Challenges

J.R. Martin
University of Sydney

Multimodality

Over the past decade, social semioticians have recontextualized discourse analysis as multimodal discourse analysis, incorporating as they have an ever-expanding range of non-verbal semiotic descriptions in their analyses (e.g. Bateman 2008, Bednarek & Martin 2010, Martinec 2005, Royce & Bowcher 2007, Ventola & Guijarro 2009). In this work, these non-verbal semiotics have been productively conceived as kinds of language, drawing on a range of metalanguages and interpretations of metalanguages, especially functional linguistics and activity theory. In this chapter I will take one informing theory, systemic functional linguistics (hereafter SFL), and from its perspective ask questions about models of semiosis assumed in multimodal research. My goal is to push towards a degree of explicitness that will help foster dialogue and catalyze future research.

The Sign

To begin, I'd like to return to Saussure's conception of the sign (1959 Baskin translation of the *Cours* used here). On my reading, Saussure's sign is constituted by an inextricable bonding of signified (hereafter *signifié*) with signifier (hereafter *signifiant*) (see Figure 12.1). It follows that signs do not realize meaning; rather they make meaning. The common sense idea that signs stand for something, so that, for example, a stop sign means 'stop', is precisely what Saussure is trying to supplant.

On this reading, the question for Saussure is not what a sign means but how it means. And it means by fusing signifié with signifiant and organizing signs into systems in which they mean in relation to what they are not. Language is thus conceived as a system of signs, in which meaning is difference (or in Saussure's terms *valeur*). It follows that in a simple traffic lights system such as

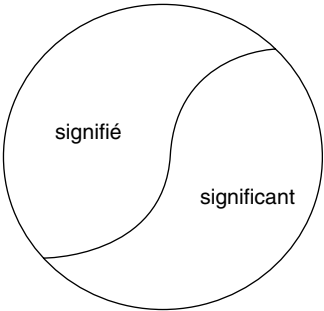


FIGURE 12.1 Bonding of signifié and signifiant in Saussure’s concept of the sign

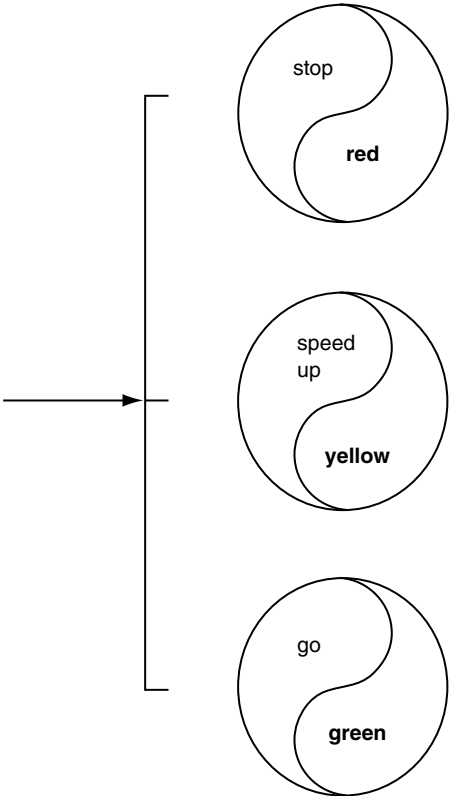


FIGURE 12.2 A simple system of signs

that in Figure 12.2, it doesn't matter whether we name signs using terms reflecting signifié or signifiant; what matters is the relationships among the signs – one sign versus another in the process of making meaning.

Based on this reading of Saussure one could ask of any multimodal analyst:

1. Do you conceive of the sign as an entity that realizes a meaning located outside itself (in the material world or in the mind or elsewhere) or alternatively as a meaning construing act?
2. Where and how, if at all, do you explicitly model valeur (i.e. the system of differences among signs)?

Realization

Developing Saussure, Hjelmslev (1961 translation of the *Prolegomena* used here) argues that language is not a simple system of signs, but a stratified system involving both a content plane and an expression plane. This can be naively read as Hjelmslev undoing Saussure's bonding of signifié and signifiant, and developing a plane (hereafter stratum) around each face of the sign. I, on the other hand, would prefer to read Hjelmslev as arguing that the bonding relation between signifié and signifiant is more complex than that articulated by Saussure and his students. For Hjelmslev language is conceived as a network of relationships; and the job of linguistics is thus to outline the nature of the complex of relationships binding signifié with signifiant. Because signifié and signifiant are mutually defining, linguistics cannot be about the signifié or signifiant per se. Linguistics has to be about the nature of the interrelationships which fuse them.

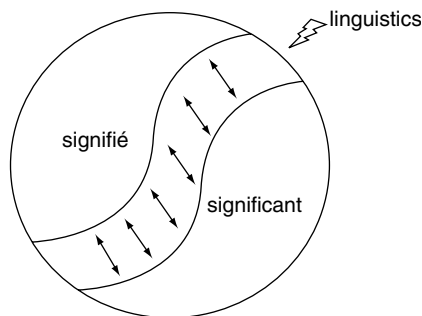


FIGURE 12.3 The focus of linguistics

Stratification

In order to explore this complexity, Hjelmslev proposes that the bonding of signifié with signifiant involves two interlocking systems of valeur – content form, which deals with systems of meaning, and expression form, which deals with systems of sound (or image or gesture if we take graphology or signing into account). In SFL terms, these two systems of mutually defining valeur are related by the concept of realization, and generally modelled as co-tangential circles, with content form subsuming expression form. In these terms Figure 12.4 is best read as a conceptualization of the bonding space outlined as the object of linguistic inquiry in Figure 12.3. Cléirigh (in preparation) refers to hierarchies of this kind as supervenient¹; Lemke (1984) refers to them as metaredundant, since content form is a pattern of expression form (a pattern of patterns in other words). Stratified systems can be conceived as evolving out of single stratum systems (such as animal language or the proto-language² spoken by infants up to around 18 months of age) through a process of emergent complexity (Matthiessen 2004).

From an SFL perspective, the emergence of grammatical metaphor and the elaboration of discourse resources for organizing meaning beyond the clause argue for a tri-stratal model of language with a stratified content plane – with discourse semantics an emergently complex pattern of lexicogrammatical patterns (cf. Halliday & Matthiessen 1999:237, Martin 1992) (see Figure 12.5).

Based on this reading of Hjelmslev and Halliday, one could ask of any multimodal analyst:

1. For a given semiotic system, how many strata are you proposing, and on which stratum is your description located?
2. Are your strata related by metaredundancy (as patterns of patterns)?
3. Are there distinct systems of valeur on each of the strata you propose?

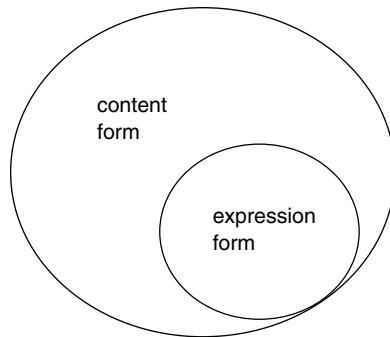


FIGURE 12.4 Expression form realizing content form in a stratified semiotic system (supervenience)

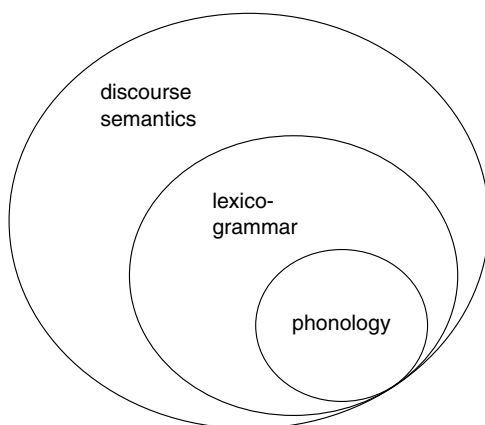


FIGURE 12.5 Stratification of content form as lexicogrammar and discourse semantics

4. Is there any ontogenetic or phylogenetic evidence suggesting that any stratified system you propose evolved from an unstratified or a less stratified system?

Rank

In SFL the principle of distinctive valeur is used to explore both relations between and within strata. Within strata, **one possibility** is that valeur is hierarchically organized in relation to units of different size, with higher-level units composed of one or smaller units, which may in turn be decomposed (Halliday & Matthiessen 2004, 2009). Distinctive levels of decomposition are referred to as ranks – for example a tone group consisting of one or more feet, a foot consisting of one or more syllables, and a syllable consisting of one or more phonemes in the phonology of a stress-timed language like English. What is critical here is that tone group systems differ from foot ones, foot ones from syllable ones and syllable ones from phoneme ones, and that the distinctive systems of valeur involved are related to one another by means of a constituency hierarchy. **The insistence on distinctive valeur constrains the number of ranks in the hierarchy, so that depth is not a simple function of the length of a unit.** The allocation of ranks to strata is partially exemplified for English in Figure 12.6.

Based on this reading of Halliday, one could ask of any multimodal analyst:

1. For a given stratum, how many ranks are you proposing, and at which rank is your description located?
2. Are there distinct systems of valeur on each of the ranks you propose?
3. Are your distinct systems of valeur related by constituency (as parts to wholes)?

