The need for a social and affordance-driven multimodal critical discourse studies

David Machin
Örebro University, Sweden

Abstract

Given the way multimodality as a field has expanded, becoming more diverse and complex, it is important to pause to identify exactly which concepts, theories and processes of multimodal analysis are more or less suitable for the needs of critical discourse analysis (CDA) and the wider field of critical discourse studies (CDS). The article argues that the field of multimodality remains fragmented both internally, with a range of divergent core interests, and externally from academic fields that have long dealt with the topics to which it is turning its interest. In this article, looking at some key ideas from visual studies, I reflect on what kind of multimodal approach best aligns with the needs of CDS. I argue for an affordance-based approach and one driven by the social and not by need to model on the basis of language.

Keywords

Introduction

While multimodality has now developed into an academic field in its own right, it is only more recently that we start to find multimodal work in critical discourse studies (CDS) journals. There are a relatively small number of books (Abousnnouga and Machin, 2013; Djonov and Zhao, 2014; Machin and Mayr, 2012; Mayr and Machin, 2012) and a special edition (Machin, 2013) on the subject. These publications begin to indicate how multimodal analysis can be best aligned with the core aims of CDA. CDA seeks to reveal

Corresponding author:
David Machin, School of Humanities, Education and Social Sciences (HumUS), Örebro University, Fakultetsgatan 1, SE-70182 Örebro, Sweden.
Email: david.machin@oru.se

buried ideologies in texts, to show how the powerful seek to re-contextualise social practice in their own interests and maintain control over ideology. These publications show how it is clear this is done multimodally, as different kinds of semiotic resources, as well as language, are deployed for ideological purposes. As regards visual communication, this has been in the form of photographs, visual design and in popular television and film.

Given the way multimodality as a field has expanded, becoming more diverse and complex, it is important to pause to identify exactly which concepts and processes of analysis are more or less suitable for the needs of CDA; what kind of multimodal approach best suits MCDA, or multimodal critical discourse analysis. This means looking inwards at concepts used in multimodality but also, importantly, as I show in this article, outwards to other academic fields, which have longer and more established traditions of visual analysis.

In recent years in fact, as within linguistics itself, there has been a wider increase in scholarly interest in visual communication, with a number of new larger edited multidisciplinary collections from the social sciences and from the humanities (Pauwels, 2012; Rose, 2012). This increase represents a growing acknowledgement of the important role played by visual communication. But it also represents a growing awareness of the fragmentation of approaches to the visual, that there is need to share knowledge of visual communication, to engage in cross-disciplinary dialogue. Pauwels (2012) argues there has been a tendency in visual research for different fields to reinvent the wheel as they operate in their own isolated networks. It is clear that MCDA has much to gain from such engagement.

Scholars have argued that one problem with work being published in multimodality is that it has, as yet, insufficient consistency or agreement in how terms are used or defined. Forceville (2010) argues that multimodality is characterised by a flourishing of terminology which remains for the most part untested, rather than by isolating and identifying clear defendable concepts. One reason for this flourishing of untested concepts is that multimodality tends to take on a huge range of topics, encompassing more and more things, rather than carrying out more localised studies around one object of study to develop robust and defendable principles and concepts. This leads, Forceville suggests, to conflating complex media such as 3D with things like design and gesture before smaller steps have been taken and firm footing established for simple forms of media. And one cause of this state of affairs is precisely that it remains fragmented from other scholarly work that deals with these specialist visual areas. Engaging with these can help provide more established theories and concepts against which we can place those from multimodality, helping to show precisely where its more systematic tools can be best used, to indicate what multimodality can and cannot do, how it can be enriched. And specifically for MCDA, its tools and approaches need to be able to clearly and compellingly demonstrate that they can play a role in revealing the dominant ideologies that lie buried in all forms of communication.

I begin this article by identifying the very different aims and concepts that are buried within multimodality itself. I show that the different sub-fields that lie under the umbrella of multimodality have concepts, tools and processes of analysis designed to do very different things, and which are generated from very different starting points, even though it is common to find a selection of these lumped together as 'multimodality' in the theory

and methods sections of journal articles. Looking at theories and concepts from visual studies and from media and cultural studies, I then go on to show that not all of these are so well aligned for a socially driven form of analysis such as CDA and that some concepts and approaches may simply cloud and distract. In the last section, I look at how MCDA can best take these ideas on board. I focus in this article only on theories of photographs. This is due to space but, I hope, will indicate the need to engage with wider fields of work in other areas.

