

## ePub<sup>WU</sup> Institutional Repository

Dennis Jancsary and Markus Höllerer and Renate Meyer

Critical analysis of visual and multimodal texts

Book Section (Accepted for Publication)  
(Refereed)

*Original Citation:*

Jancsary, Dennis and Höllerer, Markus and Meyer, Renate (2016) Critical analysis of visual and multimodal texts. In: *Methods of Critical Discourse Studies*. 3rd., SAGE, London. pp. 180-204. ISBN 978-1-4462-8241-3

This version is available at: <http://epub.wu.ac.at/6126/>

Available in ePub<sup>WU</sup>: March 2018

ePub<sup>WU</sup>, the institutional repository of the WU Vienna University of Economics and Business, is provided by the University Library and the IT-Services. The aim is to enable open access to the scholarly output of the WU.

This document is the version accepted for publication and — in case of peer review — incorporates referee comments. It is a verbatim copy of the publisher version.

# CRITICAL ANALYSIS OF VISUAL AND MULTIMODAL TEXTS

*Dennis Jancsary*

*WU Vienna University of Economics and Business  
& Copenhagen Business School, Denmark*

*Markus A. Höllerer*

*WU Vienna University of Economics and Business, Austria  
& UNSW Australia Business School, Australia*

*Renate E. Meyer*

*WU Vienna University of Economics and Business, Austria  
& Copenhagen Business School, Denmark*

This is a draft chapter. The final version is available as Chapter 8 in *Methods of Critical Discourse Studies*, 3<sup>rd</sup> edition, edited by Ruth Wodak and Michael Meyer, 2016, SAGE Publications, pp. 180-204.

## **Keywords**

Multimodality, critical discourse analysis (CDA), semiotics, visuals, visibility, visual discourse, visual analysis, sociology of knowledge, meaning reconstruction, qualitative research methods

## Introduction

This chapter is dedicated to the critical analysis of multimodal texts – i.e., texts that incorporate semiotic resources beyond verbal language. Our focus here is, in particular, on the relationship between the verbal and the visual mode of communication. In a first section, we touch briefly on the ubiquity of multimodality in contemporary society and provide definitions of core concepts. We then continue, in a second section, to systematically develop an argument for the relevance of multimodality within critical discourse analysis (CDA). The third section summarizes a number of exemplary studies that have adopted a critical approach in multimodal discourse analysis. These studies cover a variety of issues and areas of scholarly inquiry, and therefore aptly demonstrate that multimodal CDA can be applied in various forms. In the fourth section, we introduce – as a practical example – one particular methodological approach in more detail, and illustrate its various analytical steps on the basis of two selected multimodal texts. We close with a brief reflection and some concluding remarks.

### What is multimodal discourse?

Imagine yourself sitting in a Sydney beachside café on a glorious morning in 2012, browsing through a pile of newspapers. You grab the weekend edition of the *Australian Financial Review*. What you see on the front page is an article that explores how the Euro crisis has started to show repercussions in Australia. The article further expands on how the fear of the Chinese economy taking a hit has led to plummeting shares in Australia, and then continues to discuss related events in more detail, also providing concrete figures. But those facts – even if substantial – might not be what caught your attention in the first place. Indeed, the very first thing you most probably noticed was a large photograph of an experienced businessman, staring at a screen with utter bewilderment, one hand pressed close to his temple in a gesture of disbelief; this striking picture is complemented by a series of abstract graphs showing decreasing trends in different economic variables. If you had to tell a friend later what the article was about, would you still provide the same overall narrative, even if it did not contain such imagery?

Discourse studies display, by their nature, a strong affinity to language. Language, after all, is the most prominent resource for the social construction of reality and the storage of social knowledge (Berger and Luckmann 1967). Unfortunately, this also means that in actual analysis, researchers often focus on written and spoken verbal text, and ignore, or at least downplay, the importance of other information. Multimodal analysis (e.g. Kress and van Leeuwen 2001; Kress 2010) aims at addressing this shortcoming in existing research, and acknowledges the multitude of different

materials and ‘meaning resources’ that people use to create and distribute meaningful signs. Over the last decade, multimodal discourse analysis has gained considerable momentum, resulting in a number of edited volumes on the subject (e.g. Jewitt, 2009; LeVine and Scollon 2004; O’Halloran 2004; Royce and Bowcher 2007). According to Kress (2010: 79 [original emphasis]), ‘*mode* is a socially shaped and culturally given semiotic resource for making meaning. *Image, writing, layout, music, gesture, speech, moving image, soundtrack and 3D objects* are examples of modes used in representation and communication’. Note that Kress uses the term ‘mode’ in the sense of a ‘resource’, something to be employed in order to create meaning. How a particular act of communication is created, then, depends on which resources are available and regarded as appropriate in a specific social situation. Within a particular cultural domain, similar meanings can be expressed in different modes (e.g. Kress and van Leeuwen 2001). But not everything is possible, or appropriate, all of the time – the ‘pool’ of available resources as well as their meaning potential are culturally restricted. Imagine, for instance, ‘modal taboos’: some religions explicitly prohibit the depiction of particular subjects. History has seen quite a few instances of *iconoclasm*, the intentional destruction of visual references to particular people, events, values, or beliefs. But we do not even have to go that far. Think of different genres of books. Children’s books are usually heavily illustrated, to the point where *visuality* is the dominant mode, and written text becomes secondary (e.g. Kress and van Leeuwen 2006). In contrast, legal codices are mostly verbal, and even use a particular language quite different from everyday speech and writing, as well as typography in the form of clauses and articles. The reverse – photographs in legal codices or clauses in a children’s book – is unlikely to be seen.

We may therefore conclude that multimodality is governed by cultural and institutional rules – norms, conventions, and guidelines that tell us what is adequate, and what is not. Within these boundaries, people have considerable leeway with regard to the presentation of their messages. For instance, if I wish to describe to a friend the new car I have just bought, I can give her a description including more general details (e.g. type, brand, colour, or interior design features) and technical specifications (e.g. engine performance, fuel consumption, hybrid drive, or exhaust system). Of course, I can also show her a picture so that she can actually *see* it, which might give her a more immediate and holistic impression, but at the same time omits information that is not visible as such (e.g. fuel consumption). Maybe, however, a verbal or visual description is not at all what I had in mind. If I am, for instance, particularly enthusiastic about the noiseless electric power unit (or, alternatively, the sound of its sporty engine), I might want my friend to *hear* it, and invite her to join me for a ride in my new car. Still more intimate is the desire to communicate how driving my

new car *feels* (‘I can’t describe it – you have to experience it yourself’), which would involve the experience of touching it and actually sitting behind the steering wheel. All these modes communicate different aspects of the idea of what ‘my new car’ is about. I can combine some of them in my presentation to harness their particular strengths and give my friend a more ‘complete’ impression; or I focus on one mode that is, in this particular communicative situation, of singular importance for me. In any case, I make a decision, and this decision is guided by the cultural and institutional rules in place (for instance, it would be quite extravagant to send a piece of the leather seat cover to someone so that he or she could smell it) and my particular interest at the moment (for instance, the sound of the engine as a synecdoche for the power of the car). Multimodal discourse analysis pays particular attention to the different functions of each of these modes, and also to their inter-relationships (e.g. Machin and Mayr 2012; Unsworth and Cléirigh 2009).

The meaning potential of individual modes is not the same everywhere and changes over time. Put differently, it is culturally and historically contingent, and individual actors are born into a ‘socio-historical a priori’ (Luckmann 1983) that delineates the boundaries within which different modes can be used and made sense of. In the modern Western world, we have witnessed an incredible rise, for instance, in the amount and quality of visual information that we experience in our everyday lives (e.g. Kress and van Leeuwen 2006; Meyer et al. 2013; Mitchell 1994). The possibility of digitalizing visual information and the opportunity to globally diffuse it within seconds have brought about an enormous change in the way we communicate with each other. One might be tempted to speak of some sort of ‘democratization’ of communication (e.g. Kress 2010) – with all its positive, and negative, implications. In other societies and cultures, the verbal was never as strictly differentiated from the visual, meaning that multimodality might take other shapes and fulfill different functions. The particular ‘division of labour’ between the modes is a cultural construction and matches the respective social arrangements. Critical analysis has to be aware of this and must have a clear concept of what the predominance of one mode over the other means in a particular cultural and institutional setting.

## **Relevance of multimodality for critical discourse analysis**

### **A note on the meaning of ‘critical’**

There are a number of aspects that we consider vital for our understanding of ‘critical’ multimodal discourse analysis. First, critical discourse analysis (CDA) is not a method, but rather a research

program that encompasses a variety of approaches, theoretical models, and research methods (e.g. Wodak 2011a; see also Wodak and Meyer in this volume). Similarly, ‘multimodal’ CDA is not a particular analytical approach, but, on a very basic level, constitutes the acknowledgement that discourse is not just verbal, but combines a variety of modes. Second, ‘critical’ is related to questions of how things are, why they are like that, and how they could be different (e.g. Fairclough 2010; Wodak 2011a). Such questions only prove meaningful against the backdrop of a broadly constructionist epistemology, postulating that discourse is performative and constitutive, rather than representative. This means that social reality is a human achievement and could be – at least in theory – constructed differently. However, social reality also acts back on its producers, constituting them as actors, as well as their interests and potential for further meaning-making (Berger and Luckmann 1967). We have already argued that language, in such a perspective, is the most prominent resource for meaning construction. Third, what is central to multimodal CDA is a focus on how power and interest underlie particular constructions of social reality; analysis is therefore centred on the role of discourse in the (re-)production and contestation of dominant ‘truths’ (e.g. van Dijk 1993). By purposefully ‘alienating’ the researcher from the object of study, and through the extensive interpretation of texts, critical analysis facilitates the ‘unearthing’ of such structures of dominance in particular discourses, and the identification of alternative realities.