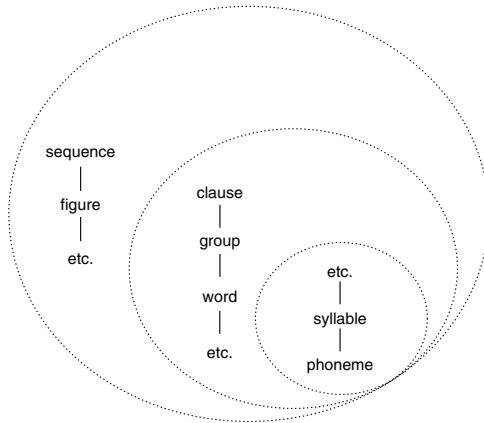


FIGURE 12.6 English rank scales (partial) by strata

Metafunction

In SFL the principle of distinctive valeur is also used to explore the organization of valeur with respect to kinds of meaning (Halliday & Matthiessen 2004, 2009). Distinctive regions of relatively interdependent systems are referred to as metafunctions – for example, ideational meaning (TRANSITIVITY), interpersonal meaning (MOOD) and textual meaning (THEME) at clause rank in lexicogrammar. What is critical here is that ideational systems complement interpersonal systems which complement textual ones. The three kinds of meaning cannot be integrated hierarchically into one super system; each perspective is partial and a comprehensive account of valeur depends on looking from three directions at the same time. When viewing Figure 12.7, it is important to keep in mind that metafunctions are not three parts of language, but three simultaneous dimensions of meaning.

Based on this reading of Halliday, one could ask of any multimodal analyst:

1. For a given semiotic system, how many metafunctions are you proposing?
2. Are there topologically distinct systems of valeur for each of the metafunctions you propose?
3. By what criteria are systems of valeur seen as relatively independent or interdependent of one another?

SFL makes further suggestions about the types of structural realization associated with different kinds of meaning (e.g. Martin 1996), with interpersonal meaning realized through prosodic structures, textual meaning through periodic structures and ideational meaning through particulate ones (orbital for experiential meaning and serial for logical meaning) (see Figure 12.8).

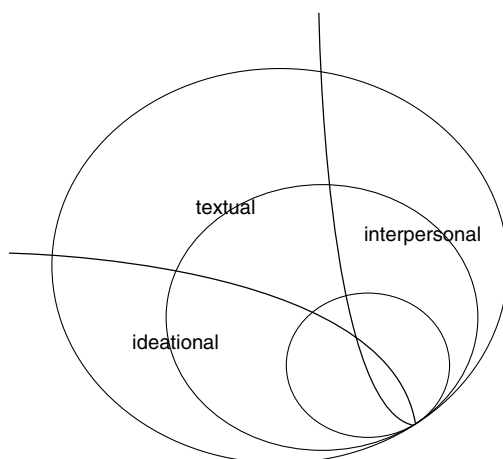


FIGURE 12.7 Metafunctions ranging across strata

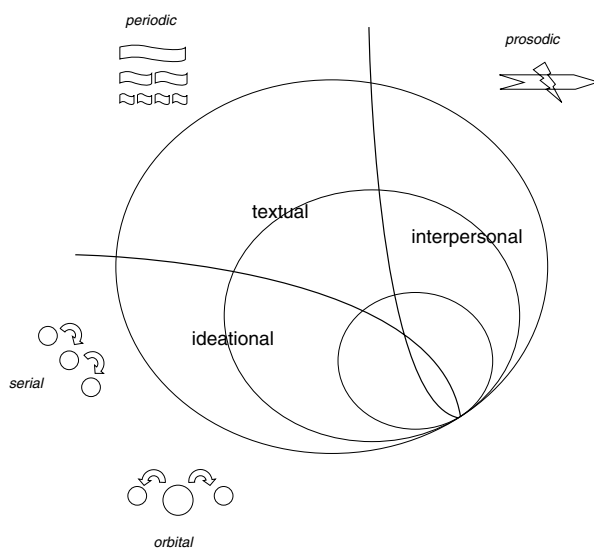


FIGURE 12.8 Types of structure in relation to metafunctions

Based on this reading of Halliday one could ask of any multimodal analyst:

1. For a given semiotic system, how many kinds of structural realization are you proposing?
2. Are the different types of realization associated with different types of meaning?
3. When analogizing from metafunctions in language to your semiotic system did you take kinds of meaning or types of structure as point of departure?

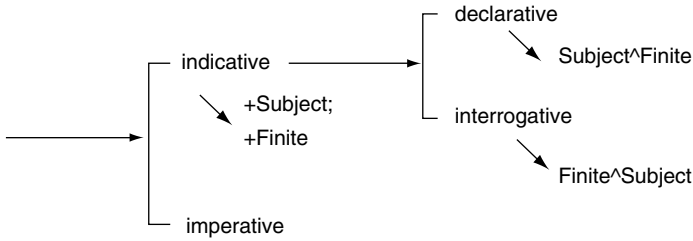


FIGURE 12.9 Realization of two English MOOD systems in structure

System/Structure Cycles

Absolutely critical to the discussion of stratification, rank and metafunction in this section is the theoretical dimension of axis, which underpins the relation of system and structure in SFL (Halliday & Matthiessen 2009:41–52). Like signifié and signifiant, system and structure are mutually defining complementarities. Paradigmatic relations (formalized in system networks) are ‘realized’ through syntagmatic relations (formalized in function structures), and conversely, syntagmatic relations constrain and motivate paradigmatic ones. A snippet of this interfacing is outlined in Figure 12.9 for English MOOD, where the choice of [indicative] conditions the presence of both a Subject and a Finite element of structure, and the more delicate choices of declarative or interrogative sequence them in relation to one another.

SFL depends on system/structure cycles of this kind to establish the ways in which systems formalizing valeur are related to one another, and emergently organized according to strata, rank and metafunction.

Based on this reading of paradigmatic and syntagmatic relations in SFL one could ask of any multimodal analyst:

1. Are your descriptions formalized as system/structure cycles, explicitly showing the relation of systemic choices to structural³ consequences?
2. How many system/structure cycles are you proposing and how are they related to one another (by strata, rank, metafunction or some other theoretical parameter)?

Instantiation

Developing Hjelmslev and Firth (e.g. Firth 1957a), Halliday argues that the hierarchy of realization outlined above has to be complemented by a hierarchy of instantiation relating the systemic potential of a language to instances of use (e.g. Halliday & Matthiessen 1999:382–387, 2009:79–82). In Hjelmslev’s terms, this is the relation of system to process (for semiotic systems in general)

or language to text (for linguistic systems), which functions alongside the relation of content form to expression form. **Whereas realization is a hierarchy of abstraction, instantiation is a hierarchy of generalization.** In meteorological terms instantiation is about the relation of long-term climatic conditions (system) to the weather patterns we experience moment-by-moment day-by-day (process). Instantiation is thus what makes it possible for a weatherman to say that the temperature today was 28 degrees (process), which was 3 degrees above average (system). This scale of generalization is outlined in Figure 12.10, which scales system in relation to genres and registers (context specific sub-potentials), text types (generalized actuals), texts (spoken, written or signed instances) and readings (our subjectified interpretations of what we see and hear).

It is commonplace to conflate the hierarchies of realization and instantiation (e.g. Martin 1992) in social semiotic theory and description. But we need to keep in mind that moving down a realization hierarchy like that⁴ outlined in Figure 12.11 from genre to phonology does not bring us any closer to a textual instance; conversely moving up a hierarchy of this kind from phonology to genre does not involve moving from 'weather to climate' as it were. All strata on the realization hierarchy instantiate (Figure 12.12). The different position of register and genre on the two hierarchies reflects the complementarity of the two hierarchies. Since genre is defined as a pattern of register (field, tenor and mode) patterns it is the highest level of realizational abstraction; but since genre (and thus register) are specializations of the meaning potential of a language as a whole, they are positioned below system on the instantiation hierarchy as subpotentializations of system.

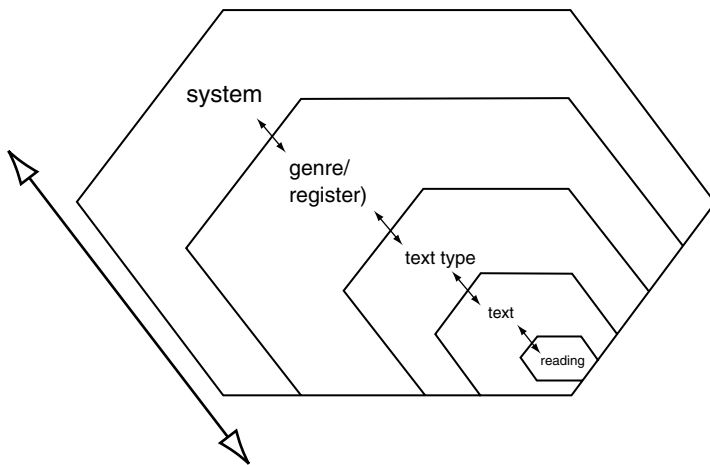


FIGURE 12.10 Hierarchy of instantiation (a cline of generalization/subpotentialization)