Multimodality: Systems or situated meaning making

Two books are credited as starting multimodality: Kress and Van Leeuwen's ([1996] 2006) *Reading Images* and O Toole's (1994) *The Language of Displayed Art*. Both were greatly influenced by the Systemic Functional Linguistics (SFL) of Michael Halliday (1978). This differed from other theories of language. It was a social theory, placing the emphasis not on a more rigid, or formal, grammar, so much as an overall system of semantic choices or alternatives made up of layers of smaller subsystems which build into the whole. This system is used by speakers in contexts to meet their communicative needs.

The main aim for SFL was to model the systems and subsystems of choices that form language which are called 'systemic networks'. These are described as being based on three meta-functions underlying meaning making. Semiotic modes (connected systems of resources) are simultaneously used to say something about the world (the ideational meta-function), to signal our relationships (the interpersonal meta-function). As they do this, they bring ideational and interpersonal meanings into a structured whole (the textual meta-function) (Halliday and Hasan, 1985).

The multimodal scholarship that has drawn from Halliday is on the one hand functional in the sense it describes meanings people make through a process of active choices in social contexts. On the other hand, it is systemic and based on an assumption that all semiotic modes could be described in the same way.

Reading Images was a highly interdisciplinary book drawing on semiotic theories of visual communication, along with theories from art history and cognitive psychology. The book was also linked to the project in critical linguistics of both the authors, so multimodality was also about how different visual representations and conventions could be shown to carry ideologies. These combined influences allowed the authors to see that a more systemic approach could contribute to existing semiotic and visual theories which tended to be less predictive and that this could also help in identifying the functional and highly motivated aspect of sign-making, placing this in social, political and historical context.

The Language of Displayed Art, less influential outside of linguistics than *Reading Images* yet providing the inspiration for much multimodal scholarship from within linguistics, also has the functional and systemic dimensions. But this book is much more driven by an interest in showing how concepts and forms of analysis used in SFL could be used to describe works of art and sculpture. Emphasis is placed on the three meta-functions and how aspects of art and sculpture, such as the shapes of figures, can be seen as fulfilling these. In this sense, this book is more systemic as is the

body of work it has inspired. It is precisely this tension between the functional/social on the one hand and the systemic part of semiosis that lies at the heart of how we can best approach MCDA on the other which becomes clear when we engage with wider theories of the visual.

The main sub-fields of multimodality

Reading Images has become routinely used by scholars to support analyses and has inspired a very specific, more interdisciplinary, more socially focused, functional kind of multimodal research. Here, the emphasis has been more on situated meaning and on the affordances of semiotic resources than on the system itself. The aim is also on understanding which resources are provided institutionally for meaning making and the ideological effects of these. For example, Jewitt (2008) has pointed to the ideological implications of the semiotic resources introduced into classrooms through new technologies.

The majority of the work which is presented as multimodal per se, however, is more closely related to *The Language of Displayed Art*. Here, the aim is to provide insights into the systems that underlie different modes of communication, in the fashion of Halliday, in order to produce system networks, or to show how different modes, in design, film, architecture, etc. fulfil Halliday's three communicative meta-functions. For example, O'Halloran (2008) shows how mathematics can be understood as multi-semiotic discourse involving language, visual images and symbolism.

Two other sub-fields of multimodality have also emerged. Scholars with an interest in cognitive metaphor theory have increasingly taken an interest in multimodal communication, mainly inspired by the work of Forceville (1996), with emphasis on metaphors in advertisements, cartoons and films. It has been taken on by some critical discourse analysts (Catalano and Waugh, 2013), for example, to show visual metaphorical patterns in the representation of immigrants in press photography.

Sociolinguists have also produced innovative multimodal work called Interactional Analysis through the work of Scollon and Scollon (2004) and Norris (2004). Here, the visual is studied in the way that it is always part of complex 'multimodal ensembles' (Norris, 2004). This work foregrounds the way that language use in settings must be seen as part of a highly subtle interplay of different semiotic resources, including things like gesture, posture and proximity, and also in relation to the rich cues in the actual immediate environment.

