Theo van Leeuwen

Multimodality as a multi-disciplinary practice

Abstract (German)

In diesem Aufsatz gibt Theo van Leeuwen einen Überblick über einige zentrale Aspekte und Entwicklungen seiner Arbeiten zur Multimodalität. Er erörtert, wie seine Forschungen von Hallidays systemisch-funktionaler Linguistik inspiriert wurden, aber auch von Disziplinen, die eine lange Geschichte in der Untersuchung der Modi und Medien haben, mit denen er sich beschäftigt, zum Beispiel Kunsttheorie und Musikwissenschaft. Sodann diskutiert er einige der Wege, durch die seine Arbeit ihren systemisch-funktionalen Ausgangspunkt modifiziert oder ergänzt haben, und zwar indem er (1) verschiedene metafunktionale Konfigurationen für verschiedene Modi und Medien und verschiedene multimodale Praktiken postuliert, (2) einen parametrischen Ansatz zur Analyse des Bedeutungspotentials von verkörpertem und materiellem Ausdruck entwickelt, (3) Praxis gegenüber Struktur in den Vordergrund stellt, und (4) die spezifischen Vorteile der verschiedenen Modi und Medien sowie die ihnen gemeinsamen Merkmale hervorhebt.

Abstract (English)

In this paper Theo van Leeuwen reviews a few key aspects of, and developments in, his work on multimodality. He discusses how his work has been inspired by Halliday's systemic-functional linguistics but also by disciplines that have a long history in studying the modes and media his work deals with, for instance art theory and musicology. He then discusses some of the ways in which his work has modified, or added to, its systemic-functional departure point - positing different metafunctional configurations for different modes and media and different multimodal practices; developing a parametric approach to analysing the meaning potential of embodied and material expression; foregrounding practice over structure; and moving to stress the specific affordances of different modes and media as well as the features which they have in common.

Keywords: Embodiment; materiality, media, metafunctions, multi-disciplinarity, multimodality, parametric approach, provenance

(1)

I see multimodality as both an object of study and a particular approach to studying that object.

As an object of study, multimodality is studied in a number of areas, including literary studies, history and anthropology, and known under different names, for instance 'intermediality', 'transmediality', 'media convergence',

'multisensoriality', and more (Elleström, 2019, Hassler-Forest & Nicklas, 2015, Jenkins, 2008, Classen, 1997).

As an approach to studying that object, multimodality differs from other approaches in that it has its roots in linguistics. Linguists have engaged with modes of communication other than language throughout the 20th century, perhaps most influentially in the Prague School, in the Paris school of structuralist semiotics, in North American work in 'non-verbal communication', and, more recently, in Hallidayan-inspired social semiotics.

But I do not consider multimodality a discipline, because, as I will discuss in more detail below, multimodality must also be 'multi' in the sense of being multidisciplinary. This does not make it theoretically incoherent. The multidisciplinary approach I practice is grounded in social semiotics, and in many ways a continuation and an extension of the semiotic theories of the Prague School and the Paris school, but it connects to other disciplines as well.

(2)

What the Prague School took from linguistics was primarily its general semiotic principles. Mukařovský (1976 [1936]: 272) argued that the model of the sign could apply to other semiotics, in all of which "the artefact functions as an external signifier", while the signified can be found in the 'collective consciousness' of the culture and society, rather than in the semantics of language. However, Mukařovský also said that:

> Language is the most fundamental set of signs […] in view of the fact that a permanent feature of the entire realm of signs is the tendency to express everything in verbal terms. (1976 [1936]: 273)

With this not all Prague school scholars agreed. According to Veltruszký (quoted in Matejka and Titunik, 1976: 281-2):

> Words cannot be fully translated into gestures, pictures, music. The meaning of a picture cannot be fully conveyed by language, music, the play of facial muscles, etc. Each of these types of signs is entirely different, each has its own unique ability to refer to certain kinds and certain aspects of reality and each is deficient in some respects.

The second main idea the Prague school took from linguistics was functionality, which had been introduced by Bühler (1965 [1934]) who had distinguished the representational function (what is spoken about), the expressive function (related to the speaker) and the appellative function (related to the person or

persons spoken to). Mukařovský (1976 [1936]: 157) argued that these functions apply to all semiotic modes: "We are easily able to distinguish traces of the three basic functions in any informational discourse".