A variety of specific approaches to CDA with divergent foci, conceptual backgrounds and analytical approaches, exist; however what they all share is the view that discourse shapes, and is simultaneously shaped by, society.

### **The contribution of multimodality to critical discourse analysis**

Following our discussion so far, it is easy to acknowledge how multimodality is a crucial topic for CDA. As Machin and Mayr (2012: 6) summarize, ‘meaning is generally communicated not only through language but also through other semiotic modes’. Power, truth and interest are, then, also represented in these other modes. Research on visual communication, for instance, has argued that visualization, due to its fact-like character, is particularly suited to supporting the truth-claims of its authors (e.g. Graves et al. 1996). Visuals are often able to ‘disguise’ power structures and hegemony as ‘objective’ representations. At the same time, power relations are inherent in practices of looking at things, and the ‘gaze’ has been identified as a disciplinary technique, disciplining both the viewer and the viewed (e.g. Foucault 1979; Kress and van Leeuwen 2006; Styhre 2010). One important point for multimodal CDA is that modes constitute conscious and unconscious *choices*

made by the author that reflect her particular social and cultural positioning as well as interests at the moment of creation (e.g. Kress 2010; Machin and Mayr 2012).

There are several ways in which multimodal CDA should assess the relationships between modes and power: First, issues of power and dominance are related to the question of *how* multimodal discourse is created. Kress argues that ‘powerful’ sign makers do not have to take into account the interests and capabilities of their audiences (Kress 2010; Kress and van Leeuwen 2001, 2006). Their sign making is strictly oriented towards their own needs, while the interpretive work is left for the audience. This, for instance, has traditionally been true in state bureaucracies, where citizens simply have to ‘learn’ the particular language of ‘officialese’ (in German, the word ‘Beamtendeutsch’ was coined for this; see, for example, the study of van Leeuwen and Wodak 1999). Conversely, if audiences possess more power than sign makers, communication and design will be more strongly oriented towards their needs and interests – although it might still try to ‘hide’ information that is detrimental to the interest of the sign maker. Contemporary corporate reporting practices, where corporations have to convince stakeholders of proper conduct and practices, are an excellent example. In such situations (multimodal) communication has to anticipate the expectations of these stakeholders (e.g. Höllerer 2013); it thus employs multiple modes to exploit their maximum persuasive potential. We could hold that this kind of power is embodied in the *form* and *design* of communication.

Second, power and power structures are, of course, created, challenged and re-negotiated in *content* – the *what* of communication. Corporate communication will usually present the board of directors in a confident stance and in a way that communicates competence and professionalism and engenders trust. Newspapers construct particular actors or actor groups in both positive and negative ways, and regularly use multimodal designs to do so. For instance, Hardy and Phillips (1999: 19) found that in editorial cartoons, immigration discourses ‘portrayed refugees as frauds, the immigration system as inadequate and the public as requiring protection’. Multimodal discourse is also used by social movements in their protest material (e.g. Philipps 2012), where visuals convey central messages and create strong emotional responses much more immediately than verbal descriptions. In all these examples, multimodal CDA provides a unique perspective that acknowledges that each mode constitutes a particular contribution to the overall signification work. It is, therefore, the central task of a student of multimodal discourse to reconstruct the ways in which the combination (or ‘ensemble’; see Kress 2010) of modes suggests particular versions of social reality that are not neutral with regard to power: they serve some interests while marginalizing others.

Finally, multimodality is linked to power in society by asking who is empowered to ‘speak’, i.e., who is granted ‘voice’ by a particular mode. Given that in modern Western cultures verbal communication is controlled more strongly than, for example, visual text or sound, these alternative modes are more susceptible to resistance and subversion by marginalized groups. Multimodality may be part of a larger shift in communication that potentially includes a systematic redistribution of power. With different modes and media of communication becoming available to a larger community of people, notions of hegemony and resistance may have to be adapted. As Kress (2010: 21) argues, ‘in all domains of communication, these rearrangements in power can be conceptualized as a shift from “vertical” to “horizontal” structures of power, from *hierarchical* to (at least seemingly) more *open, participatory relations*, captured in many aspects of contemporary communications’. The full impact of these changes, being marked structurally by more potential for participation (e.g. through the openness of the internet), and in terms of modality (e.g. legitimation of a much broader spectrum of expressive forms than just verbal language), still remains open – but the necessity and timeliness of critical research dealing with such questions seems significant.

In the following section, we will discuss research concerned with power, interest, and voice that uses multimodal data in order to capture the more comprehensive picture. The first part engages with research reconstructing issues of power more thoroughly, by accepting that multimodal discourse adds another quality to communication by making particular interests seem ‘natural’, ‘objective’ and ‘fact-like’. In the second part, we discuss how multimodal discourse provides opportunities to make otherwise marginalized voices more prominent in critical research.

## **Previous studies and exemplary research**

Since it is simply impossible to discuss the entire scope of multimodality, we make the deliberate decision to focus on the interrelationship of two particular modes: the verbal and the visual. Despite other modes being relevant in practice (see, for instance, Pinch and Bijsterveld 2012, on sound), the area of visual research has, so far, received the most scholarly attention, and provides the richest pool of concrete examples. In their literature review, Meyer et al. (2013) argue that visual research comes in many forms and shapes, and that individual studies can be roughly classified according to the role that visual material plays in the research process. In more detail, they differentiate between: an *archaeological* approach that looks for traces of meaning in existing visual discourse; a *practice* approach that focuses on the actual use and manipulation of visual material in the field; a *strategic* approach that is more psychologically oriented and studies the cognitive impact



of visual material on viewers; a *dialogical* approach, that uses visuals to initiate communication with actors in the field; and a *documenting* approach that sees visual material as an opportunity to create richer research documentation. Although all of these approaches have something to say about power and interest, we focus on two that are, from our point of view, highly promising points of departure for multimodal CDA: the archaeological and the dialogical approach.

### **‘Unearthing’ power and interest in multimodal discourse**

Research in what we have called the archaeological tradition uses multimodal artifacts as a kind of ‘window’ to gain insight into the cultural system in which they are produced, thus enhancing our understanding of the meaning structures that are created, maintained, and challenged. According to Preston et al. (1996), visuals can *reflect*, *mask/pervert*, and *constitute* social reality. While a naïve view only sees the reflection (i.e., the representation) of reality, a critical perspective acknowledges that visuals also hide aspects of reality that are not in the interest of the sign maker, and that the realities they constitute are always just one of several alternatives.

#### **Exhibit 1: The archaeological approach to multimodal CDA**

A common objective of critical archaeological approaches is to look into how particular people or groups are visually depicted, or made invisible, and what this may tell us about their status and power in society. Multimodality is important here, since the visual mode, for instance, may be purposefully used to transport messages that otherwise cannot be verbalized for legal or cultural reasons (McQuarrie and Phillips 2005).

Hardy and Phillips (1999), for example, apply such a critical approach to the study of editorial political cartoons (i.e. drawings and text) in the press, in order to reconstruct the subject-positions that this discourse assigns to different actors in the Canadian immigration system. They first reconstruct the dominant objects of discussion (the refugee, the government, the immigration system, and the public), and then analyse the meanings assigned to these objects. They found, among other things, that refugees were commonly constituted as frauds, victims, or both, and also as privileged in comparison to other immigrants.

With a slightly different focus, Schroeder and Zwick (2004) analyse aspects of masculinity in corporate advertising. They use insights from art history, visual studies, and photography to reconstruct the ‘mirror’ as a root metaphor of consumer society, creating the person as an exhibited object for visual consumption.

Visual and multimodal analysis may also be utilized to reveal fundamental discursive structures and issues of presence and absence. Höllerer et al. (2013), for instance, focus on the underlying meaning structures of visual renditions of corporate social responsibility (CSR) in annual CSR reports in Austria. They identify a set of 21 ‘discourse carrying dimensions’ that span polar opposites; these polar opposites are then clustered into a number of *topoi*. The study finds that images in the context of CSR discourse mediate spatial oppositions, bridge time, connect different institutional spheres, and help to overcome credibility gaps. Empirical results also indicate that some poles are dominant, while others are almost absent. The study is, therefore, a good example of research that reveals ‘blind spots’ in a particular discourse, and thus enables critical reflection.

The examples in Exhibit 1 illustrate how multimodality may considerably enrich traditional CDA. All of them include images in their analyses, but go beyond a discussion of visual content – the relationship between the individual visual and verbal parts is a central factor in understanding the performative power of text. It is the ‘orchestration’ (Kress 2010) of all of these elements that creates a particular version of social reality, and that serves some interests better than others.

### **Using multimodal discourse to give voice to marginalized subjects**

Power is not only manifest in discursive presences. Often, it is constituted even more distinctly in *absences*. This might sound very abstract and ambiguous at first, but think of people that are not being heard on a particular topic for various reasons. Scholars of power and domination have long recognized that power is not only exerted in direct ways, but also indirectly, by creating non-issues and preventing some topics from even entering public discussions (e.g. Bachrach and Baratz 1962; Lukes 1974). One important concept in critical research is that of voice, concerning the question of who is legitimated to speak in a certain situation and on a particular topic – and who is not. By analysing only publicly available discursive traces, social science runs the risk of overemphasizing discourses of the powerful, therefore essentially reproducing their version of reality. We wish to draw attention to the enabling and empowering aspect of multimodality and, in particular, visuality, in this respect.