In sum, these different strands of multimodality represent a shift in emphasis among language studies scholars from language as a site of meaning making, to emphasise the interlocking role of all semiotic resources. The potential is to create more predictive models of the building blocks of different forms of communication, graphic design, gesture, space, art, etc., and in turn develop a more powerful tool for analysis of the actual use of resources in context.

However, criticisms can be made of multimodality that apply in different ways to these different threads, to some extent depending on the degree of emphasis on the functional or systemic part of semiosis. In what follows, I look at these criticisms reflecting upon them with theories of the visual from visual studies.

The problem of systems versus affordances

One reason for the expansion of objects of analysis in multimodality, the drive to study all forms of diverse visual communication, can be explained partly by the nature of SFL itself. SFL uses texts and other semiotic materials in order to establish 'the grammar', or systems underlying the resources. This process and the idea that there is a universal theory for meaning making can make the systemic part of meaning making the actual object of analysis in itself. One problem with this process is that it can ignore wider, and immediate, sociological context. The signs are studied in a way that sees them as resulting from grammatical systems and disconnects them from the motivated interests of the actual sign users and the emerging power relations infused in their use (Holmberg, 2012).

The assumption that the semiotic behaviour of sign users is guided by more or less the same conventions regardless of the contexts and semiotic modes involved can create some problems. Kress (2010) argued that 'A multimodal social-semiotic approach assumes that all modes of representation are, in principle, of equal significance in representation and communication, as all modes have potential for meaning' (p. 104). This is an important statement to make to linguistics. But it is as important, as Kress also points out, to highlight that resources/modes are never exactly of the same order, or quite simply they would not have evolved. While it may be reasonable to see semiotic resources that are more elaborated, such as typography, for example, as having potential for meaning, it may not be possible to say they are the same as language or as sound, a photograph or clothing. The priority should be to not lose sight of the specific affordances that different kinds of semiotic resources carry, and in the context of CDA, therefore, why they are deployed in contexts for specific ideological purposes. In other words, it is how they are different that is crucial.

The greatest danger with the drive to approach objects of analysis with the aim of finding a system and armed with a universal theory of communication is that the analysis that takes place can appear as post hoc (Bateman et al., 2004; Machin, 2009). Therefore, it is not clear whether what is found from the analysis is simply an interpretation that is rather justified by the use of terminology. Forceville (1999) made the case that feels that examples are chosen that allow concepts and models to be illustrated. This process can suffer from producing lots of descriptive tools but fall short on showing how these reveal, what added insights are produced (Reynolds, 2012). The labelling of phenomena, as Antaki et al. (2003) point out, is not the same as actually doing analysis and showing what the payoff of that analysis actually is. In the best of CDA, this payoff should be clear and the same in the case for MCDA.

The concept of mode as an equaliser

The concept of 'mode' is itself problematic (Bateman, 2013; Machin, 2013). At a basic level, it is clear that modes are impossible to isolate. In fact, Halliday (1975) himself recognised early that modes never exist in isolation, that meaning is done through complex combinations of types of semiotic resources, that they evolve together.

One the one hand, the use of the term 'mode' can be understood in the context of linguistics where it was necessary to signal clearly that meaning was not created by

language alone. But on the other, it has helped create an equalisation of different kinds of semiotic resources, distracting from the affordances of each. It has also shifted attention from the way that modes work together, and have evolved precisely in the process of interaction. Modes have a fundamental dependency on each other (Ledin and Machin, 2015). For example, writing is always realised through forms of typography and through spatial arrangement. But meaning is not simply therefore created multimodally. The different semiotic resources produce meaning together, but each brings different affordances. It is possible to say something with one kind of semiotic resource that you cannot say with another.

Kress (2010) argued that for semiotic resources to qualify as being a mode, they must fulfil Halliday's (1978) three communicative meta-functions, as does language. But here too, there is a risk if this is taken as an indication that all semiotic resources can be treated the same, which is the case if we are in the first place looking to identify a system. Other scholars have questioned whether it is indeed appropriate to describe all kinds of semiotic resources with the three meta-functions (Van Leeuwen, 1999: 189). Bateman (2013), in fact, views the use of the three meta-functions as redundant since anything can be subjected to the weight of a set of concepts.