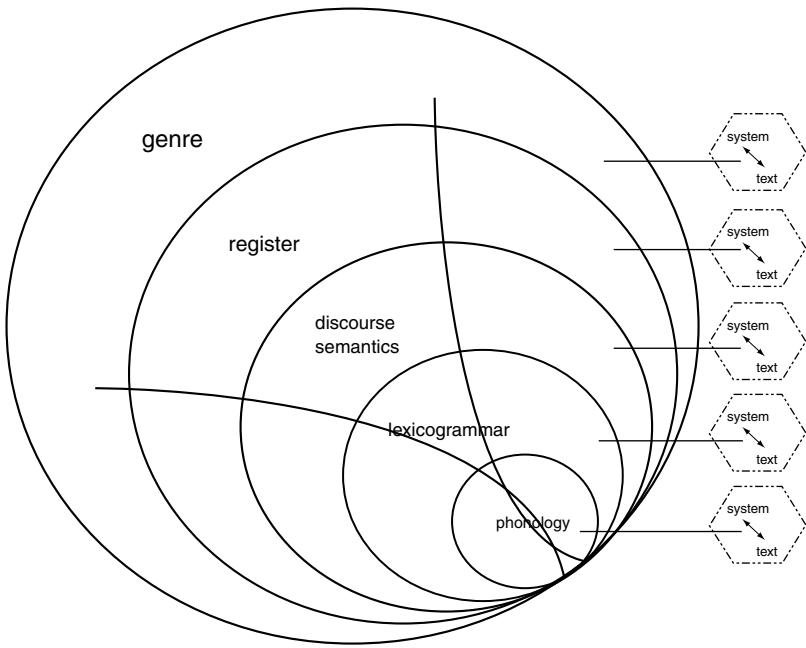


FIGURE 12.11 Realization in relation to instantiation (all strata instantiate)

Based on this reading of system⁵ and text in SFL one could ask of any multimodal analyst:

1. Is the complementarity of realization and instantiation addressed your description?
2. If so, how are you distinguishing axial realization (the defining interdependency of system and structure) from instantiation (the logogenetic unfolding of realizational resources as text)?
3. As far as the contextual specification of your system is concerned, what genres/registers/text types do you propose?

Coupling

While realization tells us what choices are available, instantiation explores which choices are taken up and how they are put together to form a text (Martin 2008a, b, 2010). The logogenetic process whereby meanings from different systems are woven together along the instantiation hierarchy is referred to by Martin (2008a) as coupling (Zappavigna et al. 2010). Coupling may involve combining choices from the same semiotic system (across ranks,

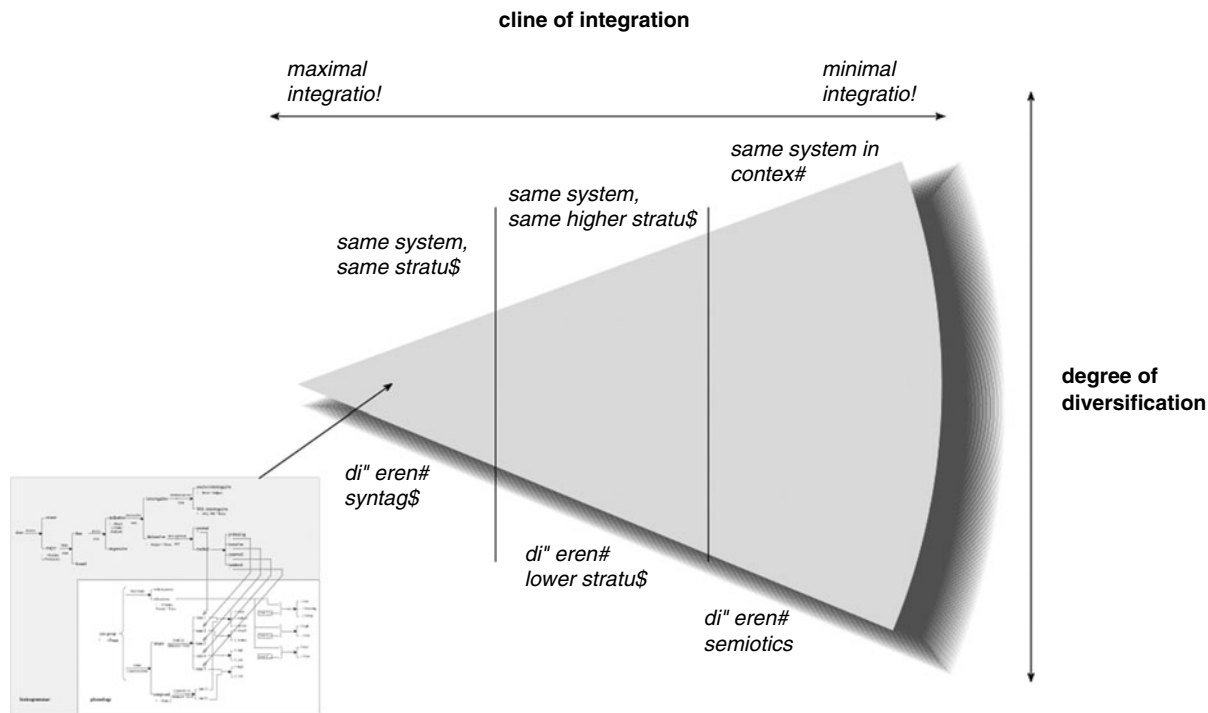


FIGURE 12.12 Matthiessen's (2009) cline of integration in relation to intermodality

metafunctions or strata or any simultaneous systems therein), or choices from different semiotic systems. **This creates a descriptive space for exploring how it is that distinct complementary semiotic systems can end up co-instantiated as a single unified multimodal text (Painter & Martin, in press).**

Matthiessen (2009:15–22, 2007b) discusses an alternative strategy for dealing with relations across semiotic systems (i.e. intermodality) involving realization. He proposes a cline of integration addressing the extent to which a common system of meaning for the two semiotics is established. Maximal integration is axial and involves setting up a single system whose axial realizations in structure draw on different modalities. Smaller degrees of integration take advantage of stratification, and posit a common system at a higher level of abstraction (e.g. semantics) whose realization is then distributed across complementary systems at a lower level of abstraction on the same hierarchy of realization. Minimal integration along this cline would involve establishing common ground only at the **highest levels of realizational abstraction** (i.e. maximally at the level of social context) and allowing for realization across different **denotative** semiotic systems (i.e. systems with their own **content/expression plane/s**; Hjelmslev 1961) in some sense sharing the same register or genre.

In relation to Matthiessen's proposed cline of integration one could ask the multimodal analyst:

1. Do you manage intermodality by proposing a single system of *valeur*, on a higher stratum or not, realized axially or inter-stratally by two or more modalities (realizational integration); or do you propose a coupling process weaving together meanings from different modalities in a single text (instantial integration)?

Coupling theory and description along the instantiation hierarchy remains in its infancy and awaits the development of animated visualization tools before real progress can be made (see Caldwell & Zappavigna, Chapter 10, Zhao 2010). Painter and Martin (in press) (cf. Chan, Chapter 7, Martin & Stenglin 2006, Royce 2007) discuss convergent and divergent coupling of image with verbiage in children's pictures book, in relation to ideational concurrence, interpersonal resonance and textual synchrony (inspired by Gill 2002). Table 12.1 exemplifies their concern with imagic character depiction and setting in relation to verbal attribution and circumstantiation, imagic facial expression and ambience in relation to verbal affect, and imagic composition (following Caple 2009) in relation to verbal information flow (see also Painter 2008 and Painter, Martin and Unsworth, Chapter 6).

As referenced in Martinec 2005, intermodal ideational relations have been modelled using categories based on the analysis of lexical cohesion (e.g. Royce 2007) or logico-semantic clause complex relations (Martinec & Salway 2005) in verbal texts. This kind of analysis in effect involves verbiage subsuming non-verbal meaning as supplementary 'as if verbal' texture, a strategy comparable to

Table 12.1 Sample convergent and divergent coupling matrix (across image and verbiage)

			convergent	\leftrightarrow coupling	divergent
metafunction	meaning potential visual	meaning potential verbal			
ideational			CONCURRENCE		
	character depiction	attribution			
	setting	circumstances			
	...				
interpersonal			RESONANCE		
	facial expression	affect			
	ambience	affect			
	...				
textual			SYNCHRONY		
	balance	Theme/New			
	array of foci	periodicity			
	...				

that deployed by Kress & Van Leeuwen 1996/2006 for compositional relations of Ideal/Real, Given/New and Centre/Margin across verbiage and image modalities – but which has arguably proved more difficult to implement for interpersonal meaning.

Based on these intermodal integration and complementarity issues one can ask the multimodal analyst:

1. Are the relations you recognize as obtaining between modalities in an intermodal text the same as those you find between units of a text in a monomodal one?
2. Do you recognize different kinds of intermodal relations depending on the kind of meaning involved (ideational/interpersonal/textual)?

Commitment

Instantiation also opens up theoretical and descriptive space for considering commitment (Martin 2008a, 2010), which refers to the amount of meaning instantiated as a text unfolds. This depends on the number of optional systems

taken up and the degree of delicacy pursued in those that are, so that the more systems entered, and the more options chosen, the greater the semantic weight of a text (Hood 2008). Commitment is one avenue for further exploring Kress and van Leeuwen's (e.g. 2001) notion of the affordances of a given semiotic system. By affordance they refer to the facility with which a certain kind of meaning is committed in one semiotic system compared to another – for example, the different ways in which verbiage and image register evaluation, through facial expression and bodily stance resources in image versus appraisal resources in verbiage (Martin & White 2005). On the one hand, language has extensive resources for inscribing affect, judgement and appreciation, whereas image arguably inscribes a narrower range of typological affectual distinctions and can only invoke, not inscribe, judgement and appreciation; at the same time, images arguably afford a **visceral somatic attitudinal punch** (cf. Martin 2001, Stenglin, Chapter 4) that can only be approximated in language through verbal imagery (i.e. lexical metaphor). Painter & Martin (in press) discuss examples of the rhetorical effect of complementary commitment of verbiage and image commitment in children's picture books.

