Perspectives on multimodality

edited by Eija Ventola, Cassily Charles, Martin Kaltenbacher.

Document Design Companion Series

DOCUMENT DESIGN COMPANION SERIES accompanies the journal DOCUMENT DESIGN and focuses on internal and external communication of medium sized to multinational corporations, governmental bodies, non-profit organizations, as well as media, health care, educational and legal institutions, etc.

The series promotes works that combine aspects of (electronic) discourse — written, spoken and visual — with aspects of text quality (function, institutional setting, culture). They are problem driven, methodologically innovative, and focused on effectivity of communication. All manuscripts are peer reviewed.

DOCUMENT DESIGN is 'designed' for: information managers, researchers in discourse studies and organization studies; text analysts, and communication specialists.

Editors

Jan Renkema Tilburg University Maria Laura Pardo

Ruth Wodak

University of Buenos Aires

Austrian Academy of Sciences

Editorial Address

Jan Renkema
Tilburg University
Discourse Studies Group
P.O.Box 90153
NL 5000 LE TILBURG
The Netherlands
E-mail: I.Renkema@uvt.nl

Volume 6

Perspectives on Multimodality Edited by Eija Ventola, Cassily Charles and Martin Kaltenbacher

Perspectives on Multimodality

Edited by

Eija Ventola

Cassily Charles

Martin Kaltenbacher

University of Salzburg

江苏工业学院图书馆 藏 书 章

John Benjamins Publishing Company Amsterdam/Philadelphia



The paper used in this publication meets the minimum requirements of American National Standard for Information Sciences – Permanence of Paper for Printed Library Materials, ANSI 239.48-1984.

Library of Congress Cataloging-in-Publication Data

Perspectives on Multimodality / edited by Eija Ventola, Cassily Charles and Martin Kaltenbacher.

p. cm. (Document Design Companion Series, ISSN 1568–1963 ; v. 6) Includes bibliographical references and indexes.

1. Semiotics. 2. Modality (Linguistics) I. Ventola, Eija. II. Charles, Cassily. III. Kaltenbacher, Martin, 1966- IV. Series.

P99.4.M6P47 2004

401'.41--dc22

2004046238

ISBN 90 272 3206 7 (Eur.) / 1 58811 595 X (US) (Hb; alk. paper)

© 2004 – John Benjamins B.V.

No part of this book may be reproduced in any form, by print, photoprint, microfilm, or any other means, without written permission from the publisher.

John Benjamins Publishing Co. · P.O. Box 36224 · 1020 ME Amsterdam · The Netherlands John Benjamins North America · P.O. Box 27519 · Philadelphia PA 19118-0519 · USA

Table of contents

Contributors	, VII
Acknowledgements	IX
Introduction Eija Ventola, Cassily Charles and Martin Kaltenbacher	1
PART I. Multimodal issues	
Chapter 1	
In between modes: Language and image in printed media Hartmut Stöckl	9
Chapter 2	
Modelling multiple semiotic systems: The case of gesture and speech Peter Muntigl	31
Chapter 3	
Problematising 'semiotic resource' Victor Lim Fei	51
Chapter 4	
Multimodality and empiricism: Preparing for a corpus-based approach to the study of multimodal meaning-making John Bateman, Judy Delin and Renate Henschel	65
PART II. Analyses and applications	
Chapter 5	
On the effectiveness of mathematics Kay L. O'Halloran	91

CHAPTER 6 Multimodality in language teaching CD-ROMs Martin Kaltenbacher	119
CHAPTER 7 The multiple modes of Dirty Dancing: A cultural studies approach to multimodal discourse analysis Markus Rheindorf	137
Chapter 8 Multimodal text analysis and subtitling Christopher Taylor	153
CHAPTER 9 Multimodality in the translation of humour in comics Klaus Kaindl	173
CHAPTER 10 Multimodality in operation: Language and picture in a museum Andrea Hofinger and Eija Ventola	193
Chapter 11 Drawing on theories of inter-semiotic layering to analyse multimodality in medical self-counselling texts and hypertexts Eva Martha Eckkrammer	211
CHAPTER 12 On the multimodality of interpreting in medical briefings for informed	
consent: Using diagrams to impart knowledge Kristin Bührig	227
Index	243

Contributors

John Bateman Renate Henschel Fachbereich 10 Fachbereich 10

Sprach- und Literaturwissenschaften Sprach- und Literaturwissenschaften

Universität Bremen Bibliothekstr. GW2 D-28334 Bremen D-28334 Bremen Germany

bateman@uni-bremen.de

Kristin Bührig Lutterothstr. 5 20255 Hamburg Germany

Buehrig@pragmatiknetz.de

Cassily Charles Unit 5/2 Ben Boyd Rd.

Neutral Bay

NSW 2089 Australia ccha7549@bigpond.net.au

Judy Delin

School of Modern Languages

and Cultures University of Leeds LS2 9TT Leeds United Kingdom j.delin@leeds.ac.uk

Eva Martha Eckkrammer Universität Salzburg Fachbereich Romanistik Akademiestr. 24 5020 Salzburg Austria

eva.eckkrammer@sbg.ac.at

Universität Bremen Bibliothekstr. GW 2

Germany

rhenschel@uni-bremen.de

Andrea Hofinger Universität Salzburg Fachbereich Anglistik Akademiestr. 24 5020 Salzburg Austria

andrea_hofinger@hotmail.com

Klaus Kaindl

University of Vienna

Zentrum fuer Translationswissenschaft

Gymnasiumstr. 50 A 1190 Wien Austria

klaus.kaindl@univie.ac.at

Martin Kaltenbacher Universität Salzburg Fachbereich Anglistik Akademiestr, 24 5020 Salzburg Austria

Martin Kaltenbacher@sbg.ac.at

Victor Lim Fei

Blk 339 Jurong East Ave 1 #02-1534

Singapore 600339 volflain@hotmail.com

vIII Contributors

Peter Muntigl Universität Salzburg Fachbereich Anglistik Akademiestr. 24 5020 Salzburg Austria Peter.Muntigl@sbg.ac.at