In multimodal work, it is possible to find photography treated as a mode. But there are reasons why an affordance-driven approach might be better, avoiding the notion of mode altogether. One of the aims of the field of semiotics (Peirce, 1984) has in fact been precisely to characterise the very different kinds of signs that compose different kinds of semiotic resources. For example, the abstract signs that comprise language are highly conventionalised and symbolic in a way that is very different from iconic signs, or indexical signs, that we find in images, which resemble or imitate the object they depict. Images can depict as well as symbolise in a way that the code of language cannot. Symbols, as in language, do not resemble what they refer to, but signify through conventions. In the case of images, of the photograph for example, the interpretation does not have to be learned in the form of a code as in language (Barthes, 1977). A photograph has no code. It can be a connotative message, but this is at the level of production and reception and not at the level of the message itself. The affordance of the photograph is that it claims to represent unmediated reality. For Tagg (1988), this is the ideological trick of the photograph. It produces an all-seeing spectator yet removes the means of its production. It is not seen per se as a selected moment, framed and angled in a very specific way.

This idea of the ideological trick of the photograph has led to a particular trajectory of critical work on the photograph in visual studies and in media and cultural studies. This provides a number of valuable insights for a multimodal approach to images, pointing to the risks of a systemic approach.

The basis of this work on the photograph has been the notion that cultural forms like photographs construct our knowledge about the world. This was part of a challenge to the perceived documentary role of the photograph. Early uses of the photograph were celebrated for their power to document and bear witness. Histories of photojournalism show how from the 19th century the ability of the image to document allowed masses of people to access conflict, suffering and other social issues in an unprecedented way. But such a view of the way images simply reflect reality shifted with the emergence of visual studies to one where we are concerned with how they come to create meaning.

One key theorist in this field has been Sontag (2004), who pointed to the way that iconic images come to stand for and replace complex processes. The Nik Ut image of the napalmed girl comes to represent the Vietnam war, at the cost of the very specific and complex political context behind the war. Such images come to trigger discursive associations and to substitute actual history (Triandafyllidou et al., 2009).

Others have made similar observations on the visual representation of war, famines and disasters: that they become represented rather by typical and memorable images that are consonant with existing news values and public understanding, which become inscribed into institutions, practices and cannons of use (Cottle, 2009). For the reporting of an earthquake, we see images that bear witness to the tragedy – a boat on a roof, a child's toy in a pool of water formed in rubble; then in subsequent images, moments of hope and heroism; then later of the community coming together in scenes of clearing up, remembrance and mourning. For Sontag (2004), we have become trained to understand complex processes by memorable and often standard types of images.

Here, we see Tagg's trick of the photograph. We can see real people in them, real grief and real destruction, so the images must bear witness to the events. But also we find that the meaning of the image lies not so much within the depicted scene itself, but in the way that it references other well trodden themes, and institutionalised uses. In this sense, what does this tell us about the affordance of the photograph? Certainly, it points to the need for an approach that is fundamentally social and asks what images are used for and by whom, and of which social practices they are a part.

Of course, such photographs can certainly be analysed as regards their content and as regards how they represent. Scholars such as Bouvier (2014) have shown how war is represented in photojournalism in a way that backgrounds violence and the overall political and economic aims of wars to foreground things like technology, humanitarianism and people's uprisings. Such studies are carried out using a combination of content analysis (asking questions of a corpus of images as regards things like how participants are represented and what they are represented as doing) and Barthesian (1973) influenced closer readings where culturally salient features are interrogated – a photograph of a dirty-faced child waving a flag represents an innocent, un-dirtied set of hopes and ideals. The child and the dirt, in this case, are not codes, but depictions that carry and evoke wider cultural associations or discourses. And the affordances of the photograph mean that at one level, unlike language, it does not have to be specified that 'this child represents the meaning of this war as a simple clash between innocent needs and family values and evil-doers'.