### **Exhibit 2: The dialogical approach to multimodal CDA**

Critical research in a dialogical tradition is primarily concerned with voice of marginalized groups, and how research can be sure to capture ‘silenced’ discourses. This is achieved, on the one hand,

by paying attention to narratives that are not represented in the (dominant) verbal mode, and, on the other, by systematically enabling actors in the field to share their experiences in a multimodal way.

An excellent example of how visibility may empower weaker groups in a particular context is provided by Bell (2012). Her study reveals how employees use visual discourse in order to resist the official, dominant verbal narrative of powerful actors within a corporation. By doing so, they create a subversive form of organizational memory that helps them to better deal with the 'death' of the organization: 'By producing images that represented Ford management as having murdered Jaguar, they presented an alternative view of the past in the present. This narrative is more tragic and sinister than could be conveyed through spoken and written words alone' (Bell 2012: 13).

A second example, the study of Slutskaya et al. (2012), systematically utilizes this empowering aspect of visuals in order to elicit richer and more adequate responses from interviewees. Their study on the 'dirty' work of butchers, and the associated threats to identity construction, met a severe challenge when established assumptions of masculinity and patriarchy, and the associated cultural and social positioning basically 'silenced' alternative voices in the profession. Purely verbal interviews failed, since interviewees did not feel comfortable in this situation, fearing that their answers would be 'insufficient'. Photo-elicitation (i.e., taking self-shot photographs as 'triggers' for conversations) proved much more useful, since photography was better suited to 'showing' the physicality of the occupation that was central to the butcher's professional identities.

Such multimodal forms of, for instance, interviewing are often more successful in surfacing discourse that otherwise would remain hidden; these novel methodological designs, therefore, provide ample potential for CDA.

The examples in Exhibit 2 show that marginalized voices that struggle to be heard in 'official' discourses use alternative routes, such as visualization, to create their version of 'how things are'. However, lack of voice is often related to a lack of rhetorical competence. As Warren (2005: 871) explains: 'Writing is a skill that is learned according to academic or literary conventions and depends, fundamentally, on the literacy of the writer, the extent of their vocabulary, knowledge of grammatical structure and, in creative writing, perhaps even prosaic construction and poetic tropes all of which are a function of education and by extension, of socio-economic circumstances'. The underprivileged are not only politically excluded from discourse, they are also often not able to put their reality into words, at least not in a way that is equally sophisticated. Genres such as photography (for instance, in the form of snapshots), however, do not require such skills. Images are often closer to people's life-worlds than sophisticated, wordy descriptions. Techniques such as

photo-elicitation or photo-voice (e.g. Warren 2002, 2005) are, therefore, a suitable way for CDA to access discourses that are usually not available for scholarly study.

Summing up, the last pages have illustrated that multimodal discourse analysis is a rather comprehensive, broad, and ‘fresh’ endeavor. On the one hand, social reality is constructed, maintained and transformed in various multimodal ways: This entails that critical scholars need to acquire the necessary ‘literacy’ to deal with such rhetoric beyond that of the written and spoken word. On the other hand, multimodal discourse is a vehicle for subversive and alternative worldviews, while verbal discourse often favours the status quo as well as dominant interests and positions. This prompts critical researchers not to restrict their attention to verbal forms of communication in their studies. Power resides in the *access to*, in the *form*, as well as in the *content* of discursive modes – and, we should add, also in their *composition*.

## **A brief demonstration of analytical procedures**

### **Some general remarks on methods of multimodal analysis**

While there is a growing number of publications dealing with multimodal analysis (e.g. Jewitt 2009; Kress 2010; Kress and van Leeuwen 2001, 2006; Machin and Mayr 2012), there is, unfortunately, still a dearth of empirical applications that explicitly address the plurality of modes in discourse *and* systematically discuss their interrelationships as a central aspect of meaning making. The examples discussed earlier acknowledge that modes beyond the verbal are relevant; however, they do not explicitly elaborate on differences and links between them. We will, therefore, exemplify a useful analytical procedure by systematically discussing two multimodal texts in this section.

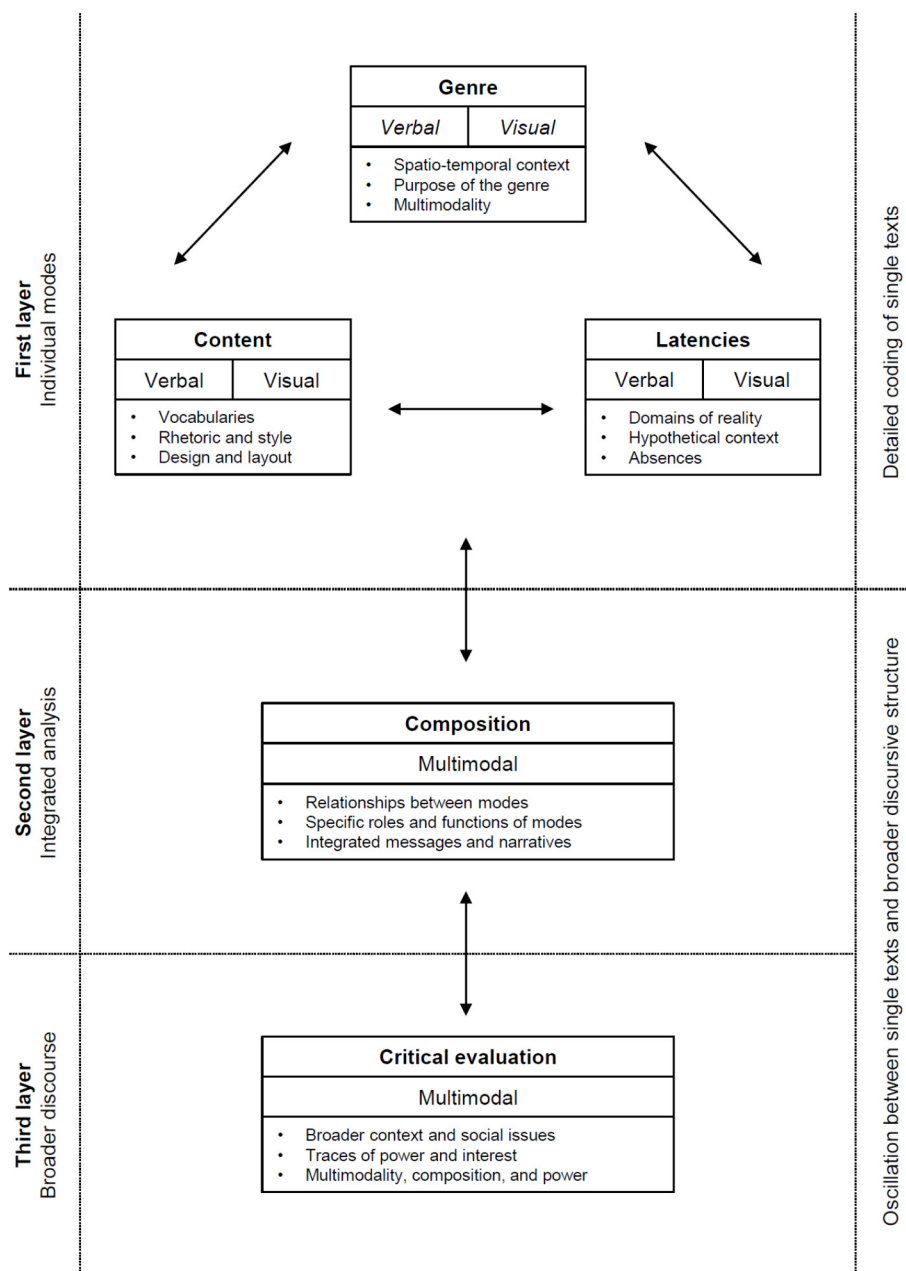
There are a number of disclaimers and caveats to be made before we present our methodological approach. First, *the* ‘multimodal method’ cannot exist, since the concrete form of analysis has to fit the particular research question, research context, and data. Second, ‘multimodal’ can mean very different things, and they all necessitate specific analytical tools. The analysis of sound (e.g. Pinch and Bijsterveld 2012), for instance, requires techniques that are very different from those of visual or verbal analysis. For these and other reasons, there are no standardized methods of analysis in multimodal research. Here, we present one specific method that is useful primarily for the analysis of multimodal material that encompasses verbal and visual text. But also in this area of application, different analytical approaches exist (and are required). A chapter on critical approaches to multimodal discourse analysis is, therefore, a rather difficult thing to write: multimodal discourse is significantly more complex than verbal discourse, and the volume at hand impressively

demonstrates how many approaches – epistemological, methodological, and analytical – exist even for verbal CDA. Moreover, in this chapter we are able to discuss only two exemplary ‘texts’. These constitute discourse fragments, and we present them in isolation, which means that we cannot assess the extent to which they are typical and characteristic for a particular discourse strand or how they are located in a particular discourse thread. The following examples should, therefore, be understood as an illustration of the more detailed analysis of particular discourse fragments, not as a complete critical discourse analysis. We focus on a number of ‘guiding questions’ for analysis, and are selective in presenting particularly striking features of the material and our interpretations of those. At the end of the chapter, we provide further readings that allow a deeper consideration of different topics and approaches in the field.