Kay O'Halloran Department of English Language and Literature National University of Singapore Block AS5, 7 Arts Link Singapore 117570

Markus Rheindorf Stolberggasse 21/32 1050 Wien Austria markusr@chello.at

ellkoh@nus.edu.sg

Hartmut Stöckl Steinstraße 3 04275 Leipzig Germany hartmut.stoeckl@phil.tu-chemnitz.de

Christopher Taylor Centro Linguistico d'Ateneo via Filzi 14 34123 Trieste Italy taylor@sslmit.univ.trieste.it

Eija Ventola University of Helsinki Department of English P.O. Box 24 (Unioninkatu 40) FIN-00014 Helsinki Finland eija.ventola@helsinki.fi

比为试读,需要完整PDF请访问: www.ertongbook.com

Acknowledgements

We wish to thank Professor Heinrich Schmidinger, Rector of the University of Salzburg, for providing the funds to invite distinguished international scholars to Salzburg to discuss new perspectives on multimodality in 2002. The present book is a direct outcome of these fruitful discussions. Furthermore, we would like to thank the Stiftungs- und Förderungsgesellschaft der Paris-Lodron Universität Salzburg for providing the necessary financial means for the preparation of this publication. Finally, we thank Dr. Anja Schwarzl for formatting the manuscript.

Introduction

Eija Ventola, Cassily Charles and Martin Kaltenbacher University of Salzburg, Austria

For a growing number of researchers of text and discourse it has over the last decades become increasingly evident, that in the pursuit of understanding communication patterns around us, the analysis of language alone is not enough. The media of communication or, more accurately, the media which provide analysts with an object of study have for the last several millennia permitted a more or less comfortable excision of language, as text/discourse, from its context. Similarly, it has been relatively easy to analyse language and its instantiation in discourse separately from other forms of meaning-making, such as the gestures and facial expressions of a conversation, the illustrations and print of a document, the music and lighting of a drama. The first displacement, that of text from context, has been addressed, or redressed, by both applied and theoretical linguists in many traditions and for many purposes, particularly in the last half-century. This has been done either through building social context into the fundamental structure of linguistic theory (e.g. social semiotic theories of language), or through including contextual observation as a supplement to 'language-only' models. However, the second displacement, that of language from other kinds of meaning-making, has only recently begun to be addressed by linguists.

Although multimodality and multimediality, when seen as combinations of writing, speaking, visualisation, sounds, music, etc., have always been omnipresent in most of the communicative contexts in which humans engage, they have for a long time been ignored, as various academic disciplines have pursued their own research agendas as research fields. Thus, it is relatively recent that the developments of the various possibilities of combining communication modes in the 'new' media, like the computer and the Internet, have forced scholars to think about the particular characteristics of these modes and the way they semiotically function and combine in the modern discourse worlds.

Changes in the technology of communication are forcing the issue for those who would understand the mediation of our lives by discourse. Multimodality, the

interdependence of semiotic resources in text, is no longer so easy to overlook in today's media - particularly in written/visual communication. Indeed, the kind of text where the 'language' can be excised as an independent unit is, in many areas, fast becoming the exception rather than the norm. Accordingly, applications are putting increasing pressure on the development of multimodal theory. The multimodal literacy needs of students are stretching the ability of teachers to support them; information is proliferating in forms which push our methods of sharing it effectively; the shape of discourse communities is changing with the changing shape of texts; and ideological currents are flowing beyond existing linguistic means of analysis and critique.

Linguistic theories of text and discourse have in many respects by now been developed to such a degree of intricacy that the prospect of 'doing it over again' to account equally well for non-linguistic meaning-making is potentially rather daunting. Indeed, some multimodality research takes the approach that what works for language is likely to work for other kinds of meaning-making, and new models need only be adaptations of those already developed for language. Other research looks to expertise in fields outside linguistics - either to disciplines such as musicology or typography, which provide models of specific meaning-making resources other than language, or to such as performance studies and film theory, where language has always been treated together with other kinds of semiosis. Research into multimodality is therefore marked at this point by a broad degree of eclecticism.

This volume therefore represents one signpost on the several paths of multimodality research and theory-building today. The chapters collected here represent a cross-section of current perspectives on multimodal discourse. The book is organised into two parts. Part I focuses on theoretical and methodological issues. Part II covers a wide range of applications for multimodal description.

Part I - Multimodal Issues - starts with a chapter by Hartmut Stöckl, who sets out to develop a system network for modalities by exploring the semiotics of language and image in printed media. He posits all semiotic resources as consisting of core modes (e.g. language), medial variants (e.g. writing), peripheral modes (e.g. typography), sub-modes (e.g. colour) and features (e.g. saturation) and demonstrates that modes can be distinguished through their semiotic properties, their cognitive orientation and their semantic potential. Semiotically, writing operates via two-dimensional arbitrary graphic forms, whereas images mirror concrete three-dimensional objects of the real world. On the cognitive level language successively integrates signs into phrases, while images must be perceived simultaneously and holistically. Semantically language is based on convention and is therefore less polysemous than images, which can only be interpreted in combination with other modes or in a narrowly defined context.

In Chapter 2 Peter Muntigl argues that language and non-linguistic semiotic systems, like gesture, share a similar metafunctional and stratal organisation and are therefore equally grammatically structured. He maintains that gestures can be broken down into a content and an expression plane sharing categories such as Participant, Process or Manner. In addition to that, he claims that the mode of an interaction determines which semiotic systems are activated in a specific situational context, since the mode also determines the channel, medium and language role in a social interaction. Genres that are likely contexts for gestures are those that embody activity sequences and specific participants (stories, procedures, explanations, descriptions) as well as meanings of location or manner.

