The Routledge Handbook of Multimodal Analysis


Reviewed by Charles Forceville
University of Amsterdam
Dept. of Media Studies

Carey Jewitt’s *Handbook* consists of a general introduction, followed by 22 chapters divided over four parts, each with its own introduction. All chapters end with a short “suggested reading” section. The references are collected in one big bibliography. There is a subject index with some items in bold; these latter (some 160 entries) are explained separately in a glossary. The book contains a lot of text; while being well-readable as it is, it could easily have been produced as a 500+ page book.

The designation “Handbook” evokes expectations of a fairly complete map of a discipline’s territory – but the word is subject to inflation now almost every self-respecting publisher has a handbook series. Rather than a map, the book is a collection of field notes. However, no reliable map was ever made without field notes, and the current collection of essays is a robust step forward in transforming “multimodality” into a serious discipline. Jewitt is well-placed to be its editor: she has substantially published on visuality and
multimodality (including Van Leeuwen & Jewitt 2001), and is managing editor, also with van Leeuwen, of the journal Visual Communication. What follows is a review by someone who both enjoys the benefits and suffers the biases of not being affiliated to the central paradigms informing the Handbook (see also Forceville 1999, 2007, 2009a). I will first briefly characterize each chapter, then list some positive trends, and end with a discussion of problems and suggestions for how to remedy them in future research.

In her introduction Jewitt proposes: “multimodality approaches representation, communication and interaction as something more than language.” Modalities, or modes, are “semiotic resources for making meaning that are employed in a culture – such as image, writing, gesture, gaze, speech, posture” (p. 1). The term “multimodality” is “strongly linked with three perspectives on representation and communication” (p. 2). These are Hallidayan social semiotics/systemic functional grammar (SFG) as elaborated and practised by Kress & Van Leeuwen; discourse analysis as exemplified by O’Toole, Baldry & Thibault, and O’Halloran; and interactional analysis, associated with work by Scollon and Norris.

The chapters in Part I, the first two by Jewitt, address theoretical and methodological issues, and aim “to provide the basis of a ‘research toolkit’ for multimodal analysis” (p. 5). This is no easy task since, as Jewitt acknowledges, new modes constantly enter the communicational landscape, while there is no agreed-upon definition of “mode.” Moreover, what counts as a mode is “inextricably shaped and construed by social, cultural, and historical factors” (p. 22). But given the modes labelled as such in the current book, researching relations across and between modes is crucial. Jewitt is aware of problems pertaining to multimodality research. One is that multimodal analysis may seem impressionistic. This, she implies, is inevitable because semiotic resources are “contextual, fluid and flexible” (p. 26). A related criticism is that multimodality “imports and imposes linguistic terms on everything.” But such criticisms “overlook the fact that much of the work on multimodality has its origins
in a particular strand of linguistics: namely, the social semiotic theory of communication first proposed by Halliday” (Ibidem). The latter indeed explains how a degree of tunnel vision arose, but of course does not justify it. Finally, Jewitt agrees that while detailed description is a sine qua non of multimodal analysis, researchers need to remain aware that description is always drive by theory and research purposes. She concludes: “The development of multimodal corpora may help to overcome some of these limitations, as might the potential to combine multimodal analysis with quantitative analysis in innovative ways” (p. 27).

Flewitt et al. usefully address both the necessity and the difficulties of having to transcribe multimodal data in language in chapter 3. On the one hand, verbal description is indispensable to ensure rigour, systematicity, and precision; on the other, the fact that non-verbal and verbal modes do different things well means that verbal transcription is necessarily imprecise. The authors evaluate a few attempts at such transcription. They point out that it is important that annotators are clear about what research question a transcription is supposed to answer, not least because it is a labour-intensive activity, and that the two modalities of the typical research paper (static pictures and written language) are often inadequate to do justice to the complexity of multimodal discourse. In chapter 4, Kress engages the vexed question what is a mode, clarifying and refining earlier analyses (Kress and Van Leeuwen 1996/2006), much of it in line with Jewitt’s opening chapters. His pedagogic interests result in an analysis of how verbal and visual information are presented in a school book section on blood circulation, and in students’ own reworkings of such information. Van Leeuwen in his chapter continues earlier fascinations with sound (1999) by discussing how dimensions of voice quality (pitch, loudness, timbre, articulation, resonance, accents) can be a semiotic resource. Marlon Brando’s “godfather” voice, for instance, exudes harshness, menace, as well as intimacy. He proposes an analysis in terms of a “parametric system” – a series of graded oppositions (tense-lax, loud-soft, vibrato-plain, …) that together characterise voice quality. In
Chapter 6, Norris introduces the concepts “modal density” and “modal configurations.” Modal density increases if/when more different modes are involved in a communicative action; modal configurations refer to their relative importance for its success (a gesture may be less essential than the spoken sentence it accompanies). I agree that these are useful concepts (cf. Forceville 2009b). However, in her sample analysis, I think Norris too easily conflates intentional and non-intentional information – an issue to which I shall return below.