### **Introducing the method/methodology**

The analytical procedure we present in this chapter is inspired by different strands of visual sociology and semiotics, especially Kress and van Leeuwen’s (2006) social semiotics, Müller-Doohm’s (1997) structural-hermeneutic symbolic analysis, and Bohnsack’s (2007) documentary method. It was initially developed for a research project on the visual (re-)contextualization of corporate social responsibility (CSR) in the Austrian context (Höllerer et al. 2013), and was further adapted specifically for multimodal discourse in a more recent study on the construction of the global financial crisis in business media (Höllerer et al. 2014). The major benefit of this analytical approach is that it enables us to work with larger quantities of visual and multimodal data without compromising the interpretive character of the overall analysis. It constitutes some sort of ‘template’ which may (and must) be adapted to the specific research question(s) and materials at hand. We suggest five ideal-typical steps of analysis and illustrate them with two different multimodal texts. For each step, we suggest a number of ‘guiding questions’ that facilitate the analysis. Subsequently, we discuss how such analysis can be extended to larger samples, and which conceptual and methodological approaches might be utilized to do so. Our approach is flexible, in that some aspects and steps may be extended or scaled down, depending on the particular research objectives. Also, the guiding questions may vary accordingly, and different hermeneutical techniques may be applied to answer the questions. Figure 1 provides a schematic overview.

In order to illustrate each step, we discuss two practical examples of multimodal texts (Figure 2 and Figure 3): a double page from a corporate annual financial report, and a front page of a business newspaper.



**Figure 1** Overview of the methodological approach

## Analysing two exemplary multimodal texts

### *Step 1:* Characterizing the genre

A text genre is part of the institutional framework of a text (see also Reisigl and Wodak in this volume) and can be understood as ‘typified communicative action invoked in response to a



recurrent situation’ (Yates and Orlikowski 1992). As such, it strongly influences the actual form and content of texts, commanding particular ‘genre rules’. Knowing these rules is essential when analysing a text, as they provide the basic framework to understand the fundamental conditions under which authors make choices and audiences interpret a text.

There are a number of central questions to ask about a genre and its key dimensions, several of which we wish to emphasize in the following:

- *What is the spatiotemporal and sociocultural context of the text?*
- *Who is/are the producer/s of a text, and who is the audience?*
- *What is the purpose of the text genre? How institutionalized is the text genre?*
- *What are the particular genre characteristics with regard to multimodality?*



**Figure 2** OMV Annual Report 2012 (reproduced with kind permission of OMV)

*Example 1.* Our first exemplary multimodal text is taken from the 2012 annual financial report of OMV, an integrated international oil and gas corporation with its headquarters in Vienna, Austria.

Its main businesses are the exploration and production of oil and gas, natural gas distribution and power generation, and the refining and marketing of oil products. OMV is, in terms of revenue, by far the largest publicly traded corporation in Austria. As such, the text belongs to the domain of economy and business. A more detailed description of the spatiotemporal and sociocultural context of the text would emphasize, for instance, an era several years after the global financial crisis of 2008, the specificities of the oil and gas industry, the particular governance model in Austria, or the more global issues such as the multiple responsibilities of business (CSR, sustainability, shareholder value, etc.).

Annual reports of publicly traded corporations are available to the general public, and regularly provided for download on the focal corporation's website. In terms of their production, texts from the genre of annual reports are usually 'collectively crafted by executive management and communication experts on behalf of the corporation' (Höllerer 2013: 586). They are, unlike, for instance, print media, directed at more qualified audiences with nonetheless varying interests (e.g. financial analysts, shareholders, competitors, banks and creditors, regulators, journalists, or NGOs). The genre of annual reports is highly institutionalized (i.e., some content is required by legal regulation, and the structure is standardized, for instance, by guidelines such as the Global Reporting Initiative); at the same time, it offers considerable leeway for individualization and creativity, especially with regard to its non-verbal parts. Annual reports primarily serve the purpose of presenting the corporation (e.g. its history, mission, areas of operation, or current strategy) and accounting for practices and results (in terms of financial performance, but increasingly also with regard to the social and ecological dimension). In this way, they aim at shaping a qualified public's perception. While content and layout are driven by a subjective agenda, such texts inhere 'truth claims', meaning that the producers claim to present 'real' facts and figures.

Corporate annual reports instrumentalize multimodality to a high degree. Visual elements such as graphs, charts and figures, but also photographs and other images, are often used to enhance, amplify, or disguise verbal text. Some corporations go even further, adding material and haptic modes to their reports (for instance, different surfaces or three-dimensional elements), or feature video clips in the electronic version of the report. However, we also see trends that emphasize the symbolic over the fact-like character of visuals (our example goes in that direction), or, in an attempt to produce highly aesthetic reports, reduce visuals to ornaments of text layout.

*Example 2.* Our second example is a cover page of the *Australian Financial Review* from May 2012. It therefore belongs to the (business) media domain. The *Australian Financial Review* is the leading Australian voice within business, finance, economics, and policy with a high intermedia agenda-



setting function. It is published by Fairfax Media in a compact format (six days a week), as well as online. Although offering a broad range of views and opinions, the newspaper has followed a consistent editorial line clearly favouring economic liberalism. An analysis of the text's spatiohistorical and sociocultural context would point to the recent financial crisis, the global financial markets, Australia's economic system and business landscape, and so on.

The media are a highly structured genre (e.g. van Dijk 1988). Business media reporting is a highly institutionalized 'story-telling' activity that conveys news in a way firmly governed by a "logic of appropriateness" based on [...] professional and craft related roles' (Cook 1998: 61) – something that has also been described as a particular 'media logic' (e.g. Altheide and Snow 1979). The design and layout of cover pages is central within this logic, and often differs across cultural contexts of production as well as target audiences. For instance, in our example, the design broadly follows an Anglo-Saxon template (e.g. by starting several stories on the cover page to be continued later in the paper, by using rather catchy short titles, or by avoiding editorial commentary on the front page), and is a crucial decision made by the senior editorial team. The audience of the daily business news very much differs from consumers of mass media, as it primarily comprises the socio-economic elite of business professionals and personal investors, among others.

Media news production has been expressively described by Tuchman's (1973) now-classic phrase of 'routinizing the unexpected'. Business and financial news are often conceptualized as providing a 'global outlook on social reality' (Berglez 2008: 847). They provide, with a particular focus on the economic dimension, access to social reality for a broader audience. Most critically, they have a gate-keeping and agenda-setting function within public discourse (e.g. the choice to report on particular issues implies the silencing of others). Apart from claiming to report on 'true' facts and figures, they also take position with regard to the issues they deal with.

PERSPECTIVE

Who's who in the China lobby

TONY WALKER 46



ARTS & REVIEW

Tina Arena's second coming

JANE ALBERT • 57

Mona Lisa's identical twin

JOHN McDONALD • 58



May 19-20, 2012 • \$3.00 • afr.com

Weekend Edition

THE AUSTRALIAN FINANCIAL REVIEW

Knock-on effects on China increase risks for big projects

# Euro fear drives share plunge

Philip Baker

Australian shares fell into negative territory for the first time this year, driven by fears that fallout from Europe will hit the Chinese economy.

Nearly \$56.2 billion was wiped from the value of the S&P/ASX 200 index this week when it fell 5.6 per cent, in line with falls in other global sharemarkets, as fears spread that Spain is the latest flashpoint in the deepening European debt crisis that could trigger a sharp global economic slowdown. Officials in Europe are beginning to talk openly about Greece leaving the currency union.

The Australian index slumped 110.9 points, or 2.7 per cent on Friday to close at 4046.5, the lowest level since November, 2011. The index is down 0.25 per cent this year after being up as much as 9 per cent earlier this month.

"We've seen this sort of volatility before and it will remain choppy for the next few weeks," said Investors Mutual investment director Anton Tagliarero. "I am surprised that Europe has caused such a reaction given it's been around for a while, but what is clear is that the best of the resources boom is over."

The dollar was hit too, dropping to as low as US97.95¢ on concerns of a fall in commodity prices. The dollar has fallen 10 per cent against the greenback since it hit \$US1.0009 on March 1.

Fears of weaker demand from China pushed copper close to a four-month low, while oil slid to a six-month low of about \$US92.56 a barrel. BHP Billiton slumped 9 per cent this week to its lowest level in a Continued page 14



## \$9bn coal port plan scrapped

Mark Ludlow and Dan Hall

A \$9 billion privately funded enlargement of Abbot Point coal terminal in North Queensland has been killed by the Newman state government, a project that would have created one of the largest coal ports in the world and helped expand the coal industry.

In another sign that Australia's resources boom is slowing, Queensland Deputy Premier and Infrastructure Minister Jeff Seeney said on Friday the new Liberal National Party government had decided to scrap the T4 to T9 expansion of the coal terminal and a \$2.3 billion 12-berth cargo facility. Other big resources projects also at risk from falling commodity prices include South Australia's giant Olympic Dam copper and uranium mine owned by BHP Billiton, which said this week it wouldn't spend \$80 billion previously budgeted for new projects. Woodside Petroleum is slowing the pace of its troubled \$35 billion Browse gas project in Western Australia.

Mr Seeney said the former Bligh Labor government tried to expand Abbot Point coal terminal too quickly. The port is designed to become the main export hub for the Continued page 4  
Matthew Stevens, page 19

## Disqualify dodgy MPs

Louise Dodson

Tony Windsor has foreshadowed moves to widen civil and criminal grounds for disqualifying MPs, indicating interest in a Greens plan to ask Parliament on Monday to set up an integrity commissioner to address principles raised by the Craig Thomson and Peter Slipper affairs. Mr Windsor said there were civil and criminal convictions that could be added to the existing grounds for being excluded from Parliament. Meanwhile, Mr Slipper has

accused former adviser James Ashby of waging an inappropriate publicity campaign against him.