In Chapter 3 Victor Lim Fei examines the nature of signs within multimodality. Similarly to Muntigl, he argues that visual images operate on a content and an expression plane. Lim claims that the difference in the semiotics of language and visual images lies primarily in the degree of arbitrariness in the relation between signifier and signified. While language signs are arbitrarily chosen to represent their signifieds, visual images are based on a high degree of iconicity. He suggests that icons are the building blocks of visual images which are pieced together to bring across a coherent message. Like language expressions, icons are context and culture dependent, and the objects and ideas they represent may change according to the contextual and cultural background.

In the last chapter of Part I, Chapter 4, John Bateman, Judy Delin and Renate Henschel argue that many established claims about aspects of multimodal discourse are only informal and interpretative, due to the lack of supportive empirical evidence. As a solution they propose multimodal corpus analysis through tools that are specially designed for analytic research into multimodal meaning. They describe first attempts at designing a multimodal annotated corpus, called the GeM (Genre and Multimodality) Model. In this project documents are not only annotated with the well-established tags of traditional corpus linguistics but with new annotation procedures that provide information about the multimodal design and hierarchical structures of individual documents, such as the layout structure of sentences, headings, photos, drawings, figures, captions, text in pictures, icons, table cells, list headers, page numbers, footnotes, etc. Such information can function as an empirical basis for the investigation of the relationship between different document genres and their realizational formats.

Part II – Analyses and Applications – is dedicated to innovative multimodal analyses in a variety of contexts and disciplines. In Chapter 5 Kay O'Halloran investigates how semiotic change in the formal language of mathematics has helped this science to become so effective in describing the physical world. She demonstrates that mathematics originally developed as a multimodal discourse, where the visual played an essential role in establishing meaning in the process of solv-

and shows that it can serve as a useful basis for the selection of the verbal elements

that should appear in the subtitles.

ing a problem. During the early Renaissance mathematical problems were often visualised through drawn pictures and diagrams to demonstrate the problem but also to engage the students. In the 16th century abstract semiotic metaphors, like lines visualising the path of flying objects, established a significant turning point in mathematics. Symbols representing physical entities became more frequent particularly with the work of Descartes and Newton. Descartes used algebraic formulae for decontextualised representations of curves, and Newton defined scientific phenomena completely through the symbolism of algebraic equations. Modern mathematics is based on complex inter-semiotic relations between language, visual image (graphs, diagrams) and symbolism (algebraic formulae).

In Chapter 6 Martin Kaltenbacher looks into the semiotics of English language teaching CD-Roms and explores which demands text-image combinations have to fulfil in order to provide positive effects for the learners. He claims that many products combine different semiotic modes in a way that may inhibit rather than foster understanding and learning. He analyses the integration of sound waves in so-called pronunciation labs, the semantics of text-picture combinations, and the use of short video clips demonstrating the articulation of model speakers. All visualisations must meet some minimal semiotic requirements: they have to be easily interpretable, they must help the learners disambiguate the meanings to be learnt, and they have to be exact representations of the structures taught. These requirements are often not met, as many visualisations are too general or too ambiguous. As a solution to this, Kaltenbacher proposes replacing complex visualisations with more discrete ones, such as icons.

In Chapter 7 Markus Rheindorf investigates the relations of specific and nonspecific modes in the film Dirty Dancing. He explores the questions whether the distribution of the modes is significant to the genre dance film, how the modes combine to realise generic structure, and how the genre can be topologically related to other genres. Rheindorf argues that the protagonists of the film construe their class as well as their gender specific identities through the mode dance. Other modes, like dress code and music, support the semiotic functions of dance. Analysing a number of crucial scenes, Rheindorf suggests that set filmic phrases share common patterns and distributions of modes, which keep occurring in certain typical contexts with similar content. Additionally, he argues that the salience of certain filmic scenes is enhanced through a strategic change in their multimodal configurations.

In Chapter 8 Christopher Taylor looks at the role multimodal text analysis can play in the process of subtitling a film in a foreign language. A particular difficulty in this process arises when the spoken text contains a word play which is specific to the source language and is made visible in the film. Here the meaning encoded in the different modes may be lost in the translation process, unless the

In the next chapter, Chapter 9, Klaus Kaindl analyses multimodality in the translation of humour in comics, with an emphasis on the comic series Astérix and Tintin. He argues that most of the humour of comics is encoded in the pictures and not in the text. Kaindl identifies eight different strategies for the translation of humour, including deleting humour, changing the modality of humour and inventing absent humour in the translation process. In addition to that, he investigates the relation between verbal text and images in the creation of word and sign play. What creates a particular problem for translation is when a sign play depends on the presentation of an object that belongs to only one culture, e.g., a rugby ball. Furthermore, Kaindl addresses the phenomena of pictorial intertextuality and pictorial allusion and investigates what other modes, like typography, colour or onomatopoeia, add to the humour of comics and how they should be tackled in the translation process.

In Chapter 10 Andrea Hofinger and Eija Ventola analyse the semiotic processes involved in designing a museum on the life of Wolfgang Amadeus Mozart in Salzburg. Hofinger and Ventola study a room in the museum, concentrating on a metafunctional analysis of a painting of Mozart's family and the accompanying audio-text and the display of some furniture in the vicinity of the painting. They focus primarily on how the choices of different ideational, interpersonal and textual resources for the text interrelate to the semiotics foregrounded in the portrait and on what these meanings tell us about the socio-cultural context in which the museum is set up. They demonstrate how careful, or careless, selection of various meaning resources can create interesting and complementary, but sometimes even contrary effects for the overall semiotic of an exhibit.

In Chapter 11 Eva Martha Eckrammer develops a range of theoretical considerations on what she calls the "inter-semiotic layering" of multimodal texts, with a particular focus on medical self-counselling texts and hypertexts. She argues that the pictorial and the verbal elements of a text never develop their meaning separately but through a process of interaction between the semiotic layers involved, though one layer may bear more prominence in the meaning-making than the others. She suggests the following four dimensions of inter-semiotic layering: transposition (image is turned into verbal text or vice versa), juxtaposition (visual and verbal elements coexist without blending), combination (visual and verbal

elements are combined in a text), and fusion (image and verbal text fuse into a new textual form). In addition to that, Eckrammer argues that hypertexts should foster multimodality due to their non-linear structure, although in reality they are conservatively verbal.