Based on this discussion of affordances and commitment one could ask the multimodal analyst:

1. How do you model the amount of meaning committed and thereby the complementary contribution of different semiotic systems in an intermodal text?
2. How does the semantic weight of a given system's contribution reflect its affordances?

Semiotic Margins

The Semiotic Margins conference and the key proceedings of which this chapter is attempting to close, addressed the question of systems of meaning, which in some sense have marginal status as semiotic systems. Unlike language, image, music, dance and space, which are generally regarded as canonical semiotic systems, these systems are often treated as somehow dependent on denotative semiotic systems. Body language (including gesture, posture, facial expression and proxemics), and paralanguage (including vocal timbre, tempo and loudness) are well-known examples. For ease of reference we'll refer to body language and paralanguage together as body language below.

Body Language (including Paralanguage)

Recently, Cléirigh has made some helpful proposals for interpreting the nature of the dependency between body language and language (see also Zappavigna et al. 2009, 2010). Cléirigh argues for a reconsideration of body

language as three distinct phenomena resulting from the ontogenetic transition from protolanguage into language (Matthiessen 2004, Painter 2009, Painter et al. 2008).

Protolinguistic body language – Cléirigh's first point is that protolanguage does not disappear from an individual's semiotic repertoire as they learn their mother tongue. Rather it hangs around and further develops as a kinological system comprising one system/structure cycle⁶ to make meaning. Cléirigh uses Halliday's protolinguistic micro-functions (Halliday 2004a) to classify the meanings construed here and their bodily expression: for example, regulatory threat ✎ raised fist, instrumental invitation ✎ extended hand, personal affection ✎ smile, interactional togetherness ✎ eye contact. Dreyfus (Chapter 3) explores the meaning potential of a meaning system of this kind for a child interacting with (languageing and protolanguageing) adults.

Linguistic body language – Second, Cléirigh notes that a range of kinological (i.e. bodily expression) resources are subsumed and developed ontogenetically by language as part of its expression plane. These resources act in tandem with prosodic phonology, participating in the realization of rhythm and intonation. As explored in Zappavigna et al. 2009, 2010, both hand and head gestures are regularly deployed in sync with salient syllables, tonic feet and coextensively with tone groups; in addition, eyebrow and hand gestures move up and down in tune with choice of tone. As with linguistic expression form in general, these kinological resources construe interpersonal and textual meaning but are not deployed representationally.

Epilinguistic body language – Third, Cléirigh notes the development of what he refers to as an epilinguistic system which develops ontogenetically alongside language to illustrate, by way of drawing in the air, language's content plane. When used in the absence of spoken language, and elaborated by specialists, this system is known as mime. This system involves system/structure cycles on the same level of abstraction organized by metafunction. Examples include: for textual meaning, exophoric reference ✎ pointing; for interpersonal meaning, modalization of probability ✎ oscillating hand; and for ideational meaning, round entities ✎ drawing circles. For further examples from face-to-face academic teaching contexts see Hood⁷ (Chapter 2).

In summary then, Cléirigh interprets kinology as developing functionally in 3 directions from protolanguage:

1. as an elaboration of protolanguage,
 2. as part of language's expression form, coordinated with phonology, and
 3. as a supplementary instantiation of language's content plane. In doing so, Cléirigh makes explicit 3 senses in which body language can be interpreted as a semiotic margin.
- As protolanguage, body language lacks distinct system-structure cycles constituting content form realized through expression form, and it

cannot combine meanings (and so lacks metafunctions); it can however mean on its own (as wordless embodied interaction such as that commonly enacted between mature speakers and babies or pets).

- As linguistic body language, body language is part of language's expression form, gesturally scaffolding rhythm and intonation, and so cannot make meanings of its own nor on its own.
- As epilinguistic body language, body language imagically co-instantiates⁸ language's content form, generally at a far lower level of commitment; it can make and combine meanings, and as mime or as wordless embodied interaction between mature speakers, it can mean on its own.

These qualifications of the meaning potential of body language aside, it is important to appreciate in positive terms the visceral contribution embodied meaning makes to human interaction; whenever, wherever and however we can see one another, we cannot help but mean **and** move.

Parametric Systems

Van Leeuwen (2009; see also 1999, Kress & van Leeuwen 2001) discusses what he calls parametric systems. These systems have the property of involving a number of simultaneous systems, consisting of two terms, which are graded in relation to one another rather than in dichotomous opposition. He illustrates systems of this kind through sound quality (Figure 12.13), where, for example, a singing voice can be more or less tense or lax, loud or soft, high or low, and so on (see also Caldwell, in press). Van Leeuwen notes that the same kind of parametric system can be proposed for typography and colour (cf. Kress & van Leeuwen 2002, van Leeuwen 2005a).

Taking Cléirigh's perspective on body language as point of departure, systems of this kind could all be explored in relation to language for their protolinguistic, linguistic and epilinguistic potential. As elaborated protolanguage they co-opt physical and material resources to construe the visceral embodied meanings that can be fashioned out of sound (cf. McDonald, Chapter 5), colour and typeface. At the same time, like linguistic body language, they can be co-opted by the expression plane of language to punctuate a phase of discourse or enhance its tone (cf. Caldwell and Zappavigna, Chapter 11). As **epilanguage** they can be deployed ideationally to represent physical or biological phenomena (e.g. bird calls), interpersonally to register feeling (e.g. imminent danger music in a film soundtrack) and textually to highlight meanings (e.g. bold-face, italics, special font, colour in text).

We also need to allow for the fact that parametric resources can interact with denotative semiotics other than language, for example, image, music or dance. The sound track for a film, for example, arguably functions in all three ways in relation to moving images, as would the music accompanying dance,

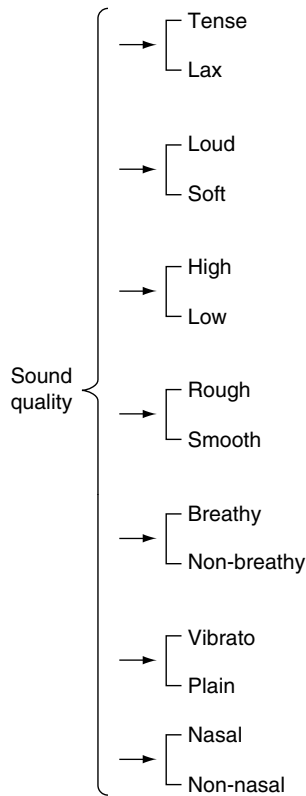


FIGURE 12.13 Van Leeuwen's (1999) systems for sound quality (a parametric system)

or the colour involved in images or buildings. In general terms then, we need to ask about the protosemiotic deployment of parametric resources, their possible function on the expression plane of denotative semiotic systems and their **episemiotic** potential.

In essence then, what we are proposing here is that the uncertain status of parametric systems as semiotic form or material substance be resolved by treating them as a semiotic co-option of physical resources, construing meaning alongside or instead of denotative semiotics as protosemiosis or episemiosis, or providing further **scaffolding for the expression form of denotative semiotics**. This implies that in order to be construed as meaning-making systems, parametric resources have to be factored out into three kinds of systems, organized by micro-function (protosemiosis) or by metafunction (denotative semiotic – textual and interpersonal meaning only; and episemiosis – ideational, interpersonal and textual meaning).

In light of this reading of Cléirigh, one could ask:

1. Is the semiotic system you are working on a denotative semiotic system, with its own content form and expression form?
2. If not, does it involve parametric resources of the kind outlined by van Leeuwen (i.e. multiple, simultaneous, graded, binary systems)?
3. If so, could it be usefully factored into protosemiotic, denotative semiotic and epilinguistic systems?

Identity (Individuation and Affiliation)

Van Leeuwen (2009) further comments on parametric resources in relation to the role they play in construing identity (see also Kress & van Leeuwen 2002, van Leeuwen 2005b on style). In effect what we are looking at here is an elaboration of protolanguage to construe personae – where such are understood in Firth's (e.g. 1957b) terms as the repertoire of personalities we assume in order to play our role in the speech fellowships to which we belong. As Firth (1957b:191–192) once quipped in relation to phonology and accent, 'It is part of the meaning of an American to sound like one'. Van Leeuwen (2009) explores this in particular in relation to singing and acting personae – for example, actor Marlon Brando's Godfather voicing and singer Bing Crosby's crooning style.

In SFL terms (Bednarek & Martin 2010, Martin 2008c; cf. Tann 2010), this turns discussion to a third hierarchy, individuation (operating alongside realization and instantiation), which brings a focus on users of language into the picture. To date, SFL researchers have explored two complementary ways of thinking about individuation. One way, inspired by Hasan's work on semantic variation (Hasan 2005, 2009, Williams 2005), interprets individuation as a hierarchy of allocation whereby semiotic resources are differentially distributed amongst users – both in terms of which options are available and of those available, which are likely to be taken up in specific contexts of instantiation. Bernstein uses the metaphor of **reservoir** and repertoire to describe the semiotic affordances of users in relation to their communities as a whole along these lines:

I shall use the **term repertoire** to refer to the set of strategies and their analogic potential possessed by any one individual and the **term reservoir** to refer to the total of sets and its potential of the community as a whole. Thus the *repertoire* of each member of the community will have both a common nucleus but there will be differences between the *repertoires*. There will be differences between the *repertoires* because of the differences between members arising out of differences in members' context and activities and their associated issues. (Bernstein 1996:157)

A second, complementary perspective on individuation looks at how personae mobilize social semiotic resources to affiliate with one another – how they share attitude and ideation couplings, in Knight's (2010) terms, to form bonds, and how these bonds then cluster as belongings of different orders (including relatively 'local' familial, collegial, professional and leisure/recreational affiliations and more 'general' fellowships reflecting 'master identities' including social class, gender, generation, ethnicity, and dis/ability). As with realization and instantiation, it is difficult to find a neutral term which privileges neither a top-down nor a bottom-up perspective. We'll adopt the term individuation for this hierarchy here, keeping in mind that it is concerned with both how semiotic resources are distributed among users (allocation) and how these resources are deployed to commune (affiliation). An outline of this user-oriented hierarchy is presented in Figure 12.14.