Mr Thomson, suspended by the ALE is preparing to defend himself in Parliament on Monday.

Reports, 3 and 7



## ATO blitz on trusts

Katie Walsh

The Australian Taxation Office is targeting baby boomers using trusts to avoid capital gains tax on wealth transfers, deputy commissioner of small and medium enterprises Michael Cranston has warned. The ATO will also use its expanded data-matching program to crack down on attempts to avoid tax. Mr Cranston said the Tax Office was seeing "some really aggressive behaviour" in this area.

Report, page 8

## Quigley ready to rumble

Exclusive John McDuling

Mike Quigley says he will not be bullied. The NBN Co chief executive is in the political hot seat - and faces rejection should the opposition, which opposes his \$36 billion

broadband project, win the next federal election. Yet he is determined not to be silenced if his critics get their facts wrong and intends to speak out, even in an election campaign.

"Frankly, I think I have a responsibility, if statements are made about the company which I think are factually wrong," he said. Mr Quigley's forceful advocacy of the NBN could make him a political asset for the Labor government.

Perspective, page 45

CHANTICLEER WHAT HAPPENS IF GREECE GOES UNDER - BACK PAGE

FBA 001

Figure 3 Australian Financial Review (cover page, 19-20 May, 2012; reproduced with kind permission of Fairfax Media)

Media discourse has a strong tradition of using multimodality. Visual elements are frequently used to draw attention, and/or to frame, complement, or counter verbal text (for a systematization of 'multimodal techniques' in the media, see Höllerer et al., 2014; for previous studies on visual

discourse in print and online newspapers, see, for instance, de Cock et al. 2011; Fahmy 2010; Knox 2007, 2009)

### *Step 2: Capturing the manifest content*

Analysing the manifest content of a text can take on a variety of forms, all of which are different approaches to content analysis. The primary function of this step is to sensitize the researchers for the ‘language’ of the text, as well as its most dominant features. As such, this step focuses on the conventional meaning of words and visual elements. We propose the following guiding questions:

- *What is the particular ‘vocabulary’ of the text?*
- *What kind of rhetorical and stylistic techniques and strategies are used?*
- *How can the ‘design’ and ‘layout’ of the overall text be described?*

In order to address these questions, researchers may rely on a variety of coding schemes taken from existing literature. For instance, in coding the visual vocabulary, a very basic start would be to look for different kinds of people (e.g. male/female, young/old), objects (e.g. mobile/immobile), actions (e.g. unidirectional/bidirectional), and settings (e.g. exterior/interior, private/commercial) in the visual text. For style, literature on photography and painting provides inspirations such as lighting, perspective, lines of sight, *mise-en-scène*, or interaction between viewer and image (and for a more detailed description for the analysis of verbal text, see for instance Reisigl and Wodak, and Mautner in this volume). Layout, finally, concerns the different ways of composing the overall text, including, among others, positioning, overlapping, or other ‘references’ between modes. We shall briefly illustrate this on the basis of our two examples.

*Example 1.* The verbal vocabulary of the text is characterized by a strong reliance on positive and technical language. Words such as ‘divestment program’, ‘cost management’, or ‘working capital’ necessitate a particular form of education. Apart from being very technical, the language is also distinctly positive, as shown by words such as ‘stabilization’, ‘growth’, ‘strengthened’, or ‘progress’. The linguistic-rhetorical means are not very elaborate, except perhaps for the reliance on metaphors (primarily ‘dead metaphors’; see, for instance, Lakoff and Johnson 1980); these will be discussed in more detail in Step 3 below.

By featuring the company’s main refinery on the outskirts of Vienna, the visual vocabulary clearly emphasizes the corporation’s industrial quality. It is, in its pure content, rather restricted, encompassing ‘industrial architecture’ in a minimalistic environment consisting almost entirely of the ‘sky’, but also comprising ‘lights’, ‘green lines’, and a form of ‘checklist’. In terms of style, the image portrays the refinery at night, which stresses the effects of different forms of lighting. Light

is, overall, a major factor in the impact of the image. Also, the viewer takes a ‘worm’s eye-view’ in relationship to the object in the image. It is taken from a long shot, enabling a ‘complete’ view of the premises, but surroundings are systematically absent. The image also seems to be digitally enhanced.

The overall composition is arranged as a two-page spread. The image of the industrial complex is the background of the composition – but also its centre. The title of the page is on the upper left part, and there is a series of four small paragraphs that are arranged in a ‘rising’ pattern from the bottom left to the upper right. Each of these paragraphs is ‘anchored’ to the page with the pictogram of a ‘checked box’.

*Example 2.* The verbal vocabulary of the newspaper page is much less restricted and also less technical. The charging is also decidedly negative. Several words point at the existence of immediate danger (e.g. ‘flashpoint’, ‘crisis’, ‘fallout’) and conflict (e.g. ‘kill’, ‘criminal’, ‘silenced’, ‘opposition’). Also, completely contrary to our first example, the sense of direction is downward (e.g. ‘slump’, ‘fall’, ‘drop’, ‘slide’). Given that it is an Australian newspaper, it is worth noting that ‘Euro’ and ‘China’ feature prominently. In addition, the vocabulary of the newspaper page also includes more numbers. Like the first example, the verbal text relies on metaphor to a certain degree. Numbers, as a form of rhetorical device, are used as for ‘operationalization’ and ‘objectivation’.

The visual vocabulary is more restricted than in the first example, and focuses primarily on the human factor. The three photographs related to the main topic of the page are all centred on ‘men’ in ‘formal dress’. In the central image, additional ‘graphs’ are added that show the dynamics of a number of measures related to stock performance and currency. Concerning visual rhetoric, gestures and countenance evoke moods and emotions. People are shown in close-up shots, making them more ‘tangible’ and ‘personal’ for the reader, but none of them looks into the camera directly.

The layout in this example is more complex than in the first one. On the one hand, there is a ‘hierarchy’ between parts of verbal text that is rather typical for the media (van Dijk 1988). The first title is by far the largest and most prominent, and it seems to create a ‘frame’ across the whole front page. The large image is in the very centre of the page, and it combines two visual elements as a collage. The two smaller portraits at the bottom of the page seem to ‘grow into’ the respective text that is aligned around the photograph. The complete article section is bounded at the top with the title of the newspaper, and at the bottom with a ‘preview’ of the next page.

### *Step 3: Reconstructing latent elements*

The description of manifest content and rhetorical and stylistic strategies are supplemented by an analysis of broader structures of meaning that underlie the text. The central objective in this step is to transcend the manifest layer of meaning within the text and grasp its latent meanings. The overall questions are similar across perspectives:

- *What parts or ‘domains’ of social reality are featured within the text?*
- *How can the hypothetical social context be characterized in which the text ‘makes sense’?*
- *What expected and unexpected ‘absences’ can be found in the text (e.g. in the sense of unrealized alternatives)?*

This means, primarily, that the analysis has to reconstruct the broader social and interdiscursive contexts that individual elements of the text refer to. Depending on text and genre, different coding strategies may be appropriate. Figures of speech, or tropes (e.g. metaphor, metonymy, irony) that were coded in the first step are a useful starting point for identifying such broader structures, as are various rhetorical (e.g. logos, pathos, ethos) and argumentative (e.g. enthymemes, topoi) structures. More generally, any form of hermeneutical analysis (e.g. Keller 2008; Hitzler and Honer 1997; Hitzler, Reichertz and Schröer 1999) may serve as inspiration and provide guidelines.

*Example 1.* The text shows a strong emphasis on professionalism and technology. Both visual and verbal vocabularies draw extensively from the domains of business and industry/technology. The organization, as depicted in the text, is agentic, capable, and ‘energetic’. The vocabulary is active, and the focus on light, energy, and movement in the visual part of the text amplifies this impression. Metaphors also draw from the domain of ‘construction’ (e.g. ‘pillar’, ‘build up’, ‘set up’) that is mirrored by the industrial building as the most central visual element on the spread. The vocabulary, therefore, creates an interesting tension between stability and movement that can be found both in the verbal as well as the visual parts of the text. The organization is constantly progressing at a substantial pace, implying that (constant) economic growth is not only possible, but a central objective. Expansion is necessary, but securing one’s assets is equally important. Inertia, however, is undesirable. The colour code of the image (blue and green) mirrors the company logo, and green can also be interpreted as an allusion to the sphere of environmentalism. Despite being engaged in a ‘dirty’ industry (oil), the corporation provides essential and substantial services (‘light’), values and implements ‘green’ procedures, and espouses environmental values.

The central absences in the text are ‘people’ and ‘nature’. This is all the more striking, as it is consistent across the verbal and visual aspects of the text. The world constructed on the two-page spread is completely devoid of any human agents, be they employees, customers, or investors, or

nature more generally. When the verbal text does not rely on passive constructions, it makes heavy use of ‘abstract’ actors, such as organizations or divisions. This reinforces the technological ‘feel’ of the text and creates an image of the organization as ‘powerful machine’.