These studies are highly influenced by Foucault (1972), who was interested in how the values of the powerful are disseminated in the most mundane ways in society in the form of discourse. This provides one way to think about how the child, the dirt and flag make meaning. Foucault's concept of discourse reminds us that discourses may never be fully present in any single text. The discourse of war as humanitarian intervention may not be found in the images themselves, but in the code. The representation of vulnerable children and of scruffy undisciplined bands of militia waving guns above their heads indexes this discourse.

The highly influential W.J.T. Mitchell (1998), in *The Last Dinosaur Book*, looked at how dinosaurs have been represented visually in different decades in ways that reflect

discourses dominant at the time. He observed a shift from those depicted as clumsy, lumbering, always pointlessly fighting, to those which were agile, intelligent, who worked in teams and communicated to work strategically. He related this to a broader shift from a society run by large state institutions to one run by the ideologies of corporatism, the rule of the team and where positions are largely unstable – the flexibility and competition that lie at the heart of neo-liberalism. The meaning of the representations of the dinosaurs here is found not so much in the grammar of the images, but in relating these to wider discourses, in the Foucauldian sense, and how they are meaningful to those who look at them.

What these studies remind us about multimodality is that it remains to be shown what additional levels of insights are being contributed through its tool kit and approach – which tend routinely to be things such as proximity, angle and modality. Would it help to show that the tails of the dinosaurs form vectors? My point here is that this would need to be carefully and systematically demonstrated, rather than added as a kind of descriptive term in a sort of post hoc analysis. Most importantly, it is clear that MCDA would have to be able to incorporate these wider insights into its own approach, how it formulates research and analysis and the tools it uses to do so. I come to this shortly.

The reception of images through ideology

There has been a long tradition in visual studies as to how we look at images, what is called 'acts of seeing'. In a sense, this is akin to thinking about the way that people are positioned in ideology to see the world through particular discourses. This throws up a number of issues as regards the problems with approaching an image from a systemic point of view.

What is of interest here is why particular viewers look in different ways and how images invite different kinds of looking. In Berger's (1972) terms, this meant thinking about how images can say things about who is looking. Lutz and Collins (1993) provide one classic example of such work showing how the photographs in *National Geographic* magazine reveal a highly romanticised view of cultures around the world related to ideas of innocence and that they represent fundamental human values, which are in fact based on highly Western assumptions rooted in Romanticism and the idea of the pure human spirit.

It has also been shown how the way we see works of art must be understood not only in terms of what lies on the canvas, but through certain cultural ways of understanding the world greatly influenced by Romanticism, and also by the workings of the art industry (Alexander, 2003). Scholars have charted the emergence of the idea of art, and the individual genius artist, as a particular set of discourses from the 18th century. The question here is not what is Art per se, but how and why do certain things become called Art? This has also been linked to issues of power and legitimation where this allows the culture of the dominant classes to be valued as superior, and attributes their sense of aesthetic values as superior (Bourdieu, 1984). The point here is that the meaning of an image is never present in the object itself. This does not mean the image itself cannot be analysed, but this should be done in the context of these wider discourses, perhaps in the fashion of Mitchell (1998), and then as regards how and why at the level of semiotic choices.

Burgin (1982) was interested in the way that the meaning of art can change and be transferred as we change our way of looking from one of looking at a documented moment to one where we look for the aesthetics, the emotion that is captured. Classic images from photojournalism, originally already selected moments, in the fashion described by Sontag (2004), become re-used, assembled alongside other iconic images, taking on whole new meanings, even used as part of advertising campaigns, which in themselves can influence the nature of discourses about the issues they represent.

Multimodality for CDS

I now want to turn to how we can think about photographic meaning as regards its affordances, through an approach that emphasises the social. I turn to the semiotics of Hjelmslev (1963) and Voloshinov (1973).

Hjelmslev was interested in the way that humans use signs to give shape to what is otherwise an amorphous mass of thought. He also described the substances in the world (like sounds, materials, shapes, things, unmarked temporality) as an amorphous mass, not yet made meaningful by humans, not yet used as signs. The sign gives order, and places meaning, in both.