As with instantiation, we need to keep in mind that as far as realization is concerned, all strata individuate (Figure 12.15). Recalling Firth, its not just part of the meaning of an American to sound like one; being American involves the coupling of identity construing choices from lexicogrammar, discourse semantics, register and genre as well. Taking all three hierarchies (realization, instantiation and individuation) into account is a challenging task; but as social semioticians, we have to keep in mind that speakers always already individuate **as** they instantiate **as** they re/deploy the realization resources of their culture.

As far as protolanguage and identity are concerned, it appears that Halliday's interactional and personal micro-functions in particular (his 'me and you' and 'here I come' functions) co-opt a wide range of parametric resources to individuate personae that affiliate as social groups. So we don't just rely on linguistic (and other denotative semiotic) resources to construe identity and community; we take advantage of all we learned about embodied belonging as we first learned to mean – and throughout life, add to our repertoire of communing affordances as parametric resources become available to us. Whether or not we should continue to refer to these elaborated resources deployed by mature speakers as 'protolinguistic' is an important question; having made the point I'll defer to the wisdom of our readers at this time.

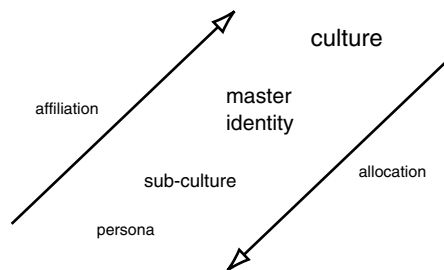


FIGURE 12.14 Individuation as a hierarchy of affiliation and allocation

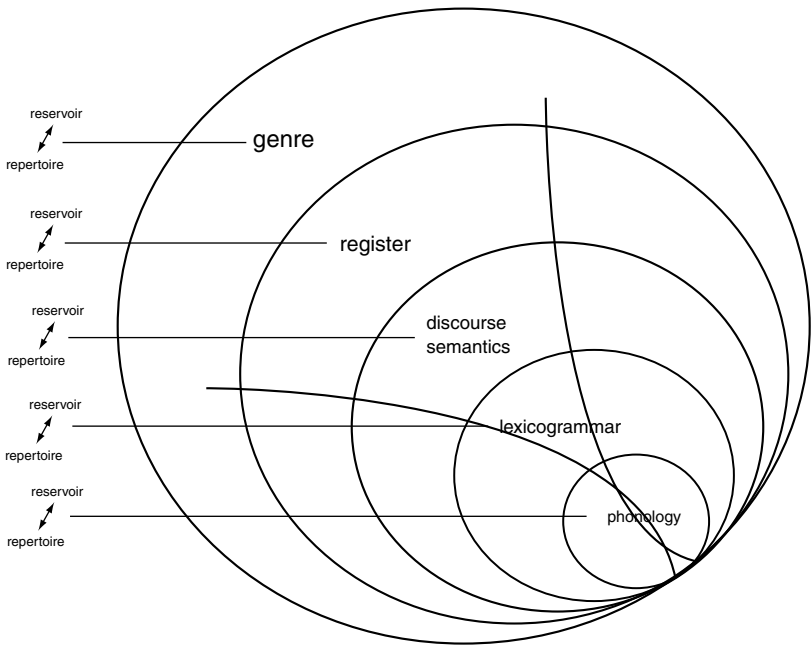


FIGURE 12.15 Realization in relation to individuation (all strata individuate)

In light of these concerns with identity and affiliation, one could ask:

1. How do you describe the allocation of the semiotic resources you are focusing on to repertoires of users?
2. How do these repertoires engender communities of such users?
3. Is there a distinctive role for denotative semiotic, protosemiotic and episemiotic systems in this process?

Beyond Semiosis (Circumvention)

This brings us finally to the question of the limits of semiosis. Is there meaning beyond language (and other denotative semiotics), protosemiosis and episemiosis? Or to put this more helpfully, where and when do we draw the line between semiosis and the biological systems⁹ out of which it evolved (where these in turn evolved out of physical systems)? In Cl  irigh's terms, a smiling face, for example is not just a physical act; it also means. And with the evolution of smileys, it comes to function in electronic discourse as an epilinguistic resource construing affect. And as smileys become more common still, variegated to register a range of feelings and finding conventional expression on traditional keyboards

as say :-) or :-(, one might even argue that they have been fully co-opted by language's expression plane as part of graphology (cf. Knox 2009a, b). Over time, the overwhelming trend is to make things mean, and co-opt more and more of the somatic and physical environment of semiosis into the semiosis itself, as technology affords. Currently electronic communication and 'clinical' intervention (e.g. plastic surgery, piercing and tattoo) are key technologies shifting the borders of what counts as in or outside of a social semiotician's gaze.

However excluded, somatic and physical systems are the material context in which we mean, so straddling the borders of Figure 12.16, on some kind of interdisciplinary basis or other, is ever a useful corrective to social semiotic analysis alone. We also need to consider the possibility of construing somatics **as if** it were semiosis; Martinec's work on action, for example (1998, 2000a, b, c, 2001; cf. Martinec 2004 on what is termed here), might well be considered an endeavour of this kind by those (not including the author) who view its purview as somehow beyond semiosis.

In light of these concerns with the limits of semiosis, one could ask:

1. On what basis do you distinguish between the semiosis you are considering and its biological and/or physical environment?

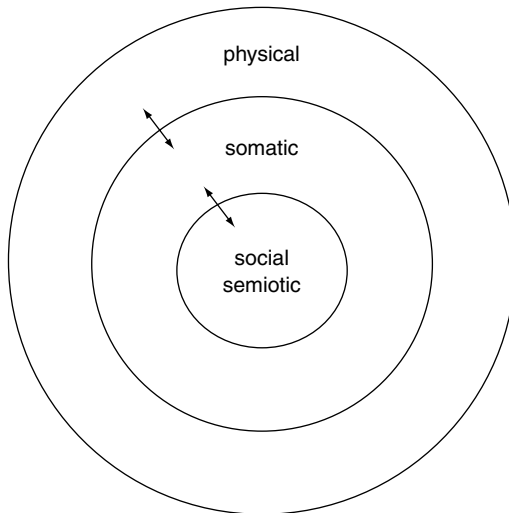


FIGURE 12.16 Circumvention (semiosis embedded in biological systems embedded in physical systems)

Note: Cleirigh in his current work in fact models semiotic systems as supervening on neurobiological ones, which are in turn suprevene on physical ones; I am borrowing an earlier term of his here (circumvention) to highlight the difference between supervenience within a semiotic system and the relationship of one order of system to another. Once again apologies to Chris for the now deliberate 'misreading'.

2. To what extent do you feel that interdisciplinary research involving neurobiologists and/or physicists is necessary to give a full account of the discourse you are considering?
3. Are you deliberately treating aspects of biological and physical materiality as if they were semiosis?

Challenging Theory

Let me make just two points in closing. The first is the importance of bringing time into the picture, which I have not had space to pursue here. Clearly, multimodal texts unfold through time (Zhao 2010), and what was referred to as instantiation above is a logogenetic process – snowballing meaning. In addition, clearly, identity is something that develops throughout the lifetime of an individual, and what was referred to as individuation above is an ontogenetic process accumulating logogenesis as repertoire – seasons of meaning (Painter 2009) (see Figure 12.17). And finally, clearly, systems evolve, as reservoirs of meaning adapt phylogenetically to changing technologies and environmental

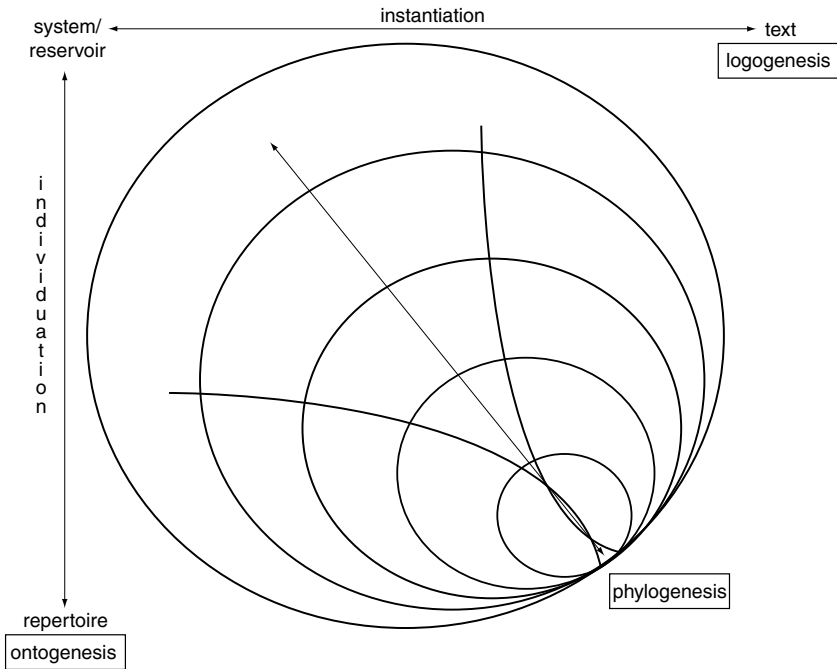


FIGURE 12.17 Instantiation and individuation in relation to genesis (logogenesis, ontogenesis and phylogenesis)

conditions – glacial meaning (Halliday 2004b). So there are clearly many questions to be asked about a given non-verbal semiotic system in relation to these dimensions of time.