*Example 2.* In the newspaper article, domains are much more clearly cued in the verbal text. The language is largely one of disaster and conflict, sometimes reaching into the rather extreme vocabulary of armed conflict (e.g. ‘kill’, ‘blitz’). It is completely focused on (negative) dynamics, with no substantial reference to stability. The hypothetical social context is strongly formal and serious. Money is quite literally the currency. There is nothing playful or creative about this world, especially now that it is in crisis. The sombre and dark (mostly black) colours reinforce the serious and dramatic impression of the overall text. Countenance and gestures of people support such interpretation. The three main personae exhibit desperation, helplessness and determination, respectively. This is also not a world for everybody. In the world of money and politics, male white elites dominate. Decisions are mainly taken by a chosen few, and the amounts of money at stake just baffle the average citizen.

The newspaper page has one striking absence – that of any form of diversity. The text creates an impression of the metaphorical ‘old boys’ club’ that consists entirely of older white men in suits. This impression is even amplified when looking at the complete page. The only two women featuring on the page are both related to art rather than business – one as an artist, the other as a model. All depicted persons are Caucasian. While it might be argued that those images ‘just represent reality’, this is hardly ever the case. Images are always selections, particular ‘framings’ of social reality. Homogeneity in depiction is an editorial decision.

#### *Step 4: Composition*

A fourth step focuses on reconstructing the effects of ‘composing’ multimodal texts in particular ways. Kress (2010) has stressed such composition as one of the central aspects of multimodal meaning-making. While the previous steps have analysed the different modes largely in isolation from each other, we now turn to a more integrated approach.

- *How do verbal and visual elements relate to each other?*
- *What are the particular ‘roles’ and ‘functions’ of the verbal and the visual within the text?*
- *What integrated ‘messages’ or ‘narratives’ are created through this composition?*

In order to address these questions, researchers may focus on various aspects of the interrelationships between modes, such as hierarchical arrangements (‘Is one mode on a higher or lower level of text structure than others?’), or issues of emphasis and dominance (‘Which mode is

at the focus of attention?'). In addition, similarities, differences, and references may also be grasped in terms of content and latent meaning ('How similar is their vocabulary?'; see steps above). It must be noted that multimodal texts rarely encompass only a single narrative. Rather, their multivocality is a central asset, connecting them to a multiplicity of divergent narratives at the level of the overall discourse. It is, therefore, impossible to present exhaustively the narratives that our examples evoke. We can offer only an indication of the most dominant stories they tell us.

*Example 1.* The verbal and visual parts of the text mutually reinforce and support each other distinctly. The verbal mode is used to 'inform'. It provides the context and the more detailed information. While the image is strongly stylized, the verbal text is supposed to tell the 'truth' and provide the 'facts'. The visual mode, on the other hand, 'personalizes' the message through the depiction of the concrete corporate building. It also adds those attributes that verbal text is not as easily able to communicate, like dynamics, and even a form of 'industrial romance'. Overall, it makes the message more 'tangible'. Layout, as a third mode, ties the other two modes together and adds emphasis and hierarchy. The composition as a whole creates an impression of potency, growth and movement, to the extent that even the way in which paragraphs are arranged imply an upward trend, and the checkboxes allude to objectives reached. This basically implies a bright future, since the corporation has chosen the right path. In addition, the corporation is presented as a 'well-oiled machine' that 'never sleeps' and provides essential services.

*Example 2.* In terms of mood and atmosphere, the verbal and visual aspects of the text strongly reinforce each other. The task of creating 'credibility' and 'facticity' is equally divided here between the verbal and the visual mode. While words provide the cast of actors, the sequence of events, and some numbers, the graphs show an 'objective' representation of developments. The visual mode adds 'emotion' in a slightly different from the first example since it is more closely tied to actual people. In the first article, which is focused on the fear of an economic downturn, such fear and worry is perfectly mirrored in the behaviour of the central actor in the image, and the diagrams make it rational and measureable. In the other articles, the faces of people depicted show aggravation and determination, accordingly. Overall, the page shows a 'triangle' structure of functions. The large image is the most symbolic and emotional one, connecting falling share prices to individual suffering and despair. The graphs 'rationalize' the message, and the two people on the bottom 'personalize' the issue. This creates an intricate equilibrium between metaphor and facticity in the story about the struggle and impotency of the societal (male) elite against the overwhelming forces of global (and possibly imported) recession and economic downturn.

### *Step 5: Conclusions and critical evaluation*

All these previous steps build upon and inform each other in an effort to reconstruct the ‘patterns’ of social meaning within the text. In terms of critical analysis, the final set of questions we need to direct at the text are concerned with questions of interest and power. Of course, in the limited space provided here, it is by no means possible to present a fine-grained and exhaustive analysis. Rather, we simply point to some of the most striking elements and encourage readers to continue with their own analysis.

- *What does the analysis tell us about broader social issues and the particular institutional and cultural context in which the text is embedded?*
- *How can we describe the different traces of interest and power that we find within presences and absences?*
- *How do the different modes and their overall composition reinforce, challenge, or conceal such power?*

*Example 1.* We may reasonably infer that a composition like this points at a context where corporations are increasingly challenged to legitimize their own conduct towards differentiated audiences. The multimodal composition expresses power and speed, but also the potency to responsibly manage such dynamics. It perpetuates the ‘meta-narrative’ of unlimited growth and suggests that the corporation is able to sustain such growth. At the same time, such pursuit is presented as a not inherently self-interested behaviour – rather it serves the basic needs of society (energy; ‘lighting up the dark’; ‘being active while others sleep’), and does so in an environmentally responsible way (stressing the colour green, ‘romantic’ imagery). This is very much in line with recent findings that corporations use multimodal texts in an attempt to de-problematize seemingly contradictory expectations of their stakeholders (Höllerer et al. 2013).

Traces of interest and power in the composition vary in their visibility. The most obvious form of power is openly depicted: the power of the successful, expanding corporation that uses technology to provide essential services to its customers, but also assumes its broader responsibilities. Focusing on the particular ‘absences’ in the text, one can find more subtle aspects. By making the people that keep the machine running ‘invisible’, for instance, the composition ascribes power and agency not only to technology itself, but also to the corporation as an abstract, legal construct. The absence of people is especially striking since this particular corporation usually makes extensive use of people in their visualizations. A second aspect that is made invisible – usually a rather controversial aspect in the oil and gas industry – is the ‘dirty’ side of production.

The verbal part of the text is the actual ‘report’: it presents objectives met and challenges overcome, which is reinforced by the visual element of the ‘ticked boxes’. The image provides these achievements with a ‘face’ – although an abstract one. Also, due to its immediate facticity, the image



is able to portray a refinery as something actually 'beautiful' and aesthetic. By constant reinforcement and repetition (ticked boxes, upward alignment of paragraphs, speedy movement in the foreground, optimistic verbal text), the central message found in the heading ('Profitable Growth is well on track') is communicated through all modes simultaneously and, therefore, made much more persuasive.

*Example 2.* There are at least two related aspects to the context that seem striking. First, there is an – explicit as well as implicit – focus on measurability, precision and causality. Explicitly, this is shown, on the one hand, through the focus on precise numbers in the text, but also in the graphs and images. Even someone unable to understand the precise meaning of the graphs will 'get' this importance of precision. To a degree, such precision is also mirrored in the 'de-contextualized' depiction of people on the bottom of the page. The absence of any 'noise' in the image leaves no questions about who is central in these photographs. A second aspect that builds on such measurability, precision and causality, is constituted by the need to assign responsibility, accountability and blame. The big headline is very clear on what causes the 'share plunge', and in the other articles, it becomes equally clear who the culprits and the heroes are.

Interest and power are more subtle in our second example. It is primarily constituted by presences and absences. There is an implicit claim to 'truth' within the multimodal text that fits the context of 'serious' business media. Such a claim is represented, on the one hand, in 'realistic' photography, and, on the other hand, through the extensive use of numbers and charts. The visual part of the text also contains additional information on 'who is relevant'. We have noted before that 'business', in this text, is exclusively 'male' and perpetuates the discourse of 'male' leadership, especially in the situation of crises and the fight against 'the forces'. The placement of the females in the domain of the Arts even reinforces this impression. The images, in this sense, are much more than just illustrations of the verbal text. The combination of text and visuals specifies 'what kind of men' are the main characters (as leaders, heroes, or culprits): here, we are dealing with an elite group of white, older men wearing suits.

### **Extending the analysis to larger samples**

So far, we have only illustrated the procedure for single multimodal texts. While some qualitative and hermeneutical approaches explicitly claim that single texts are sufficient for the study of particular questions, since every text contains the social and discursive structures in which it is embedded (see, for instance, the objective hermeneutics of Oevermann et al. 1979), often a larger

sample is needed in order to reconstruct broader parts of the discourse thread, provide a comparative analysis of different threads, find patterns on the field level or across genres, or study developments over time.

Basically, such field-level analysis would use the same procedures as those of individual texts. Steps 1-3 would remain unchanged, they would be just applied to a larger set of texts. Starting with step 4, researchers would start a process of constant ‘oscillation’ between the single text and the broader discursive structure. The challenge here is to find useful concepts to work with in linking individual textual elements to discursive structure. The options are manifold. In our own research (e.g. Höllerer et al. 2013, 2014), we have, for instance, used framings (e.g. Gamson and Modigliani 1989; Meyer 2004; Meyer and Höllerer 2010), topoi (e.g. Jancsary 2013; Wengeler 2003), discourse-carrying dimensions (e.g. Bublitz 2011; Link 1997), and narratives (e.g. Czarniawska 2004; Rowlinson et al. 2014). Other analytical concepts include *Deutungsmuster* (e.g. Meuser and Sackmann 1991; Oevermann 2001), or legitimation strategies (e.g. Meyer and Lefsrud 2012; Vaara and Monin 2010; van Leeuwen and Wodak 1999). What all these approaches have in common is that they allow for a differentiated analysis of meaning structures in a particular field or on a specific topic. Social reality and meaning are seldom monolithic, but divided into particular ‘zones’ of meaning (Berger and Luckmann 1967). Multimodal research provides additional ways of understanding such zones and the ways in which they emerge, and are maintained or challenged.