In the final chapter, Chapter 12, Kristin Bührig investigates how visual materials, such as diagrams and charts, are integrated in the interpretation process at hospitals, where non-native speakers are briefed about their pending operations by German doctors with the help of a non-professional interpreter. One major obstacle in this type of communication is the linguistic barrier established through the use versus the lack of specialist language and expert knowledge by the parties involved. Bührig focuses on the different discursive roles a diagram plays in the doctor's attempt to communicate knowledge about the operation and the interpreter's attempt to pass on this knowledge in the target language. She finds that the professional doctor uses the visuals systematically to build up the patient's medical knowledge about his diagnosis and for refocusing the thematic elements in his discourse. The interpreter faces problems in transferring the medical information into the target language and uses the diagram to support the new information in the rhematic parts of the translation.

The impetus for this book were discussions which took place during the First International Symposium on Multimodal Discourse at the University of Salzburg in 2002, organised by the editors. We hope that this volume will be a means of opening up the discussion to others who will bring their interest and contribution to the rich and growing area of multimodal discourse.

Salzburg, 1st May, 2004 The editors: Eija Ventola, Cassily Charles and Martin Kaltenbacher PART I

Multimodal issues

CHAPTER 1

In between modes

Language and image in printed media

Hartmut Stöckl Technical University Chemnitz, Germany

The present chapter addresses theoretical and text-analytical issues of the language-image-link in printed media. After modelling multimodality as a networked system of core modes, medial variants, peripheral modes, sub-modes and features, the chapter goes on to characterise the linguistic and the pictorial mode according to semiotic, cognitive and semantic criteria. Central to this chapter is the idea that modes and sub-modes shift or blend (mode overlapping) and mix (mode mixing) in multimodal communicative events. This is demonstrated on two advertising texts sketching out the levels and criteria of analysis necessary to adequately describe language-image-links. The two sample texts show that there is a strong pictorial element in language and a linguistic element in images. The chapter concludes by looking at possible semiotic principles operating across modes.

1. Multimodality – the late discovery of the obvious

Whether as the reflection of a changing communicative landscape (i.e. stronger reliance on modes other than language) or a practical tool in text and discourse analysis, multimodality is currently gaining academic ground. Beyond its mere hype-aspects¹ fostered by an incessant quest for novel research topics, it addresses a phenomenon which is as old as representation itself and crucial to an understanding of almost all forms of communication: multimodal refers to communicative artefacts and processes which combine various sign systems (modes) and whose production and reception calls upon the communicators to semantically and formally interrelate all sign repertoires present. Spontaneous face-to-face talk relies heavily on non-verbal means (e.g. gestures, posture, body language) and has our visually perceptible environment as a constant topic of communication. Often it even has the immediate manipulation of objects as its target. Written language, on the other hand, incorporates images and – through typography and layout – wields strong pictorial powers. More recent media, like film, television or computer, mul-

tiply semiotic potentials by integrating moving images, language (spoken and written), sound and music. The same can already be said of older mediated experience like theatre, drama or opera. I would go as far as to argue that the purely mono-modal text has always been an exception while the core practice in communication has essentially been multimodal all along.² The dominance of linguistics, however, and the concentration on language as the central mode, paired with a lack of adequate models for the analysis of other modes, made verbal mono-modality appear to be the standard and dominant form of communication.

Considering that multimodality research has been around only for some 30 years,³ its achievements are quite remarkable. For one thing, research into multimodality has resulted in various 'grammars' of individual non-linguistic modes.⁴ For another, the social semiotic school as one major driving force behind multimodality research has succeeded in showing that 'phylogenetically' some modes seem to displace others in the development of media and genres. Social semiotics has also hinted at possible causes for this,⁵ and it has amply illustrated that the specific usage of one or the other mode is guided by socially determined intentions and realises group interests, subjective points of view or ideological stances. A third strand in multimodal research, namely studying the interrelations between various modes, is – as far as I can see – underrepresented. We seem to know more about the functioning of individual modes than about how they interact and are organised in text and discourse.⁶

In the present sketch I want, therefore, to turn towards the interface between two prominent modes, language and image in printed media. First of all, I would like to raise the question of what a mode actually is and how many there are. This involves the tricky business of neatly delineating various sign systems. I hope to show that there is a whole network of heavily interdependent modes and submodes and that in textual practice modes can shift and blend into one another (Sections 2 and 4). Secondly, I seek to demonstrate how language and image can be distinguished from one another looking at their semiotic structure, their semantic characteristics and the cognitive operations they typically entail (Section 3). Thirdly, analysing two sample texts from advertising and describing various facets of inter-modal relations, I will show that there is a pictorial element in language and a linguistic element in images (Sections 4 and 5). Finally, I would like to address the question of whether there are common semiotic principles operating across modes – as Kress and van Leeuwen (2001) have suggested – and what exactly they could be (Section 6).

2. How many modes is 'multi'?

'Mode' would seem to be an easily definable term. Following Jakobson's idea of 'code' (Jakobson 1971), it is often glossed as 'sign system', 'sign-repertoire' or 'semiotic' from which communicators can pick their signs to realise their communicative intentions. Apparently, we intuitively gauge the meaning of 'code' as the quality or type of signs used in a communicative event. Along with 'mode' comes the notion of a 'grammar', i.e. signs belonging to one mode are seen to be governed by a common set of rules that state how these signs can be combined to make meaning in particular situations. In practice, however, things turn out to be less straightforward than this. Let us look at modes a little more closely.