The chapters in Part II will illuminate five significant factors mediating multimodality: “history and timescales, the facilities of technology and their usage, social and culture factors, identity, and literacy” (p. 95). O’Halloran addresses how mathematical symbolic notations have developed historically, under the influence both of technology and the organizations and institutes in which the notations have been deployed. Dimensions include the shift from finger-counting (the basis of the decimal system) via notched bones and clay tokens to written numerical systems, the latter fulfilling institutional needs to keep records of mathematical operations. O’Halloran concludes with a section on how computerization has had an impact on mathematical knowledge. Jones’ chapter on the affordances and constraints of a specific social network site’s format aimed at gay men exemplifies a number of qualities of good multimodal analysis. Jones first discusses how newspaper noticeboards in Chinese parks have traditionally been meeting places for cruising gay men. He then notes how using another “site of display” (here: a personal profile on a Hong Kong gay website) offers new opportunities for action, and loses others. Analysing the categories and choices the format provides for creating a profile (length, weight, colour etc. as well as photographs, “fan” tokens, and other possibilities for interactivity), Jones demonstrates how a discourse using two modalities (text and pictures) on a specific “site of display” (a website) in a discourse whose purpose is crystal clear (attracting the sexual attention of other men) co-determines the projection of identity. Such a clear research design invites other scholars to extend work on identity projection by
using other variables: other modes, other gay sites, heterosexual dating sites, other corpora of “sites of display.” Leander & Vasudevan’s “multimodality and mobile culture” chapter describes three instances of how new technologies affect multimodal communication and identity formation. The discussions of the examples – a cell phone exchange, an online game, a Myspace profile – are mildly interesting but do not exceed the descriptive level. Lemke’s chapter is a thoughtful essay inspired by phenomenological approaches on how our degree of temporal control over media consumption and our locus of access (compare seeing a film on your own DVD player, allowing for rewinding, slow-motion viewing, freeze-framing etc. versus seeing the same film in a cinema) affect our experience of these media, and our opportunities for identification. Encounters with many media, moreover, encourage “traversals” (p. 145) to other, related media (official URLs, fan-fiction sites, spin-off games, merchandise). “Global producers market identities” (p. 150), Lemke concludes, and this calls for critical analysis of underlying power relations. The last chapter in Part II, by Unsworth and Cléirigh, zooms in on word-image interaction in pedagogical texts. Considerably refining Barthes’ (1986/1964) notions of “anchoring” and “relay,” they propose various ways in which, given each mode’s potential for creating meaning, written texts can complement or qualify the images they accompany and vice versa, and then discuss one schoolbook example where the interaction between the modes is badly organized. Again, because of the clear focus this chapter truly provides tools that can be used, tested, and if necessary refined.

Part III gives a platform to approaches to multimodality beyond SFG, social semiotics, and discourse studies. Scollon and Scollon briefly sketch the research trend from monomodality (= language) to multimodality. They emphasize particularly the spatio-temporal dimensions of language use, highlighting how the cycles the human body experiences (from heart-beats to seasons) are to be taken into account in meaning-giving processes. Somewhat alarmingly, these authors want to accommodate yet another meaning of
“modality” in multimodal research: the clues that reveal a person’s stance toward a linguistic statement (cf “I want to/should/might/can eat dinner”). Machin’s chapter considers various criticisms of multimodality research, in the service of bolstering it. A central question is what counts as a sign. This is especially problematic for what Peirce calls “icons” – signs that communicate via resemblance, such as photographs. Even if they are considered signs, how, if at all, is interpretive freedom constrained? Specific criticisms by Forceville (1999a) and Bateman et al. (2004) of Kress & Van Leeuwen’s social semiotics approach are addressed, although somewhat euphemistically. Machin for instance fails to mention my fundamental objection that often it is specifically these authors’ eagerness to conduct a critical, socially engaged kind of scholarship that tends to prejudice their interpretations (see also Bateman 2008: 44). Street et al.’s chapter aims to show how the traditions of multimodality research and “New Literacy Studies” (broadly defined as “an approach to literacy learning as a social practice – shaped by context, texts and practices used to make meaning” in the glossary) intersect, focusing on ethnographic perspectives as the area of overlap. Ivarsson et al. (chapter 15) illustrate the claim that writing, images, and other semiotic resources provide both tools to externalize experience and to communicate about it. They discuss multimodal interactions between people playing interactive games and between a teacher and a student about the latter’s architectural design. Their conclusions – that representational tools are interconnected; that meaning-making is always relative to social practices; that language use in situated activities deserves more research – are uncontroversial. Luff et al. also “go 3D,” specifically linking multimodality to “workplace studies.” The authors are doubtful about opportunities for synergy, however: “the term ‘multimodal’ provides perhaps a useful contrast to the ‘linguistic,’ but it may prove an unfortunate way of characterising communication and serve to impoverish analytic developments” (p. 222), not least because the various modes (images, gestures, gazes, speech) are often so difficult to separate from each other. The anthropologist
Howes further complicates matters in his chapter by showing how the Desana Indians experience their world in a synaesthetic way: “an odour will bring to mind a clour; a colour, an odour, along, perhaps, with a temperature and a vibration” (229), arguing that this is a culturally, not neurologically governed form of synaesthesia. After some brief remarks on the other less theorized senses, taste, and touch, the author concludes, “There are many potentially fruitful directions for further research on multimodality in anthropology, but he first task must be the production of detailed case studies of the sensory vocabularies and practices of as many of the world’s cultures as possible” (p. 234; see also Plümacher and Holz 2007).