The second has to do with the fact that no matter how assiduous a researcher one assigns to answering the questions proposed above, they would inevitably find themselves frustrated by a relative lack of explicitness in extant multimodal work. In part this stems from the fact that not all research is informed by an SFL model of language and so the kinds of questions and answers entailed by a theory of this kind do not arise.¹⁰ For SFL research, on the other hand, this appears to stem from the nature of pioneering work, where breakthroughs may depend on deliberately restricting one's gaze – asking questions about metafunctions, for example, but not about all strata or ranks at the same time, or asking about types of structure rather than kinds of meaning, or asking about one kind of meaning rather than another. Now that the frontiers of knowledge have been opened up, however, it is time to turn the full power of the theory back onto what has been achieved, and at the same time to begin to renovate the theory in terms of what has already been found. The work on instantiation referred to at several points throughout this chapter, for example (Bednarek & Martin 2010), would not have occurred had it not been for intermodality and the vexing puzzle of how complementary semiotic systems are woven together instantially as unified multimodal texts.

For those about to so gaze, I salute you. And for the peerless pioneers authoring chapters in this volume, my deepest thanks – for the extreme provocations giving rise to the questions I am posing here.

Acknowledgement

I am deeply indebted to Chris Cléirigh for many of the concepts developed in this chapter, and for his guidance in averting some of the worst of my misunderstandings in relation to his work; the recontextualization of some of his ideas here is my responsibility alone.

Notes

¹ Supervenience sits in contrast to circumvenient systems discussed in this chapter. Supervenient systems are in a relation of realization, whereas circumvenient systems are in a relation of embedding.

² Halliday (e.g. 1975) in fact describes infant's protolanguage a bi-stratal in spite of the fact that for languages of this kind *valeur* on the two strata would have to be identical. I prefer an emergent complexity model here in which stratification cannot be proposed in the absence of distinct *valeur* (and thus metaredundancy); see Painter 1984:34–36 and Matthiessen 2007a:516–519 for discussion (for

Matthiessen the two SFL theoretical dimensions of axis and strata are conflated in protolanguage).

- ³ In answering this query we have to keep in mind that structural realization may involve a single unit; in Tagalog, for example, the realization of interrogative from a MOOD network like that in Figure 12.9 would be the question particle *ba*, not a syntagm like Finite^Subject. It is also crucial to distinguish axial realization from instantiation; both dimensions of axis, paradigmatic system and syntagmatic structure are instantiated in texts as they logogenetically unfold – the structural realization of a system is NOT its instantiation!
- ⁴ To this hierarchy I have added the strata of register and genre (after Martin 1992, Martin & Rose 2008), by way of incorporating social context as higher levels of abstraction; in Hjelmslev's terms register and genre are connotative semiotic systems, defined as systems which take another semiotic as their expression plane (versus denotative semiotics which have their own expression plane).
- ⁵ Note that I am now, like most systemicists, in the uncomfortable position of having used the term system in two ways – with respect to axis as the paradigmatic complement of syntagmatic structure, and with respect to instantiation as the meaning potential specialized in texts. This dual usage of the term in SFL is too sedimented to undo here, and in any case reflects the privileged position given to paradigmatic relations are far as modelling the systemic reservoir of meanings in a culture is concerned.
- ⁶ Halliday (e.g. 2004a) and Matthiessen (e.g. 2007a:516) prefer an interpretation of protolanguage in which axis is conflated with strata (i.e. system conflated with content form and structure with expression form) and thus refer to protolinguistic systems as bi-stratal.
- ⁷ Hood's work on gestures illustrating abstract concepts in disciplinary discourse suggests a possible extension of Cl  irigh's conception of epilinguistic systems to diagrams, which co-opt images in order to co-instantiate language's content plane, by drawing on a page (versus drawing in the air); a powerpoint presentation or instruction manual consisting solely of diagrams would be thus akin to mine.
- ⁸ Treating epilinguistic body language as co-instantiating language's content form is an extrapolation by the author from Cl  irigh's work.
- ⁹ Matthiessen (e.g. 2004, 2007a), following Halliday, proposes four stages of evolution, physical, biological, social and semiotic; I find it hard to imagine a social system without some form of communication, so prefer the physical, biological and social semiotic genesis implied in Figure 12.16 here. It might be possible, however, to draw a line between social systems dependent on 'protolanguage' type systems alone and those additionally deploying stratified linguistic systems, which may be what Halliday and Matthiessen have in mind.
- ¹⁰ This would be especially true for activity theorists, who instead of seeing action as a kind of meaning see meaning as a kind of behaviour involving verbal artefacts (e.g. Norris & Jones 2005).

References

- Bateman, J. (2008). *Multimodality and genre: A foundation for the systematic analysis of multimodal documents*. London: Palgrave Macmillan.

- Bednarek, M. & Martin, J.R. (Eds) (2010). *New discourse on language: Functional perspectives on multimodality, identity, and affiliation*. London: Continuum.
- Bernstein, B. (1996). *Pedagogy, symbolic control and identity: Theory, research, critique*. London: Taylor & Francis. [Revised Edition 2000].
- Caldwell, D. (in press). Making Many Meanings in Popular Rap Music. In A. Mahboob & N. Knight (Eds), *Directions in Applicable Linguistics*. London: Continuum.
- Caple, H. (2009). Playing with words and pictures: Text-image relations and semiotic interplay in a new genre of western news reportage. PhD Thesis, Department of Linguistics, University of Sydney, Sydney.
- Cléirigh, C. (in preparation). *The life of meaning*. Draft manuscript.
- Firth, J.R. (1957a). A synopsis of linguistic theory, 1930–1955. In *Studies in Linguistic Analysis* (Special volume of the Philological Society), pp. 1–13. London: Blackwell. Reprinted in Palmer, F.R. (Ed.) (1968). *Selected papers of J R Firth, 1952–1959* (pp. 168–205). London: Longman.
- Firth, J.R. (1957b). Personality and language in society. *Papers in linguistics 1934–1951* (pp. 177–189). Oxford: Oxford University Press.
- Gill, T. (2002). Visual and verbal playmates: An exploration of visual and verbal modalities in children's picture books. BA Hons Thesis, Department of Linguistics, University of Sydney, Sydney.
- Halliday, M.A.K. (1975). *Learning how to mean: Explorations in the development of language*. London: Edward Arnold (Explorations in Language Study).
- Halliday, M.A.K. (2004a). *The language of early childhood*. London: Continuum (Vol. 4 in the Collected Works of M.A.K. Halliday).
- Halliday, M.A.K. (2004b). *The language of science*. London: Continuum (Vol. 5 in the Collected Works of M.A.K. Halliday). London: Continuum.
- Halliday, M.A.K. & Matthiessen, C.M.I.M. (1999). *Construing experience through language: A language-based approach to cognition*. London: Cassell.
- Halliday, M.A.K. & Matthiessen, C.M.I.M. (2004). *An introduction to functional grammar* (3rd edn). London: Edward Arnold.
- Halliday, M.A.K. & Matthiessen, C.M.I.M. (2009). *Systemic functional grammar: A first step into theory*. Beijing: Higher Education Press.
- Hasan, R. (2005). *Language, society and consciousness*. London: Equinox (Vol. 1 in The Collected Works of Ruqaiya Hasan, edited by J. Webster).
- Hasan, R. (2009). *Semantic variation: Meaning in society and sociolinguistics*. London: Equinox (The Collected Works of Ruqaiya Hasan, edited by J. Webster).
- Hjelmslev, L. (1961). *Prolegomena to a theory of language*. Madison, WI: University of Wisconsin Press.
- Hood, S. (2008). Summary writing in academic contexts: Implicating meaning in processes of change. *Linguistics and Education* 19, 351–365.
- Knight, N. (2010). Wrinkling complexity: Concepts of identity and affiliation in humour. In M. Bednarek & J.R. Martin (Eds), *New discourse on language: Functional perspectives on multimodality, identity, and affiliation* (pp. 35–58). London: Continuum.
- Knox, J.S. (2009a). Visual minimalism in hard news: Thumbnail faces on the *smh* online home page. *Social Semiotics*, 19(2), 165–189.
- Knox, J.S. (2009b). Punctuating the home page: Image as language in an online newspaper. *Discourse and Communication*, 3(2), 145–172.