If we take sign quality to mean the different sensory channels via which signs can be perceived and meaning made by communicators, there is justification to subdivide visual, auditory (or audial, cf. Iedema 2003:32), tactile, olfactory and gustative sign(-mode). While this seems a valid and largely unchallengeable systematisation, it is too rough to be of much use to a theory of multimodality. Language or the linguistic code, for instance, which we justly feel to be a mode in its own right, would thus fall both into the visual and the auditory category as it can be medially realised as either speech or writing. Both are governed by the grammar of language, but as different medial or material varieties of one mode they entail a number of concomitant, additional sub-modes. So speech - besides being linguistic - also employs volume, intonation, timbre, rhythm, speed or pausing, all of which are design features of language in its spoken form and are often termed para-verbal. Furthermore, speech is accompanied and crucially shaped by what has come to be called the non-verbal mode, i.e. gesture, posture and body language. Similarly, writing as the visual counterpart of speech entails typography, which can be seen as the written variant of para-verbal means. What intonation, speed and rhythm are to speech, typography is to writing. However, to complicate matters, some aspects of typography like layout or paper quality are rather nonverbal as they seem further removed from language as such but still accompany writing and contribute to its meaning beyond the linguistic. Moreover, despite being an integral part of writing, the quality of a document's paper transcends the visual mode and has a clear tactile quality to it.

At least three points can be generalized from these brief observations. Firstly, modes cut across sensory channels, so the nature of a sign is not sufficiently characterised by looking at its path of perception. Secondly, one mode can be realised in different media thus creating medial variants of one mode (e.g. speech and writing as variants of the linguistic mode). As any one variant has distinct materiality, it in turn commands individual sets of concomitant sub-modes facilitating or accompanying the variant. This is, among others, a reason why media and

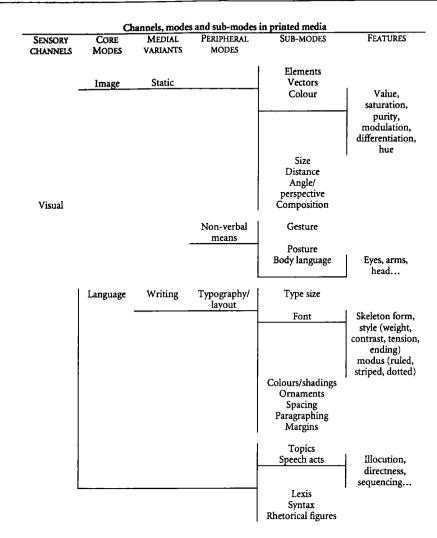


Figure 1. Network of modes, sub-modes and features in printed media

modes ought not be confused but neatly kept apart and regarded in their interdependences. Thirdly, and perhaps most importantly, the range of existing modes represents a hierarchically structured and networked system, in which any one mode can be seen to fall into sub-modes which in their turn consist of distinct features that make up the sub-mode. Let me exemplify these points looking at the image as another major signing mode.

Images, just like language, are not purely visual, they also have a tactile quality to them, which may be reflected in the meaning we construe from them. So while essentially visual, the material and techniques used in the production of the pic-

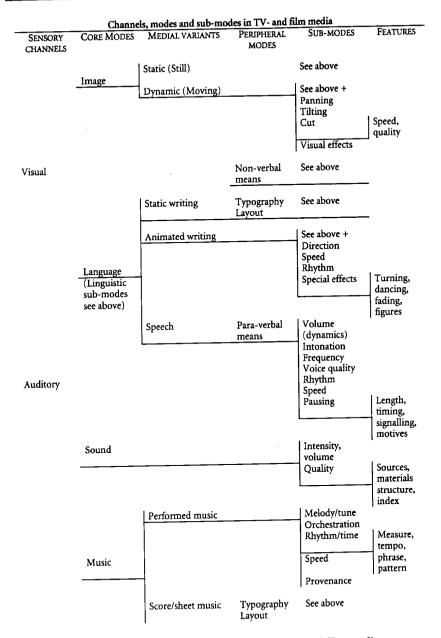


Figure 2. Network of modes, sub-modes and features in TV- and film-media

ture are also subject to touch. The nature of the pictorial sign can, however, not be gauged from its visual quality alone because this also pertains to written language, as we have seen. In this respect both image and language are equally visual. Images are realised in different media, the static printed image and the dynamic moving image of film and television being the most prominent examples. Besides sharing some sub-modes (e.g. elements, vectors, distance, angle, colour etc.), both medial variants also differ in the set of sub-modes they entail. So the moving image commands such specific sub-modes as, for instance, panning or tilting, it can have narrative⁷ and has at its disposal sub-modes originating in post-production like cut or visual effects (e.g. slow-motion, time-laps). Sub-modes in their turn can then be seen to comprise sets of distinctive features, i.e. specific aspects of one sub-mode, which are both phenomena in perception contributing to an overall gestalt as well as analytical categories which help to theoretically come to grips with sub-modes. Colour, for instance, can be decomposed into six distinctive features: value, saturation, purity, modulation, differentiation and hue (Kress & van Leeuwen 2002: 354ff.). Any concrete colour may then be specified by determining values on the scales of those distinctive features and their individual combination.

Applying this kind of systematisation of mode to textual genres in printed media and TV- or film-media, I suggest the following schematically represented network (cf. Figures 1 and 2).

The graphic representations of how modes are structured (cf. Figures 1 and 2) are necessarily formalised simplifications, which demand some comments on inherent problems and limitations.

- 1. While the columns of core modes contain central sign-repertoires that are deeply entrenched in people's popular perceptions of codes and communication and can stand on their own, peripheral modes come as inevitable 'by-products', as inherent elements of a core mode's specific medial realisation. This is not to say that core modes are more important than peripheral modes or more powerful in terms of their internal grammar and resulting communicative potential. Rather than being a major/minor distinction the differentiation indicates that some modes only come into being along with others and depend on them to some extent. Non-verbal means are ambivalent here as they can also function independently of language. On the one hand, then, the non-verbal is a concomitant aspect of language (cf. Müller 1998), on the other hand, it is part of a communicator's image and thus also relevant to pictorial analysis.
- 2. In another sense core modes are also abstract modes that need to be instantiated in a specific medial variant. The grammar of language must be realised either in speech or in writing it is only in these medial variants that peripheral modes pertain to them.
- 3. Sub-modes⁸ constitute a mode in that they provide the building blocks of a mode's grammar. It would be wrong, I believe, to view any of the sub-modes in isolation or as dominant in the make up of a mode. It is rather all sub-modes

in conjunction and their manifold interrelations that establish a mode and facilitate its realisation in a communicative event. As a gestalt in perception, the concrete materialisation of a mode in text and discourse is always more than the sum of its parts. No claim to completeness is made here, both for reasons of space as well as for the incomplete nature of relevant research.