Hjelmslev, like Saussure (1974), attributes the sign as having two faces, the signifier and the signified. But he usefully breaks these into two further dimensions: substance and form. Substance at the level of signifier means the substance chosen to create meaning, so this could be sound as in language, or colour or a photograph. Substance at the level of the signified constitutes the ideas and discourses that this can communicate. But this does not tell us how a sign-maker knows which substance creates what kinds of ideas. This comes from the level of form. Form at the level of signified is the repertoire of meaning potentials. This could be the abstract system in the case of language, the meanings assigned to colours in a particular culture or the collected meanings gathered for the photograph. Form at the level of the signified is the actual selections made from this repertoire and the rules that have become established for doing so – this might include sentence form, verb tense or established forms of design, or associated cannons of use for the photograph. The sign-maker makes a selection as regards substance based on the meanings and shapes available in form to communicate ideas and discourses. But since these four components are in fact one and the same thing, faces of the sign, the ideas and discourses at the same time feed back into form – although the term 'feedback' is misleading as both substance and form are the sign itself.

Importantly for Hjelmslev, the sign-function itself gives shape to consciousness, and gives order to the material world. This creates room for a more Foucauldian (1972) sense where discourse is not always entirely present in any individual act of communication or text. The world of substance shaped by the sign-function, which can include photographs, buildings, landscapes, is already made meaningful by discourse, by the sign-function. It also indicates that there will be a strong relationship between the nature of consciousness, at a particular time, and the available forms of expression. In one sense for Hjelmslev, the world of substance and consciousness itself is also part of the sign-function. Practically speaking, in this model, function and system become much more as one. On the one hand, I would suggest that the nature of thought here, of how it is

organised in our heads and how language may differ from things like images, has long been a core question of philosophy (see Dennett, 1987; Goodman, 1968). But what I think it makes clear is that there are some things we must deal with.

Hjelmslev's view leads us to view the sign as fundamentally motivated. Signs, as was argued by the Marxist linguist Voloshinov (1973), must never be studied in an abstract or in an unhistorical manner. The meaning of semiotic resources is part of a struggle over the definition of reality where the powerful in society will seek control over this process. For Barthes (1977), it was this sensitivity to such a history that was important for denaturalising signs. Signs, like the photograph of the dirty-faced child carrying the flag, carry traces of power relations. These power relations are inscribed into the institutional processes of the news organisations which produce them and into the discourses which they realise, which themselves are infused into culture, into consciousness as ideology already appearing as a natural form in the world of substance.

It is for these reasons that when dealing with something like a photograph, the emphasis in the first place for MCDA should be on the social: Why is this photograph being deployed? How is it already discursively shaped and governed by practices of representational use? Is it claiming to represent the nature of a war? Is it claiming to be artistic? Is it claiming to be educational? If we look for the codes, with no clear sense of what the image claims to be doing as regards a concrete social issue, then we can fall into pointless description.

For MCDA, the question here as regards the photograph would be as follows: What affordances does it bring to the overall representation which serves the interests of an institution or an ideology, for example? With all semiotic resources, this means asking what they have evolved to do. MCDA would be interested in how the institution regulates the use of such images, and also how these are experienced by those who look at them. For the analysis of the photograph, we may well look at who is represented and what they do. We may ask Barthesian questions about the cultural meanings of the objects and people they depict, or carry out a social actor analysis. But we would also be mindful of the way that the photograph comes loaded with many of the kinds of cannons of use and the discourses of truth and Art revealed in visual studies.

We can also put this in a different way. In MCDA, like CDA, it is important to show how discourses seek to control and shape social practices in the interests of dominant ideology. In other words, discourses can be thought of as being used to re-contextualise social practices (Van Leeuwen, 2009). This can be done through removing, adding or substituting elements such as processes, causalities or identities, or shifting about the actual order of sequence. So it is important to ask how specific affordances play a role in re-contextualising social practice.

In the case of the photographs of war that depicted dirty-faced children waving flags, we might be aware that the news outlet wishes to represent a typical news frame of 'hope and healing' that is often used to maintain reader interest when it is felt they are no longer interested in other aspects of the situation itself. The picture editor chooses images of children that resonate with reader understandings of innocence and hope. The images, for the reader, may play a role of documenting a situation, although if it is a special feature, there may also be expectation that they are in some ways recognisable as being of artistic merit, usually signified by a typical set of techniques. In this process, the actual

ongoing multi-sided civil conflict becomes one re-contextualised as where resolutions are tangible, where the 'ordinary person' with simple needs is easily identifiable. At the same time, foreign policy, as in the case of the conflicts in Iraq and Libya, either grossly underestimates the complexity of the situation, or silently goes on in the background forging deals for the oil now in the hands of a more coooperative, hastily assembled, government.