- Kress, G. & van Leeuwen, T. (1996). *Reading images: The grammar of visual design*. London: Routledge [Revised second edition 2006].
- Kress, G. & van Leeuwen, T. (2001). *Multimodal discourse – The modes and media of contemporary communication*. London: Edward Arnold.
- Kress, G. & van Leeuwen, T. (2002). Colour as a semiotic mode: Notes for a grammar of colour. *Visual Communication*, 1(3), 343–368.
- Lemke, J. (1984). *Semiotics and education*. Toronto: Toronto Semiotic Circle (Monographs, Working Papers and Publications 2).
- Martin, J.R. (1992). *English text: System and structure*. Amsterdam: Benjamins.
- Martin, J.R. (1996). Types of structure: Deconstructing notions of constituency in clause and text. In E.H. Hovy & D.R. Scott (Eds), *Computational and conversational discourse: Burning issues – an interdisciplinary account* (pp. 39–66). Heidelberg: Springer (NATO Advanced Science Institute Series F – Computer and Systems Sciences, Vol. 151).
- Martin, J.R. (2001). 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). Brepols (Semiotic & Cognitive Studies X).
- Martin, J.R. (2008a). Tenderness: Realisation and instantiation in a Botswanan town. *Odense Working Papers in Language and Communication* (Special Issue of Papers from 34th International Systemic Functional Congress edited by Nina Nørgaard) (pp. 30–62).
- Martin, J.R. (2008b). Intermodal reconciliation: Mates in arms. In L. Unsworth (Ed.), *New literacies and the English curriculum: Multimodal perspectives* (pp. 112–148). London: Continuum.
- Martin, J.R. (2008c). Innocence: Realisation, instantiation and individuation in a Botswanan town. In N. Knight & A. Mahboob (Eds), *Questioning linguistics* (pp. 27–54). Cambridge: Cambridge Scholars Publishing.
- Martin, J.R. (2010). Semantic variation: Modelling system, text and affiliation in social semiosis. In M. Bednarek & J.R. Martin (Eds), *New discourse on language: Functional perspectives on multimodality, identity, and affiliation* (pp. 1–34). London: Continuum.
- Martin, J.R. & Rose, D. (2008). *Genre relations: Mapping culture*. London: Equinox.
- Martin, J.R. & Stenglin, M. (2006). Materialising reconciliation: Negotiating difference in a post-colonial exhibition. In T. Royce & W. Bowcher (Eds), *New directions in the analysis of multimodal discourse* (pp. 215–238). Mahwah, NJ: Lawrence Erlbaum Associates.
- Martin, J. & White, P.R.R. (2005). *The language of evaluation: Appraisal in English*. London: Palgrave (Chinese translation in preparation for Peking University Press).
- Martinec, R. (1998). Cohesion in action. *Semiotica* 1, 20(½), 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. (2000c). Construction of identity in M Jackson's 'Jam'. *Social Semiotics*, 10(3), 313–329.
- Martinec, R. (2001). Interpersonal resources in action. *Semiotica*, 135(1/4), 117–145.
- Martinec, R. (2004). Gestures which co-occur with speech as a systematic resource: The realisation of experiential meaning in indexes. *Social Semiotics*, 14(2), 193–213.

- Martinec, R. (2005). Topics in multimodality. In R. Hasan, C.M.I.M. Matthiessen & J. Webster (Eds), *Continuing discourse on language* (Vol. 1) (pp. 157–181). London: Equinox.
- Martinec, R. & Salway, A. (2005). Image-text relations in new (and old) media. *Visual Communication*, 4(3), 337–371.
- Matthiessen, C.M.I.M. (2004). The evolution of language: A systemic functional exploration of phylogenetic phases. In G. Williams & A. Lukin (Eds), *The development of language: Functional perspectives on species and individuals* (pp. 45–89). London: Continuum.
- Matthiessen, C.M.I.M. (2007a). The ‘architecture’ of language according to systemic functional theory: Developments since the 1970s. In R. Hasan, C.M.I.M. Matthiessen & J. Webster (Eds), *Continuing discourse on language* (Vol. 2) (pp. 505–462). London: Equinox.
- Matthiessen, C.M.I.M. (2007b). The multimodal page: A systemic functional exploration. In Royce & Bowcher, *New directions in the analysis of multimodal discourse* (pp. 1–62). Mahwah, NJ: Lawrence Erlbaum Associates.
- Matthiessen, C.M.I.M. (2009). Multisemiosis and context-based register typology: Registerial variation in the complementarity of semiotic systems. In E. Ventola & A.J.M. Guijarro (Eds), *The world told and the world shown: Multisemiotic issues* (pp. 11–38). London: Palgrave Macmillan.
- Norris, S. & Jones, R. (Eds) (2005). *Discourse in action: Introducing mediated discourse analysis*. London: Routledge.
- Painter, C. (1984). *Into the mother tongue: A case study in early language development*. London: Pinter.
- Painter, C. (2008). The role of colour in children’s picture books: Choices in AMBIENCE. In L. Unsworth (Ed.), *New literacies and the English curriculum: Multimodal perspectives* (pp. 89–111). London, Continuum.
- Painter, C. (2009). Language development. In M.A.K. Halliday & J. Webster (Eds), *Companion to systemic-functional linguistics* (pp. 87–103). London: Continuum.
- Painter, C. & Martin J.R. (in press). *Intermodal complementarity: Modelling affordances across verbiage and image in children’s picture books*. In Ilha do desterro: A journal of English language, literatures in English and cultural studies xx (Special Issue on Multimodality ed. F. Veloso).
- Painter, C., Derewiaka, B. & Torr, J. (2008). From microfunction to metaphor: Learning language and learning through language. In R. Hasan, C.M.I.M. Matthiessen & J. Webster (Eds). *Continuing discourse on language* (Vol. 2) (pp. 563–588). London: Equinox.
- Royce, T. (2007). Intersemiotic complementarity: A framework for multimodal analysis. In Royce & Bowcher, *New directions in the analysis of multimodal discourse* (pp. 63–111). Mahwah, NJ: Lawrence Erlbaum Associates.
- Royce, T. & Bowcher, W. (2007). *New directions in the analysis of multimodal discourse*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Saussure, F. de (1959). *Course in general linguistics* (C. Bally & A. Sechehaye, Eds and W. Baskin, trans.). New York: McGraw-Hill.
- Tann, K. (2010). Imagining communities: A multifunctional approach to identity management in texts. In M. Bednarek & J.R. Martin (Eds), *New discourse on language: Functional perspectives on multimodality, identity, and affiliation* (pp. 163–194). London: Continuum.

- van Leeuwen, T. (1999). *Speech, music, sound*. London: Macmillan.
- van Leeuwen, T. (2005a). Typographic meaning. *Visual Communication*, 4(2), 137–143.
- van Leeuwen, T. (2005b). *Introducing social semiotics*. London: Routledge.
- van Leeuwen, T. (2009). Parametric systems: The case of voice quality. In C. Jewitt (Ed.), *The Routledge handbook of multimodal analysis* (pp. 68–77). Oxon/New York: Routledge.
- Ventola, E. & Guijarro, A.J.M. (Eds) (2009). *The world told and the world shown: Multisemiotic issues*. London: Palgrave Macmillan.
- Williams, G. (2005). Semantic variation. In R. Hasan, C.M.I.M. Matthiessen & J. Webster (Eds), *Continuing discourse on language* (Vol. 2) (pp. 457–480). London: Equinox.
- Zappavigna, M., Cléirigh, C., Dwyer, P. & Martin, J.R. (2009). Body language in NSW Youth Justice Conferencing. Paper presented at the Towards Restorative Justice Conference, University of Sydney, 7–9 December 2009.
- Zappavigna, M., Cléirigh, C., Dwyer, P. & Martin, J.R. (2010). The coupling of gesture and phonology. In M. Bednarek & J.R. Martin (Eds), *New discourse on language: Functional perspectives on multimodality, identity, and affiliation* (pp. 219–236). London: Continuum.
- Zhao, S. (2010). Intersemiotic relations as logoegenetic patterns: Towards the restoration of the time dimension in hypertext description. In M. Bednarek & J.R. Martin (Eds), *New discourse on language: Functional perspectives on multimodality, identity, and affiliation* (pp. 95–218). London: Continuum.

Index

- absolutist 108
- action 263
- actualization 41
- affiliation 11, 12, 260–2
 - social process of affiliation 14–15, 17, 23, 24–5, 26*n*
- allocation 261
- ambience 97, 130
- analyst 102, 105
- appraisal 7, 11, 12–13, 17, 43–5, 78, 234–5
 - see also* attitude, engagement, graduation
- arc diagram 230, 232–4, 236–9, 241
 - see also* visual representation, text visualisation
- assessment 144, 164
 - item difficulty 161, 164
 - stimulus texts 145, 149, 151–6
 - test items 157–8
 - test materials 144
- attitude 11, 12–14, 17, 20, 23, 34, 43–4, 49
 - affect 13, 43, 115
 - appreciation 13, 15, 23–4, 43
 - judgement 13–14, 22–5, 43
 - inscribed 13–14, 44
 - invoked 13–14, 44
- augmentative and alternative communication 59
- axis 250–1
- balance 126, 134–41
- binding in 3D space 78
 - Bound 79, 87, 88, 94
 - Too Bound 78, 83, 86
 - Unbound 79, 86, 87, 88, 89, 90
 - Too Unbound 78, 89
- binding in image 97, 127
 - Bound image 127, 130–3
 - Unbound image 127–30
- body 103, 117, 119
- body language 8, 10–11, 31–5, 37–49, 256–9
 - epilinguistic body language 33–5, 41, 43, 48, 257–8
 - linguistic body language 33, 43, 257–8
 - protolinguistic body language 33, 257–8
 - see also* facial expression, gaze, gesture, posture
- bonding 9, 10, 14, 26*n*, 79
 - attitudinal re/alignment 80, 82, 90–1
 - classification & framing 80, 84–5, 87, 88–9, 91
 - contact 94, 96
 - icons 79, 83, 86, 95
- causation 114
- change 114, 115
- circumvention 262–3
- cohesion 125, 173
 - cohesive relations 146–8
 - see* image-text relations, modelling
- commitment 255–6
- communication partner 61
- communities 14, 15, 22–3, 24–5, 26*n*, 102, 104
- composer 102, 103, 109
 - performer 102
- composition 134
 - centrifocal 134–6
 - iterating 134, 140
 - polarized 136–9, 140
 - triptych 136
 - see also* textual metafunction
- content plane (content form) 245