- 4. If sub-modes are the building blocks of a mode's grammar, then features are the shapes of the blocks and the patterns from which they can be combined. Features can be conceived of as property scales on which certain values can be adjusted in the design of a communicative product/event. Again, it is the sum total of all features and their values that make up the specific realisation of a sub-mode. As they are far from always being binary oppositions, it seems advisable to dispense with the attribute 'distinctive'. For reasons of space I have confined myself to indicating features for only one sub-mode per mode. Within the features listed, again, completeness was not my objective.
- 5. It is my impression that some sub-modes, like colour, will allow for a relatively finely-grained differentiation into features, whereas others will only command few features or none. Those that consist of a larger variety of features seem to have a more flexible internal grammar and can due to their complexity function more autonomously in communication. Such sub-modes will therefore also be able to shift across modes. For instance, colour is not just relevant for pictures but also for typography. Still other sub-modes, like rhythm for instance, are so basic they pertain not just to music, but to speech, sound, animated writing and moving images. Generally, individual sub-modes can apply to more than one mode when the modes in question share some essential design feature, like in the example of rhythm a time-based structure, that is a linear sequencing as its underlying principle. Yet, it would be wrong, I feel, to grant such strong sub-modes the status of modes in their own right as they are always part of a whole configuration of sub-modes, without which the individual sub-mode, however strong, would not be able to operate. 9

でする ないない ないない 大きない

6. Most importantly, the text is the locus where all modes, sub-modes and features are realised. So it is the dynamics of text production and reception, the complex chain from discourse and design to production and distribution (Kress & van Leeuwen 2001: 1–23) that determines how we deploy modal resources and how they in turn are construed in reception. In textual practice one mode is often made salient while others recede into the background – both in perception and production. Genres or text types differ a great deal in how they foreground and background modes and sub-modes. Whereas typography is largely "automatized" (Iedema 2003: 40) in many text types, because it follows established conventions, advertising makes creative use of it and thus puts it centre-stage. Taking a textual perspective on modes and multimodality

means to also pay attention to "the socio-semiotic histories and transitions" (Iedema 2003:48) of texts, i.e. to uncover how in production and reception modes shift and blend and how producers and users translate or 'transpose' meaning from one mode to the other. In personal e-mail communication or in some types of advertising, for instance, texts are medially realised in writing, while conceptually they are oral in nature. So typically oral textual characteristics normally realised with the help of para-verbal means in speech may be transposed to the written medium using typography and its sub-modes.

Mode distinctions

When 'reading' a multimodal text, average recipients will normally become only dimly aware of the fact that they are processing information encoded in different modes. The manifold inter-modal connections that need to be made in order to understand a complex message distributed across various semiotics will go largely unnoticed. All modes, then, have become a single unified gestalt in perception, and it is our neurological and cognitive disposition for multimodal information processing that is responsible for this kind of ease in our handling of multimodal artefacts. A theory of multimodal communication, however, has to meticulously dissect an apparently homogeneous and holistic impression. It has to sensitise us for the essential differences of the modes involved and make us aware of the textual work we invest in building inter-modal relations so crucial to understanding. Multimodal theory also needs to ask in how far there are systematic similarities and ties between the modes involved.

In what follows, I want to demonstrate that it is at least on three different levels that modes can be distinguished from one another: semiotic properties, cognitive orientation, and semantic potential. While I will here confine myself to a juxtaposition of the two core modes language and image, 10 the three theoretical perspectives can be applied to all other modes and sub-modes.

3.1 Semiotic properties

The semiotic properties of a mode refer to its internal structure and to the general ways in which users can make meaning with a mode's signs. Language has what linguists call double articulation, i.e. discrete signs on two levels of organisation, phonemes and morphemes, which combine to form words and utterances. This design feature explains the boundless flexibility and resourcefulness of language. Images, in contrast, have no distinct signing units. There are no rules that would explain how pixels yield higher-level units when combined. What comes closest to a repertoire (alphabet) of pictorial signs is the elements that can be depicted in images and that we are familiar with from our visual experience of the world. The rules of how to combine them are tantamount to the natural order of things in worlds real or imagined. Consequently then, it is the iconic nature of signs that enables and determines their meaning. Because pictorial signs can access the same mental models that real-world objects access, we can understand pictures. It is the similarity between optical impressions gained from picture viewing and real-world vision that facilitates meaning-making in images. 11 Whereas images thus represent an analogue code, language is deeply symbolic and digital. We need to recode visual perception into abstract sounds or graphs that bear no resemblance to the objects depicted and vice versa. Although writing and pictures share some similarity, as both use the surface area of a medium for representation, they radically differ in that writing utilizes two-dimensional arbitrary graphic forms to represent speech sounds, whereas pictures systematically evoke the three-dimensionality experienced in the perception of objects.

3.2 Cognitive orientation

These and other semiotic properties result in different cognitive operations demanded or afforded by language and images. Most importantly, language is a linear mode that calls for the successive integration of signs into phrases, whereas images are rather based on simultaneous and holistic gestalt-perception. Consequently, images can be regarded as a quick mode relative to language as they do not necessitate parsing. We know from psychological experiments that images are far more likely to be attention-getters in perception than language and can also be memorised much more easily and effectively. Both have to do with their analogue code characteristics - no recoding needs to take place and pictures can therefore be regarded as a code close to reality or - as some semioticians have argued - a 'language' without a code. The speed of pictorial perception is usually put down to the simultaneity of gestalt formation, whereas the communicative impact of images is seen in the fact that they directly tap into the emotions and provide immediate sensory input.