The chapters in Part IV are case studies. Björkvall’s data are IKEA tables. He proposes as their most important variables whether one can put one’s knees under them (“intense” tables) or not (“restrained tables”), and the forms and materials of the table tops. Partly on the basis of some interviews with buyers as well as IKEA personnel he discusses issues such as that round tables allow for a more democratic seating arrangement than rectangular ones, as well as that aesthetic preferences for types may be culturally determined.

With Jaworski and Thurlow’s chapter we move outside: they video-recorded three hours of tourists gesturing (particularly pointing) and moving on the Tower of Pisa campo. The authors underline the mediated and embodied nature of the tourist experience, and argue that the meaning of gesture and movement interacting with space should not be studied as mere “context for talk” (p. 255), but deserves research in its own right. The authors’ indebtedness to social semiotics’ engaged character surfaces in their claim that “ultimately, the practices of tourism – whether verbal or nonverbal – reinscribe an ideology of conquest, of control and of posssession” (Ibid.). Mavers’ chapter persuasively suggests how school children’s verbo-visual “maps” of events are good pedagogic sources of information, and allow for more or less systematic analysis. For this to function, however, the map-making must be governed by
explicit tasks and questions, and the analysis must be supported by monitoring of the children’s creations as well as their post-hoc explanations. Stenglin continues her earlier research on the role of museum space, and trajectories (“path”) and goals (“venue”) within that space, by analysing a multimedia installation of the artist Ellen Dreyfus in a gallery exhibition. Central variables are “prominence” (pertaining to elements that attract visitors’ attention and therefore guide their paths) and degrees of (pleasant or uncomfortable) boundedness – or lack thereof. Inevitably, these elements are often recruited to negotiate power and social distance; but they can also be used for informing the design of 3D spaces for display. The last chapter, by West, is a treasure-trove of examples of how sound and music can be used and manipulated to create meaning in film, design, and musical education, and contains a sample analysis of the latter. Five functions of speech-music-gesture communication (testing/inquiring, instructional, analytical, accompanying, and expressive) were used to code every message unit in student-pupil interactions, with the aid of software, in order “to search for emerging patterns” (p. 291).

A number of shared ideas & trends emerge from the book, helping to build the discipline. Here is a staccato list: purely monomodal (mostly: verbal) communication is the exception rather than the rule. Since spoken and written language are the best-theorized modes of communication they provide stimulating ideas for theorizing non-verbal communication, but it is crucial to be aware that each mode has its own “affordances and constraints” for communication. Where necessary, mode-specific tools for analysis must be, and are, developed. While the insistence on exposing ideology features prominently in various chapters, the idea that a relatively unbiased search for potentially meaning-carrying variables in multimodal discourse is a good starting point for analysis seems to be gaining ground. There is a growing awareness that theorizing multimodality requires charting what, in a given discourse, are the possible choices that makers have at their disposal, both within and across
modes, to create meaning – and analysing which choices were actually made. The need for, and the difficulties of, verbal transcription of non-verbal communication are acknowledged. Video-recording becomes a more common tool aiding preciseness and systematicity in transcription, but the standard scholarly-article-on paper is often felt to be inadequate for reporting; clearly, in this respect there is a great future for online (versions of) journals that can accommodate moving images as well as sound and music. Many chapters discuss the centrality of the human body and of spatio-temporal dimensions in meaning-making, which allows for building bridges with cognitive linguistic models (e.g. Lakoff and Johnson 1980).