Conclusion

As many scholars in multimodality are becoming aware, there is the need to avoid tunnel vision and reinventing the wheel. If MCDA can make a meaningful contribution to the study of the visual, this needs to be through an engagement with wider scholarly work.

Now that multimodality is emerging as a field in its own right, it is vital that more care is given to the kinds of tools and approaches that are required for different purposes – and tools that are more robust and defendable. CDA has a very specific set of aims. Of course, in CDA itself there are a range of different approaches and tools, so this should be the case in MCDA. But it is important to establish what kinds of approaches and tools for analysing the visual are best suited for these aims. For MCDA, I have suggested that we must favour an approach which better locates the sign both as motivated and as having form, but also which roots it in ideology and how this shapes the way the world appears to us – what Foucault realised about discourse and power.

For MCDA, it is important to work more on the idea that discourse is itself always present in the sign at every level and that the sign itself is what forms the ideological consciousness, which Althusser (1977) described as persons already being formed through this ideology in complex ways. Analysing the use of a photograph means having a greater sociological imagination in order to understand the wide discourses and institutional processes and dynamics of hegemony that shape the choice to use them. It is through this sociological imagination that we bring the social to MCDA and at the same time allow MCDA to offer something clearly unique and exciting.

Declaration of conflicting interests

Funding

References

Abousnnouga G and Machin D (2013) *The Language of War Monuments*. London: Bloomsbury.
Alexander VD (2003) *Sociology of the Arts*. London: John Wiley & Sons.
Althusser L (1977) Ideology and ideological state apparatuses (Notes towards an investigation). In: Althusser L (ed.) *'Lenin and Philosophy' and Other Essays*. London: New Left Books, pp. 127–186.

Antaki C, Billig M, Edwards D, et al. (2003) Discourse analysis means doing analysis: A critique of six analytic shortcomings. *Discourse Analysis Online* 1: 1–22.

Barthes R (1973) *Mythologies*. London: Fontana Press.

Barthes R (1977) *Image, Music, Text*. London: Fontana Press.

Bateman J (2013) *Review of T. van Leeuwen (2009) the Language of Colour: An Introduction*. London: Routledge. *Linguistics and the Human Sciences* 9(1). Available at: http://www.equinoxpub.com/journals/index.php/LHS/article/view/20898

Bateman J, Delin J and Henschel R (2004) Multimodality and empiricism. *Perspectives on Multimodality* 6: 65–87.

Berger J (1972) *Ways of Seeing*. London: Penguin Books.

Bourdieu P (1984) *Distinction: A Social Critique of the Judgement of Taste*. Cambridge, MA: Harvard University Press.

Bouvier G (2014) British press photographs and the misrepresentation of the 2011 'uprising' in Libya: A content analysis. In: Machin D (ed.) *Visual Communication*. Berlin: De Gruyter Mouton, pp. 281–299.

Burgin V (1982) Looking at photographs. In: Burgin V (ed.) *Thinking Photography*. London: Palgrave Macmillan, pp. 142–153.

Catalano T and Waugh L (2013) The ideologies behind newspaper crime reports of Latinos and Wall Street/CEOs: A critical analysis of metonymy in text and image. *Critical Discourse Studies* 10: 406–426.

Cottle S (2009) *Global Crisis Reporting*. Milton Keynes: Open University Press.

Dennett D (1987) *The Intentional Stance*. Cambridge, MA: MIT Press.

Djonov E and Zhao S (eds) (2014) *Critical Multimodal Studies of Popular Discourse*. New York: Routledge.

Forceville C (1996) *Pictorial Metaphor in Advertising*. London: Routledge.

Forceville C (2010) Review of C. Jewitt (ed.) *The Routledge Handbook of Multimodal Analysis (2009)*. *Journal of Pragmatics* 42: 2604–2608.

Foucault M (1972) *The Archaeology of Knowledge*. London: Tavistock.