## Summary

In this chapter we have aimed at elaborating the main aspects, dimensions, and implications of multimodal CDA. We wish to emphasize that it is not so much a particular *variant* of CDA; rather, it encompasses a broad range of discourse-analytical approaches that deal with the multiple *ways and resources of (re-)constructing social reality*. Multimodal CDA engages with different forms of *data*, and therefore also employs a well-stocked conceptual and methodological toolbox. As a matter of fact, this chapter could not offer a comprehensive overview of all tools available. We provide, however, a selection of additional readings in the appendix that enable an in-depth engagement with more specific topics related to multimodal CDA.

We presented, in some detail, one specific methodological approach that is particularly suited for the analysis of large samples of multimodal material (i.e., visual and verbal elements) in order to detect the broader, underlying meaning structures that organize discourse and social reality. We are well aware that our methodological suggestions here do, by no means, provide a standardized

'schema' according to which multimodal CDA should proceed. It should therefore be understood as what it is: an illustration of how one could proceed in doing multimodal CDA. It is, nevertheless, our hope that in offering ideas in a more systematic way, we can inspire our readers as to how such analysis should be conducted. Also, we wish to stress that the strengths of our approach lie, particularly, in its flexibility, adaptability and its applicability to larger corpuses of multimodal data (as opposed to a methodology that excels in in-depth interpretation of single cases or small samples). While primarily developed for the analysis of verbal and visual text, our guiding questions can be adapted relatively easily for studies that aim at different modes of discourse.

Of course, multimodal analysis also has to face a number of important challenges. First, it entails a rather strong dependence on data and documentation. Some forms of data collection (e.g. interviews, surveys) are not tailored for modes beyond the verbal. Also, actors in the field might be reluctant to provide multimodal accounts of their experiences (e.g. photographs, videos, drawings). Second, multimodal research strains researchers' abilities to deal with a variety of modes at the same time, all of which require particular, and potentially very divergent, sets of analytical skills. Third, since different modes create meaning in rather specific ways, comparison is not trivial. Fourth, contemporary publication outlets are often ill equipped to deal with other modes than the verbal. Still, in order to capture contemporary social reality that is increasingly constructed, mediated, reproduced and challenged by a multitude of discursive modes that become ever more accessible, multimodal literacy becomes, in our view, a necessity for researchers of the social world.

## Further reading

*Kress, G. (2010) Multimodality. A Social Semiotic Approach to Contemporary Communication.* Abingdon: Routledge.

In his book, Kress provides a rich and detailed introduction into multimodal discourse from a social semiotic perspective. The volume entails an elaborate theoretical discussion of meaning and communication, and makes ample use of examples and illustrations in order to make concepts and ideas more accessible.

*Kress, G. and van Leeuwen, T. (2006) Reading Images. The Grammar of Visual Design.* London: Routledge.

Focusing on the visual mode of communication, this book provides an in-depth engagement with the way visuals ‘work’ and how they can be more systematically understood. The authors engage with various aspects of the visual, from content to style to latent meaning. It is an invaluable resource for a better understanding of the visual elements in discourse.

*Machin, D. and Mayr, A. (2012) How to Do Critical Discourse Analysis. A Multimodal Introduction.* London: Sage.

Machin and Mayr make multimodality an explicit part of their version of CDA. Their book is an accessible and excellently structured overview of different aspects that such multimodal CDA encompasses. In a systematic way, it covers topics such as speech and speakers, representing people and action, absences, persuasion and ‘truth’. Their book is rich with illustrations that exemplify their approach.

*Meyer, R. E., Höllerer, M. A., Jancsary, D. and van Leeuwen, T. (2013) The Visual Dimension in Organizing, Organization, and Organization Research: Core Ideas, Current Developments, and Promising Avenues.* *Academy of Management Annals* 7: 487-553.

This article provides a detailed and systematic overview of research on visuals and visibility in organization and management research; it also touches on related disciplines that have dealt with visibility extensively (such as, for instance, psychology, communication studies, or philosophy). Meyer and her colleagues suggest that visuals may play a multiplicity of different roles in (critical)

research, and present a typology of approaches that also serves as inspiration for future research in this area.

*Rose, G. (2007) Visual Methodologies. An Introduction to the Interpretation of Visual Materials.* London: Sage.

Rose presents an elaborate overview of the most prominent and promising methodological approaches to visual analysis. The book covers content analysis, semiology, psychoanalysis, audience studies and anthropological approaches. It explicitly devotes two chapters to visual discourse analysis.

## Tasks

*(A) Choose a cover page of any tabloid newspaper that seems of interest to you. Select one article on that page that encompasses some kind of visual element.*

1. Take a look at the ‘vocabulary’ of the verbal text. What are the most striking nouns, verbs, and adjectives? Do they belong to a particular ‘domain’ (e.g. war, love, family, sports)?
2. In the same vein, take a look at the visual ‘imagery’. What elements can you identify (e.g. people, objects, actions)?
3. Try to summarize the visual and the verbal text in a short narrative of no more than two or three sentences.
4. How do these stories relate to each other? Do they support or contradict each other, or do they seem to be unrelated? Is there an overall story that can be told across modes?
5. In whose interest is it to tell the story in this particular way? Can you identify winners and losers?

*Some additional things to think about:*

- Does it make any difference whether you perceive the verbal or visual part of the text first? If yes, what exactly changes?
- Is it important how the verbal and visual parts of the text are spatially positioned in relation to each other? If yes, what changes when you rearrange them?

(B) Choose a scene from one of your favourite movies (preferably one where a dialogue between characters happens).

1. What modes can you identify (e.g. spoken word, visual impressions, body language, composition of the shot)?
2. Try to define, for each of the modes you identified, what its role or function is in the overall composition. How does their impression on you differ from the others? One way to assess this could be to imagine how the scene would affect you if one mode was absent.
3. What information do you get through the particular interplay of modes that is not explicitly 'said' or 'shown' in the scene?

*Some additional things to think about:*

- Is there a particular 'sequence' to the use of modes in the scene (e.g. is there a strong visual impression first, and spoken text comes later)? How does that influence your understanding of the scene?
- Could you completely change the meaning of the scene by changing individual modes? If you could, then how?

## Acknowledgements

We gratefully acknowledge financial support from the Danish Research Council: DFF-1327-00030.

## REFERENCES

- Altheide, D. L., and Snow, R. P. (1979). *Media logic*. Beverly Hills, CA: Sage Publications.
- Bachrach, P., and Baratz, M. S. (1962). Two faces of power. *American Political Science Review*, 56: 947-952.
- Bell, E. (2012). Ways of seeing organisational death: A critical semiotic analysis of organisational memorialisation. *Visual Studies*, 27: 4-17.
- Berger, P. L., and Luckmann, T. (1967). *The social construction of reality. A treatise in the sociology of knowledge*. New York: Anchor Books.
- Berglez, P. (2008). What is global journalism? Theoretical and empirical conceptualizations. *Journalism Studies*, 9: 845-858.
- Bohnsack, R. (2007). Die dokumentarische Methode in der Bild- und Fotointerpretation. In: R. Bohnsack, I. Nentwig-Gesemann and A.-M. Nohl (eds), *Die dokumentarische Methode und ihre Forschungspraxis* (pp. 69-91). Wiesbaden: VS Verlag.
- Bublitz, H. (2011). Differenz und Integration. Zur diskursanalytischen Rekonstruktion der Regelstrukturen sozialer Wirklichkeit. In: R. Keller, A. Hirsland, W. Schneider and W. Viehöver (eds), *Handbuch Sozialwissenschaftliche Diskursanalyse. Band 1: Theorien und Methoden*. Wiesbaden: VS Verlag.