3.3 Semantic potential

Semiotic and cognitive characteristics determine what users of a mode can do with it in terms of specific meaning-making resources. Although debatable, it has generally been accepted that the semantics of language is less vague and polysemous than that of the image. While language provides scope for double meaning, it has conventional semantics attached to words and utterances. Images, on the contrary,

are seen to be inherently vague and ambiguous and can only be made to mean and communicate specific contents by a combination with other modes or the embedding into narrowly defined communicative situations. Most importantly, images lack a definite speech act repertoire, which is why their illocutions remain cloudy unless they are complemented by language. Language, on the other hand, counts as less rich in information than images, which carry a welter of sensory information and are particularly intense in terms of connotation. Conversely, language is at a great advantage as for its potential to communicate all sensory modalities, whereas pictures clearly are confined to visual information. 12 Similarly, the self-referential capabilities of images are weak, whereas they are basically unlimited with language. Finally, language can be used to make just about any utterance imaginable. This huge semantic flexibility, which results from the linguistic design principle of double articulation and an elaborate set of rules, is contrasted by some obvious semantic restrictions of images. Some meaning relations like causality cannot be expressed, negation and affirmation are impossible and the utterances construable from images are usually additive. To sum up, language has its strength in the depiction of events and states-of-affairs in time, whereas images are particularly suited to the representation of objects in space and their physical characteristics (cf. Kress 1998:68f.).

Mode integration: Overlapping and mixing

The comparison between language and image showed up a number of essential differences, which shape the ways these two central signing modes can be made use of in communication. The danger inherent in contrasting two modes, however, is that we tend to somehow look at one mode in terms of another. So, mostly, due to language's dominance, we seem to be asking which linguistic properties images have. Thus we run the risk of overlooking some important design features of images which are outside the linguistic perspective. While a possible denigration of images resulting from this must be avoided, the metaphoric stance of a pictorial language is engrained in our naturally logo-centric take of communication.

More importantly, we need to realise that the mode distinctions pointed out are relative, as language and image are inseparably intertwined both in concrete forms of communication (mode mixing) as well as cognitively, semantically and historically (mode overlaps). Producing and understanding images presuppose verbal categorization of visual elements just as producing and understanding language relies heavily on conjuring up and manipulating mental images, the latter being most poignantly illustrated by the vast store of phrases that use a literal image to make metaphoric meaning. The history of alphabetic signs is a testimony

to the fact that the graphic articulation used to represent speech sounds has its origins and immediate precursors in the pictorial. The strongest argument for the innate tie between language and image, after all, is their co-presence in almost all forms of communication, a symbiotic mode integration (mixing) which is guided by the principle of reciprocally balancing out limitations and weaknesses of the modes combined.

The textual interface between language and image

There are two basic ways in which the linguistic and the pictorial mode can come together in a text. Firstly, a verbal text can itself acquire image qualities by means of typography and layout. In this case a peripheral mode (typography) of a medial variant (writing/language) is employed for a partial transfer from one core mode (language) to another (image). Here, the carrier of the linguistic mode emulates the pictorial. Secondly, and this is the more common option, a verbal text is combined with an image. The two core modes are semantically and formally integrated so that each mode strategically employs its range of sub-modes thus unfolding the specific semiotic potential of each mode and contributing to an overall communicative gestalt. A specific type of this language-image-combination would result if the verbal text contained language that was itself pictorial or figurative and established a semantic or formal link with the accompanying visual image. Let me illustrate both types of language-image-link by discussing two sample texts drawn from the advertising genre. My aim in this will be to show on which levels such interfaces between two modes can be analysed.

5.1 The typographic image

In an advertisement of the RSPCA (the Royal Society for the Prevention of Cruelty to Animals) for free-range eggs (cf. Figure 3) the verbal text is typographically designed to yield the visual form and appearance of a supermarket receipt. Although, of course, the language contained in the text is not what we would expect to read on a receipt, conjuring up the image of a receipt is possible, because this specific text type, like many others, comes with a built-in range of graphic features that can be imitated. Wehde (2000) calls such configurations of typographical and layout properties which form a set of visual expectations tied to a particular text type (format) a "typographisches Dispositiv", which could be rendered in English as the 'typographic repertoire' of a text type. Which typographic/layout sub-modes, then, have been employed to give the impression of the typographic repertoire of a receipt?

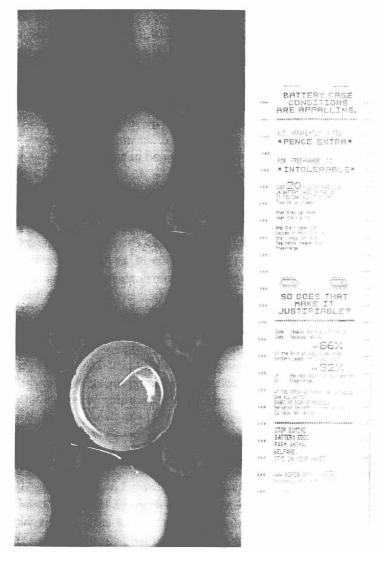


Figure 3. "A few pence extra", RSPCA (THE TIMES MAGAZINE 15.03.03, p. 10)

Battery conditions are appalling. But apparently, a few pence extra for free-range is intolerable. Over 20 million hens live in battery cages in the UK. If you can call it living. They're so crammed in that they can never even open their wings. And their bones can become brittle. But hey, their eggs can be a few pence cheaper than free-range. So does that make it justifiable? Some people obviously think it does. Because, while 86% of the British public say that battery cages are cruel, only 32% of the eggs sold in Britain are barn or free-range. At the RSPCA we think the UK

should ban all battery cages as soon as possible. Margaret Beckett is considering it. So help her decide. Stop buying battery eggs. Farm animal welfare. It's in your www.rspca.org.uk/eggs Registered charity number 219099