Goodman N (1968) *Ways of Worldmaking*. Indianapolis, IN: Hackett Publishing.

Halliday MAK (1975) *Learning How to Mean: Explorations in the Development of Language*. London: Edward Arnold.

Halliday MAK (1978) *Language as Social Semiotic*. London: Edward Arnold.

Halliday MAK and Hasan R (1985) *Language, Context, and Text*. Oxford: Oxford University Press.

Hjelmslev L (1963) *Prolegomena to a Theory of Language*. Madison, WI: University of Wisconsin Press.

Holmberg P (2012) Kontextsomaktivitet, situationstypochpraktik. Enkritiskanalysavkontextbegreppetisystemisk-funktionellteori [Kontext as activity, situations type and practice. A critical analysis if context understanding in systemic functional theory]. *Språk och stil* 22: 67–86.

Jewitt C (2008) Multimodality and literacy in school classrooms. *Review of Research in Education* 32: 241–267.

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

Kress G and Van Leeuwen T ([1996] 2006) *Reading Images: The Grammar of Visual Design*. London: Routledge.

Ledin P and Machin D (2015) How lists, bullet points and tables recontextualize social practice: A multimodal study of management language in Swedish universities. *Critical Discourse Studies* 12: 463–481.

Lutz A and Collins JL (1993) *Reading National Geographic*. Chicago, IL: University of Chicago Press.

Machin D (2009) Multimodality and theories of the visual. In: Jewitt C (ed.) *The Routledge Handbook of Multimodal Analysis*. London: Routledge, pp. 181–190.

Machin D (2013) What is multimodal critical discourse studies? *Critical Discourse Studies* 10: 347–355.

Machin D and Mayr A (2012) *How to Do Critical Discourse Analysis*. London: SAGE.

Mayr A and Machin D (2012) *The Language of Crime and Deviance*. London: Bloomsbury.

Mitchell WJT (1998) *The Last Dinosaur Book*. Chicago, IL: University of Chicago Press.

Norris S (2004) *Analyzing Multimodal Interaction*. London: Routledge.

O'Halloran KL (2008) Systemic functional multimodal discourse analysis (SF-MDA): Constructing ideational meaning using language and visual imagery. *Visual Communication* 7: 443–475.

O'Toole M (1994) *The Language of Displayed Art*. Teaneck, NJ: Fairleigh-Dickenson University Press.

Pauwels L (2012) An integrated conceptual framework for visual social research. In: Margolis E and Pauwels L (eds) *SAGE Handbook of Visual Research Methods*. London: SAGE, pp. 3–23.

Peirce CS (1984) *Writings of Charles S. Peirce: A Chronological Edition, Vol. 2: 1867–1871*. Bloomington, IN: Indiana University Press.

Reynolds E (2012) Review of S. Norris, Identity in (inter)action: Introducing multimodal (inter)action analysis. *Discourse Studies* 14: 805–817.

Rose G (2012) *Visual Methodologies*. London: SAGE.

Saussure F de (1974) *A Course in General Linguistics*. New York: McGraw-Hill.

Scollon R and Scollon SW (2004) *Nexus Analysis*. London: Routledge.

Sontag S (2004) *Regarding the Pain of Others*. London: Penguin Books.

Tagg J (1988) *The Burden of Representation*. Basingstoke: Palgrave Macmillan.

Triandafyllidou A, Wodak R and Krzyżanowski M (eds) (2009) *The European Public Sphere and the Media: Europe in Crisis*. Basingstoke: Palgrave Macmillan.

Van Leeuwen T (1999) *Speech, Music, Sound*. Basingstoke: Palgrave Macmillan.

Van Leeuwen T (2009) *Discourse and Practice*. London: Routledge.

Voloshinov V (1973) *Marxism and the Philosophy of Language*. Cambridge, MA: Seminar Press, in liaison with the Harvard University Press and the Academic Press Inc.

Author biography

David Machin is Professor of Media and Communication at Örebro University, Sweden. His publications have widely theorised and applied multimodal critical discourse studies, also with an interest in how these can connect with design practice. His recent books include *The Language of War Monuments* (2013) and *Visual Journalism* (2015). He is co-editor of the journals *Social Semiotics* and *Journal of Language and Politics*.