- Cook, T. E. (1998). *Governing with the news. The news media as political institution*. Chicago, IL: The University of Chicago Press.
- Czarniawska, B. (2004). *Narratives in social science research*. London: Sage Publications.
- de Cock, C., Baker, M., and Volkmann, C. (2011). Financial phantasmagoria: Corporate image-work in times of crisis. *Organization*, 18: 153-172.
- Fahmy, S. (2010). Contrasting visual frames of our times: A framing analysis of English- and Arabic-language press coverage of war and terrorism. *International Communication Gazette*, 72: 695-717.
- Fairclough, N. (2010). *Critical discourse analysis. The critical study of language*. London: Routledge.
- Foucault, M. (1979). *Discipline and punish: The birth of the prison*. New York, NY: Vintage Books.
- Gamson, W. A., and Modigliani, A. (1989). Media discourse and public opinion on nuclear power: A constructionist approach. *American Journal of Sociology*, 95: 1-37.
- Graves, O. F., Flesher, D. L., and Jordan, R. E. (1996). Pictures and the bottom line: The television epistemology of U.S. annual reports. *Accounting, Organizations and Society*, 21: 57-88.
- Hardy, C., and Phillips, N. (1999). No joking matter: Discursive struggle in the Canadian refugee system. *Organization Studies*, 20: 1-24.
- Hitzler, R., and Honer, A. (eds). (1997). *Sozialwissenschaftliche Hermeneutik. Eine Einführung*. Opladen: UTB.
- Hitzler, R., Reichertz, J., and Schröer, N. (eds). (1999). *Hermeneutische Wissenssoziologie. Standpunkte zur Theorie der Interpretation*. Konstanz: UVK Universitäts-Verlag.
- Höllerer, M. A. (2013). From taken-for-granted to explicit commitment: The rise of CSR in a corporatist country. *Journal of Management Studies*, 50: 573-606.
- Höllerer, M. A., Jancsary, D., and Grafström, M. (2014). 'A picture is worth a thousand words': Visually assigned meaning and meta-narratives of the Global Financial Crisis. Working paper, WU Vienna.
- Höllerer, M. A., Jancsary, D., Meyer, R. E., and Vettori, O. (2013). Imageries of corporate social responsibility: Visual re-contextualization and field-level meaning. In: M. Lounsbury and E. Boxenbaum (eds), *Institutional logics in action, part B (Research in the Sociology of Organizations, Vol. 39B)*. Bingley: Emerald. pp. 139-74..
- Jancsary, D. (2013). *Die rhetorische Konstruktion von Führung und Steuerung: Eine argumentationstheoretische Untersuchung deutschsprachiger Führungsgrundsätze*. Frankfurt a.M.: Peter Lang.
- Jewitt, C. (ed.). (2009). *The Routledge Handbook of multimodal analysis*. Abingdon: Routledge.
- Keller, R. (2008). *Wissenssoziologische Diskursanalyse. Grundlegung eines Forschungsprogramms*. Wiesbaden: VS Verlag.
- Knox, J. (2007). Visual-verbal communication on online newspaper home pages. *Visual Communication*, 6: 19-53.
- Knox, J. (2009). Punctuating the home page: image as language in an online newspaper. *Discourse & Communication*, 3: 145-172.
- Kress, G. (2010). *Multimodality. A social semiotic approach to contemporary communication*. Abingdon: Routledge.
- Kress, G., and van Leeuwen, T. (2001). *Multimodal discourse. The modes and media of contemporary communication*. London: Hodder Education.
- Kress, G., and van Leeuwen, T. (2006). *Reading images. The grammar of visual design*. London: Routledge.
- Lakoff, G., and Johnson, M. (1980). *Metaphors we live by*. Chicago: The University of Chicago Press.
- LeVine, P., and Scollon, R. (eds). (2004). *Discourse & technology. Multimodal discourse analysis*. Washington, D.C.: Georgetown University Press.

- Link, J. (1997). *Versuch über den Normalismus. Wie Normalität produziert wird*. Opladen: Westdeutscher Verlag.
- Luckmann, T. (1983). *Life-world and social realities*. London: Heinemann.
- Lukes, S. (1974). *Power. A radical view*. London: Macmillan.
- Machin, D., and Mayr, A. (2012). *How to do critical discourse analysis. A multimodal introduction*. London: Sage Publications.
- McQuarrie, E. F., and Phillips, B. J. (2005). Indirect persuasion in advertising: How consumers process metaphors presented in pictures and words. *Journal of Advertising*, 34: 7-20.
- Meuser, M., and Sackmann, R. (eds). (1991). *Analyse sozialer Deutungsmuster. Beiträge zur empirischen Wissenssoziologie*. Pfaffenweiler: Centaurus.
- Meyer, R. E. (2004). *Globale Managementkonzepte und lokaler Kontext. Organisationale Wertorientierung im österreichischen öffentlichen Diskurs*. Wien: WUV.
- Meyer, R. E., and Höllerer, M. A. (2010). Meaning structures in a contested issue field: A topographic map of shareholder value in Austria. *Academy of Management Journal*, 53: 1241-1262.
- Meyer, R. E., Höllerer, M. A., Jancsary, D., and van Leeuwen, T. (2013). The visual dimension in organizing, organization, and organization research: Core ideas, current developments, and promising avenues. *Academy of Management Annals*, 7: 487-553.
- Meyer, R. E., and Lefsrud, L. M. (2012). Science or science fiction? Professionals' discursive construction of climate change. *Organization Studies*, 33: 1477-1506.
- Mitchell, W. J. T. (1994). *Picture theory. Essays on verbal and visual representation*. Chicago: University of Chicago Press.
- Müller-Doohm, S. (1997). Bildinterpretation als struktural-hermeneutische Symbolanalyse. In: R. Hitzler and A. Honer (eds), *Sozialwissenschaftliche Hermeneutik: Eine Einführung*. Opladen: Leske + Budrich. pp. 81-108.
- O'Halloran, K. L. (ed.). (2004). *Multimodal discourse analysis*. London: Continuum.
- Oevermann, U. (2001). Die Struktur sozialer Deutungsmuster - Versuch einer Aktualisierung. *Sozialer Sinn*, Heft 1, 35-81.
- Oevermann, U., Tilman, A., Konau, E., and Krambeck, J. (1979). Die Methodologie einer "objektiven Hermeneutik" und ihre allgemeine forschungslogische Bedeutung in den Sozialwissenschaften. In: H.-G. Soeffner (ed.), *Interpretative Verfahren in den Sozial- und Textwissenschaften*. Stuttgart: Metzler. pp. 352-434.
- Philipps, A. (2012). Visual protest material as empirical data. *Visual Communication*, 11: 3-21.
- Pinch, T., and Bijsterveld, K. (eds). (2012). *The Oxford handbook of sound studies*. New York: Oxford University Press.
- Preston, A. M., Wright, C., and Young, J. J. (1996). Imag[in]ing annual reports. *Accounting, Organizations and Society*, 21: 113-137.
- Rowlinson, M., Casey, A., Hansen, P. H., and Mills, A. J. (2014). Narratives and memory in organizations. *Organization*, 21(4), 441-446.
- Royce, T. D., and Bowcher, W. L. (eds). (2007). *New directions in the analysis of multimodal discourse*. Mahwah, NJ: Lawrence Erlbaum.
- Schroeder, J. E., and Zwick, D. (2004). Mirrors of masculinity: Representation and identity in advertising images. *Consumption Markets & Culture*, 7: 21-52.
- Slutskaya, N., Simpson, A., and Hughes, J. (2012). Lessons from photoelicitation: Encouraging working men to speak. *Qualitative Research in Organizations and Management: An International Journal*, 7: 16-33.
- Styhre, A. (2010). *Visual culture in organizations: Theory and cases*. London: Routledge.
- Tuchmann, G. (1973). Making news by doing work: Routinizing the unexpected. *American Journal of Sociology*, 79: 110-131.



- Unsworth, L., and Cléirigh, C. (2009). Multimodality and reading: The construction of meaning through image-text interaction. In: C. Jewitt (ed.), *The Routledge Handbook of multimodal analysis*. Abingdon: Routledge. pp. 151-163.
- Vaara, E., and Monin, P. (2010). A recursive perspective on discursive legitimation and organizational action in mergers and acquisitions. *Organization Science*, 21: 3-22.
- van Dijk, T. A. (1988). *News as discourse*. Hillsdale: Lawrence Erlbaum.
- van Dijk, T. A. (1993). Principles of critical discourse analysis. *Discourse & Society*, 4: 249-283.
- van Leeuwen, T., and Wodak, R. (1999). Legitimizing immigration control: A discourse-historical analysis. *Discourse Studies*, 1: 83-118.
- Warren, S. (2002). "Show me how it feels to work here": Using photography to research organizational aesthetics. *ephemera*, 2: 224-245.
- Warren, S. (2005). Photography and voice in critical qualitative management research. *Accounting, Auditing & Accountability Journal*, 18: 861-882.
- Wengeler, M. (2003). *Topos und Diskurs. Begründung einer argumentationsanalytischen Methode und ihre Anwendung auf den Migrationsdiskurs (1960-1985)*. Tübingen: Niemeyer.
- Wodak, R. (2011a). Critical discourse analysis. In: K. Hyland and B. Paltridge (eds), *The Continuum companion to discourse analysis*. London: Continuum. pp. 38-53.
- Yates, J., and Orlikowski, W. J. (1992). Genres of organizational communication: A structural approach to studying communication and media. *Academy of Management Review*, 17: 299-326.

## **Glossary**

**Mode.** A mode can be defined as a culturally available resource for constructing meaning. This entails a mode being something to be employed in communication. There is no comprehensive list of modes: the most common, however, are verbal text, visuals, gesture, spatial layout/design, and sound. Whether a specific mode is available in a social situation, and what kind of meaning may be expressed in a particular mode, is largely influenced by the specific institutional and cultural context in which communication happens to occur.

**Multimodality.** The term multimodality describes the fact that most of our communication does not just include a single mode – but utilizes a multiplicity of them. Multimodal discourse analysis, therefore, has to acknowledge that people use different materials and meaning resources simultaneously and/or for different objectives. It therefore focuses on the various functions of each of these modes, their composition and orchestration, and their specific contribution to meaning (re-)construction.

**Visual.** The visual mode basically refers to meaning resources that we primarily experience with our sense of sight. In contrast to verbal language that works according to principles of sequence and linearity, visual structures create meaning primarily through immediacy and spatial arrangements. Vision, however, is also socially constructed/regulated (i.e., social rules and conventions influence what we are able and allowed to ‘see’). The visual encompasses a large variety of different expressive forms and artefacts. This includes more ‘physical’ genres such as photographs, pictures, paintings, drawings, and sketches as well as ‘non-physical’ ones, such as charts, diagrams, models, and typography. The visual comprises still images as well as motion pictures.