- context 103, 106, 179
 - stratified model 179
 - context of culture (genre) 179
 - anecdote 183–5
 - story 181–3
 - context of situation (register) 179
 - field 76–7, 82, 86, 179–85
 - mode 179–85, 193
 - tenor 179–85
- contextualized image 127, 129
- coupling 14–15, 22–5, 26*n*, 215, 252, 254
 - concurrence 255
 - resonance 255
 - synchrony 255
- crescendo 236, 240–1
 - see also* intensification
- decontextualised image 127, 129
- deixis 112
- discourse semantic systems 7, 11, 16, 17, 20, 25
 - see also* appraisal, negotiation
- dynamic vector 115
- embodiment 103, 115, 117, 119
- emotion 114, 116, 185–91
 - and learning 185–7
 - and reading 187–91
 - and scaffolding interaction cycles in parent-child reading 191–3
- empirical 110
- end-rhyme 230, 233–4, 236–41
 - see also* rhymes
- engagement 43–4, 46–9
 - heteroglossic 46–8
 - monoglossic 44
- esthetic 102, 109, 110, 111
- exchange structure 64–6
- experience 112, 113
- expression plane (expression form) 101, 104, 109, 245
- expressionist 108–9
 - expressionist-absolutist 110, 112
- face-to-face classrooms 31, 34, 48–9
 - face-to-face teaching 31, 34, 43
 - see also* pedagogy
- facial expression 31, 32, 44, 47
- force 235–6, 238–40
 - see also* intensification
- formalist 106, 108
 - formalist-absolutist 110
- framing 81, 82, 86, 126, 127–34, 141
- gaze 31–6, 38, 193
- gesture 31–6, 38, 41–2, 44–9, 53, 54
 - oscillation 47
 - prone hand position 46–8
 - supine hand position 46–7
 - see also* body language
- gloss 63
- graduation 13, 43–6, 49, 234–6
 - focus 45
 - force 45, 235–6, 238–40
 - intensity 45, 48
 - quantification 45, 48
- hierarchy of
 - instantiation 32–3
 - periodicity 134
 - realization 32
- humour 7, 12
 - convivial conversational humour 12, 15, 17, 20, 22, 23–5, 26*n*
 - humorous tension 14–16, 22–5
- ideational meaning
 - complementarity 147, 153–7
 - concurrence 147, 149–52, 161, 162
 - participant-process configuration 148, 150
 - see also* metafunctions
- identification 35–9, 48, 49
- ideological 101, 107
- idiolect 55
- image-text interaction 163
 - see also* image-text relations
- image-text relations 146, 157
 - augmentation 153, 154, 159, 161, 162
 - distribution 153, 155, 162
 - divergence 156
 - equivalence 149–51, 161, 162
 - exemplification 152
 - exposition 151–2

- image-text relations, modeling 146–7, 157, 164
 - cohesive relations 146, 147
 - inter-semiotic relations 146, 148, 162
 - intra-semiotic relations 148
 - topological meaning 162
 - typological meaning 162
- immanent (neutral) 109, 110, 111
- individuation 260–2, 264–5
- instantiation 251–2, 264–5
- intensification 235, 241
 - see also* crescendo
- intermodal
 - integration 126
 - relations 147
 - see also* image-text relations, coupling
- interpretation 101, 109, 110, 111
- inter-semiotic relations 146, 148, 162
 - see also* image-text relations, modelling
- intra-semiotic relations 148
 - see* image-text relations, modelling
- intonation 33, 43
- language, compared with music 101, 104, 111, 114, 115, 116
- laughter, ontogenetic development of 9–11
 - phylogenetic origins of 8–11
 - relation to smiling 8–9
 - sound potential of
 - see also* system network of laughter sound potential
- linguistic service 64
- listener 105, 109
- listening 103, 104, 105, 109, 119
- logico-semantic relations 146
 - elaboration 149–52, 157, 162
 - enhancement 157
 - extension 153–6, 157, 161
 - projection 157
- logogenesis 212–15, 236, 264
 - dynamic modelling 213
 - syntagmatic patterning 214–15
- meaning, extra-musical 109, 111
- meaning, intra-musical 111, 112
- meaning, of music 101, 102, 103, 106, 107, 108, 109, 111, 112, 113, 115, 116, 117
- meaning potential 105
- metafunctions 33–4, 40, 45, 48, 75, 101, 102, 248–9
 - ideational (and experiential) meaning 33, 41, 75–8, 125, 126, 149
- interpersonal meaning 7, 8, 9, 11, 12, 25, 33–4, 41, 43, 45, 48, 78–80, 125, 229, 234
- textual meaning 33–5, 39, 81–2, 125, 126, 141
 - see also* composition
- metaredundancy 246
- motion 114, 117
- movement 117–18, 119
- multimodal analysis 146–50
 - frames 148
 - multi-semiotic frames 148, 150
 - rank 147, 148, 247–8
 - units of analysis 148–50
 - visual and verbal elements 148, 158
- multimodal reading comprehension 145
 - integrative reading 160
 - reading comprehension 160, 162
 - see also* pedagogy
- multimodal texts 144, 175–6
 - hybrid texts; comics 146
- multimodality 9–11, 31, 75
 - modalities 31
 - modalities of learning 193, 194–5
 - modes 31, 75, 102, 174
 - see also* context
- music 101, 105, 106, 110, 114, 116, 117
 - absolute 101, 103
 - as activity 116
 - cognition 105
 - as language 108
 - notation 117
 - as object 116
 - structures 105, 111, 112–13
 - syntax 112
 - text 111, 112
 - works of 116, 117
- musical event 105, 117
- musicking 117

- negotiation 7, 11, 16–17, 22, 25
 - move 11, 14, 15, 16–17, 22, 23, 24, 25
 - speech function 15, 16
- non-linguistic semiotic behaviour 59
- non-semiotic behaviour 59
- normative 110, 111
- ontogenesis 264
- paralanguage 10, 11, 256–7
- parametric systems 258–60
- particulate structures 77–8
 - orbital 78
 - serial 77–8
- patterns, text 229–30
 - linguistic 232–3, 237
 - logogenetic 241
 - multimodal 230
- pedagogy 31, 48, 49, 193–202
 - initiate-response-feedback (IRF) 187
 - pedagogic genre 193
 - reading to learn 194–202
 - redesigning pedagogy 204–6
- performance 113
- phases 40, 43, 181, 188–9
- phenomenological 102
- phylogenesis 264
- poietic 102, 109, 110, 111
- pointing 35, 47, 48
 - delineation 38–9
 - directionality 35, 48
 - particularity 38
 - specificity 37–8
 - vectors 35–6
 - see also* identification
- position 40, 43
- posture 33, 43, 49
- protolanguage 9–11, 60
- rank *see* multimodal analysis, stratification
- rap music 233–5
 - virtuosity in 330, 333–4, 336, 339
- realisation *see* stratification
- referentialist 109–10
- repertoire 260
- repetition 230–2, 234–8
 - discourse semantic 235–6
 - paralinguistic 235–6
- representational meaning
 - analytical images 148
 - conceptual images 148
- reservoir 260
- rhyme 233–4, 236–7, 239–41
 - see also* end-rhyme
- rhythm 33, 43, 169–73
- semiotic behaviour 59
- sign 243–4
- signifiant 243–6
- signifié 243–6
- sign language 53
- social semiotics 31, 32, 48, 75, 102, 146
- solidarity 12, 24, 25
- somatic 263
- sound 101, 102, 103, 104, 106, 114, 115–16
 - materiality 101
 - in relationship 104, 105
 - semiotics of 104, 105–6
 - in space 104
 - in time 104
 - vocal 115–16
- speech disorder 53
- speech function 63–4
- spoken language 31, 41–2, 178, 179, 180, 181
- stratification 246–7, 264–5
- structure 250
 - types of structure 248–9
- supevenience 246
- symbolic attribute 129
- system 250
- systemic functional linguistics (SFL) 32, 43, 75
 - systemic functional semiotics 33
- system network 7, 17, 18, 19, 21, 106–7
 - of laughter sound potential 17–25
 - articulation system 18, 19
 - prosody system 18, 21
- text visualization 229–30, 232–4, 241
 - animated networks 223–6
 - overview 215–16

- streamgraphs 219–22
- text arcs 217–19
 - see also* arc diagram
- Theme 126, 134
 - hyper-Theme 134, 140
 - macro-Theme 134, 140
- tone groups 41
 - tonic 111
- transmodal communication 55
- values 11, 14–15, 22, 24, 25, 26ⁿ
- valeur 245
- visual representation 229, 231
 - see also* text visualization
- visual and verbal elements 158
 - see* multimodal analysis
- visual-verbal relations 147
 - see* image-text relations
- voice quality 8, 20, 21