All this agreement is a reason to rejoice, but there are also less happy tendencies to be detected in the chapters. The most worrying of these are that there is no clear agreement on what should count as a “modality” or “mode,” and that it is often difficult to separate the analysis of modes (e.g., gaze, gesture, movement, and talk in face-to-face conversation). Add to this the correct insight that virtually all communication is multimodal, and that (since multimodal discourse is dynamic) context is always crucial for analysis, and we are dangerously close to having to conclude that systematic research on multimodality is a *contradictio in terminis*: there might simply never be enough stability to allow for generalizations.

This brings me to another critical point. If multimodality is everywhere and thus infinite, it is difficult to complain that there are things missing from this *Handbook*. Nonetheless, I find it odd and unfortunate that despite occasional references, not a single chapter (although Stenglin’s comes close) is devoted to an *artistic* genre or medium. This may be a sign of times in which humanities departments are under increasing pressure to teach their students “useful things” (i.e., things that allow them to earn money as soon as possible) while “all art,” as Oscar Wilde pointed out long ago “is quite useless.” It would be sad if the shift toward multimodality were to result in a further diminishing of interest in, and
knowledge of, artistic discourse. Art deserves academic attention for its own sake, but apart from that, sustained research of artistic discourse offers profound insights into communication, culture, and cognition – including their multimodal aspects. Specifically, I find it surprising that no chapter has been devoted to film, which is a multimodal medium *par excellence*, and has yielded a library full of pertinent texts pertaining to multimodality (although not under that label). Film (and TV and theatre, for that matter) offers numerous ready-made specimens in which the interaction between language, image, sound, music, movement, gesture, posture, gaze etc. can be studied in great detail *ad libitum* either to better understand the medium itself or as a training ground for the analysis of multimodality “in the wild.” An added bonus is that one can increasingly manipulate the role of certain modalities (e.g. via language menus on DVDs, or by switching off soundtracks, or by reducing the quality of visual information via software programs) to discover more about “modal density” and “modal configurations” (Norris). Similar arguments hold for the rapidly evolving scholarly study of comics, and of cartography. I am also disappointed that there are so few references to the discipline of narratology (e.g., Bal 2009, Bordwell 1995, Ryan 2004, Chatman 1992), let alone a chapter devoted to it. It seems a waste to ignore decades of robust and exciting scholarship that has so much to contribute to the theorizing of multimodality.

My concern is that there is a serious chance that this handbook, the first of its kind, will be unproblematically used as … well, a handbook, in courses that have “multimodality” in their title – and I foresee that their number will rapidly grow. The risk is that future students will be unaware of pertinent humanities disciplines and approaches that are not, or barely, represented in this *Handbook*. Let me end by putting in my pennies’ worth to help focus the discipline.

(1) It is crucial to distinguish between, on the one hand, phenomena that are intentionally used by a person or institution to transfer more or less specific information to a
more or less specific audience and, on the other hand, phenomena that carry information symptomatically (cf. Bordwell & Thompson 1997: 76), that is, willy-nilly – think of blushing, or a Freudian slip. I submit that it helps enormously to consider only the former as “communication,” although I am well aware (a) that we here have a continuum, not an either/or categorization; and (b) that symptomatic behaviour is worthy of research. More generally formulated, multimodal research needs to be embedded in a theory of communication; only then can issues whether something is, or is not, a “sign,” and where freedom of interpretation begins and ends, be resolved. My candidate is Relevance Theory (Sperber and Wilson 1995; see also Forceville 1996: chapter 5, 2005, 2009c; Yus 2008).

(2) There are substantial differences between multimodal discourses in 2D and in 3D, and between time-based and non-time based discourses. By and large, the number of modes in 3D and time based media is larger than in 2D and in non-time based discourses, and so is the complexity of studying their interaction. Theorists of multimodality may do well to focus their attention more on relatively simple discourses in order to develop and refine tools; even if this entails, for practical reasons, temporarily excluding pertinent modes from analysis.

(3) I fully support Howes’ call for more case studies. If we want to develop and refine “tools for analysis,” I suggest we do so by systematically analysing corpora of discourses (1) belonging to the same genre; (2) communicated in the same medium; (3) drawing on the same combinations of modes; (4) in light of a clearly formulated research question. Fine examples of this are Jones’ and Unsworth & Cléirigh’s chapters in this volume and Bateman (2008). Only such “piecemeal theorizing” (Carroll 1996: 58) allows for generalizations that help the discipline come into its own.

The Handbook provides much exciting food for thought and shows that useful tools and promising approaches are being developed. But more than anything else it makes clear
the need for further subdivision of work, for systematic rigour, and for many, many more corpus-based case studies.

References