The visual image of a receipt is mainly formed by narrow margins, which have been marked by lines of three stars each. The text body is heavily subdivided into small portions of variously aligned print. This has been achieved by paragraphing and spacing as well as by lines formed of the same stars as at the margins. The font clearly betrays its provenance as having been produced by the typical cash-desk printer. This is signalled through the formation of the characters from individual dots, a system also used in digital displays. Contributing to this is the blue colour of the print as well as its irregular quality and the blurred print blotches in between, which by association also indicate the low but functional quality of the paper and the printing technology. The enumerative and strongly portioned character of the receipt also materialises through bold print, tab stops in the middle of lines, capitalisation and the use of numbers so typical of receipts. Besides being heavily paragraphed, receipts also come as parts of a continuous text (paper roll) from which they are torn when handed out to the customer. This continuous character of the text is borne out by the recurring logo on either side and the centrally aligned text body right underneath, which usually communicates the name and address of the supermarket or some other standing detail. Continuity is also expressed by the cut off logos at the top and bottom of the text.

When typographical repertoires are exported from one text type (receipt) to another (advertisement), the resulting effect is not merely pictorial as in our example, where the receipt is sort of reified as a textual object. More importantly, exported or emulated typographical repertoires have a semantic impact. In the RSPCA advertisement the receipt-like character of the text adds to the meaning of the verbal text. It supports the central argument of the ad, which says that purchasing behaviour can make all the difference in the battle for more free-range farming. The receipt image of the text makes the pivotal point that it is in the supermarket where farming policies are shaped via the price of the eggs and consumer behaviour. The readiness to overcome one's own meanness and spend a few pence extra on eggs as the target of the advertisement is suitably transported in the visual image of a receipt, as it directly taps into the knowledge script of shopping and thus reduces the mediated effect of communication and makes it more direct.

5.2 The verbal image and the language-image-link

Print advertising is a textual genre whose reliance on language-image-combinations is almost obligatory. In the design of such inter-modal links creators of ads exploit the whole range of semantic, formal and pragmatic relations, which can be



Figure 4. "No small fry", Toshiba (The Sunday Times Magazine 14.09.97, pp. 2f.)

anything from simple to complex.¹³ Linguistic text analysis, however, is invariably confronted with problems when it is to tackle such two-mode genres. Let me, therefore, suggest some levels and criteria of analysis suitable to shed light on the interface between language and image.

An advertisement for Toshiba laptops (cf. Figure 4) combines an image with a long copy, whose main feature is the accumulation of a number of metaphoric phrases (bold), which lend salience to the verbal text structure. Taken together, the figurative expressions and their core lexical elements (e.g. ocean, waves, shoal, swim, plunge; catch, small fry, haul, bait, canned) help to form two contiguous mental images or knowledge frames: the sea and fishing. The visual image (a half-open tin showing a keyboard inside and some computer features on the outside, like CD-compartment or connections for periphery devices) helps to focus the mental image of fishing and fish processing, but also links it to the main subject of the text, portable computing. The pivotal question we need to ask is how the two modes can come together in the make-up of the text and which pragma-semantic and cognitive operations underlie the meaning-making processes in this specific instance.

We've squeezed a desktop into a portable. No catch (1). The only difference between a Toshiba portable and a desktop is scale. When it comes to size and weight the two are oceans apart (2). And for power, function and features our portables are no small fry (3). Features like the best chips, Intel MMX, the largest screens, 13,3" and the fastest CD-ROM drives, mean our portables are a good haul (4). And the PC Card Slots on every Toshiba mean our portables can take the bait of most networks (5). Whilst our innovative firsts have always created waves (6) amongst desktop and portable manufacturers alike. It's because we would rather lead the shoal than swim with it (7), that makes us the world leader in portable computing. Our portables have consistently landed prizes from PC Magazine for service, reliability and technical innovation. If you'd like to take the plunge (8) and

see why more people are working wherever and whenever they like, call us on ... Hopefully you'll agree, Toshiba really do have portable computing canned (9)...

The first level to scrutinize is the semantic and pragmatic tie between language and image. It is obvious here that both the image and a whole network of linguistic expressions function metaphorically. The central metaphor could be spelt out as LAPTOP = FISH CAN (visual image) or BUILDING LAPTOPS = CATCHING FISH/CANNING FISH (verbal text). Whereas the visual image provides the overriding metaphor which serves as a frame of orientation for the interpretation of the verbal metaphors, the literal meanings of the metaphorical phrases (in italics) detail and structure the emerging mental images (sea, fishing). The result of this kind of visual-verbal metaphorical play is the mental mapping of source domain features (sea, fishing) onto target domain features (computer manufacturing). So here analogies are built between

- (1) the sea and the market (oceans apart, create waves, take the plunge)
- (2) the quality/quantity of the catch and the quality of the computers (no small fry, good haul)
- (3) fish processing (canning) and quality computer assembly (have portable computing canned)
- (4) fish/fishing and computer firms (rather lead the shoal than swim with it, no catch)
- (5) instruments of fishing (bait) and compatibility with networking standards (take the bait from most networks).

On a second level, we could ask which cognitive operations are afforded by the design of the language-image-link. Clearly, what is intended here is the oscillation between literal and metaphoric meaning-making throughout the text, but also the successive integration of verbal phrases into a whole. On the one hand, all figurative expressions can be read literally (i.e. with recourse to sea and fishing), an interpretation facilitated mainly by the visual image, which provides the concept of a can with all its physical properties. On the other hand, of course, the phrases can be decoded in their metaphorical meanings, which support the persuasive intention of the text. Both on the literal and metaphorical plane of meaning the phrases combine to build a structured network. The network's elements cohere because on a literal level they build paradigmatic sense relations and all add to a common mental image. On a metaphorical level the phrases are networked as they all contribute to realising typical advertising speech acts and establishing the advertising text pattern. So, no small fry and good haul describe and evaluate the advertised product, created waves, rather lead the shoal than swim with it and have portable