Paris School semiotics emerged in the late 1950s, with Roland Barthes as a pivotal figure as far as multimodality is concerned. Rather than focusing on the arts, like the Prague School, he focused on popular media and popular culture, which had grown in importance as a result of the post-war importation of American popular culture into Europe. His linguistic inspiration came primarily from Saussure and Hjelmslev, as well as from French linguists such as Martinet, Benveniste and Mounin. But while Saussure saw linguistics as part of semiotics, Barthes saw semiotics as part of linguistics. In contrast to Mukařovský he saw "the world of signifieds as none other than that of language" (Barthes, 1967: 10) and language "not only as a mode, but also as component, relay or signified" (ibid).

Like the Prague school, the Paris school primarily relied on linguistics for general semiotic concepts, "analytical concepts which we think a priori to be sufficiently general to start semiological research on its way" (ibid: 2). For Barthes these were 'langue' and 'parole', 'signifier' and 'signified', 'syntagm' and 'system', and 'denotation' and 'connotation'. He did, however, also say that applying these concepts to other modes of communication might require adapting them: "It is not presupposed that semiology will always be forced to follow the linguistic model closely" (ibid: 11).

Like Barthes, Christian Metz, with whom I studied in the early 1980s, also focused, in his early work on the semiotics of the cinema, on general principles of linguistic analysis - "commutation, analytical breakdown, strict distinction between the signifier and the signified" (1974: 107). Although systemic-functional linguists would later analyse images in terms of ranks of different magnitude (O'Toole, 1994; Boeriis and Holsanova, 2012), for Metz, "the cinema has no phonemes, nor does it, whatever one may say, have words" (1974: 65). It only has shots and these, Metz thought, make meaning at a level analogous to the clause. But unlike clauses, shots cannot be analysed into 'parts of speech'. They are "blocks of reality", "actualized in their total meaning" (ibid: 11). It is only on the level of discourse that patterns and regularities can be found, such as those which Metz identified in his 'grande syntagmatique' (ibid: 108-146), a typology of film sequences primarily based on the conjunctive relations between the shots in these sequences.

Norris, in her 4-volume anthology of multimodality papers (Norris, 2016), rightly sees 1950s and 1960s North American work in 'non-verbal communi-

cation' as multimodality research. Foreshadowed by Sapir, who had described gesture as "an elaborate and secret code that is written nowhere, known by none and understood by all" (Sapir, 1949: 556), scholars like Birdwhistell (1970) followed the model of the American descriptive linguistics of the period quite closely, positing 'kinemes' as nonverbal phonemes, and arguing against universalist psychological accounts of body language and against the idea of non-verbal communication as 'paralanguage', which he referred to as a "modifier temptation" holding "that all other modes of communication are to be studied as subsystems subordinate to it [i.e. to language]" (ibid: 175). Instead, he showed, there are many forms of kinesic behaviour in which verbalization is absent, and the modes in the multimodal mix of everyday interaction are ordered in complementary rather than hierarchical ways.

Here are some of the questions all this raised for me:

- Is semiotics part of linguistics or is linguistics part of semiotics?
- Is multimodality confined to taking only general principles from linguistics, or can it also adapt the fine-grained forms of analysis that have been one of the key strengths of linguistics?
- Can language express the whole of a culture's semantics or not?

My approach to exploring these questions is deeply grounded in linguistics, yet, over time, I began to question some aspects of the close correspondence between language and other modes Gunther Kress and I had theorized in our early work on visual communication (Kress and van Leeuwen, 1990) – and the same applies to Kress (see e.g. chapter 4 of Kress, 2011).

(3)

From its beginnings, in the late 1980s, social semiotics has seen language as one mode among others. In our early work on visual communication (Kress and Van Leeuwen, 1990), Kress and I nevertheless closely followed the model of systemic-functional linguistics, not only with respect to general semiotic concepts and methods, such as its assumption that all semiotic modes have resources for representing what is going on in the world (the ideational metafunction), for enacting interactions (the interpersonal metafunction) and for creating coherent texts (the textual metafunction), but also in doing what Metz had said cannot be done, analysing clause-level units into their constituents. This was possible because Halliday analyses clauses in terms of functional-semantic constituents such as 'Actor', 'Process' and 'Goal', rather than in terms

of formal constituents such as 'nominal group' and 'verbal group'. 'Visual grammar', Kress and I argued, can realize the same functional constituents, but in different ways – by means of 'volumes' and 'vectors', formal aspects of visual composition[1] which we took from the art theory of Rudolf Arnheim (1982), rather than by means of nominal groups and verbal groups. In this way, we could analyse Metz's "blocks of reality" into their constituents after all, showing that reality, in images, is constructed rather than reproduced. We then applied the same approach to the construction of interpersonal and textual meaning – the recognition that images can construct visual 'speech acts' ('multimodal acts') was, at the time, innovative, since image analysis had, for the most part, focused only on representation.

By recognizing functionally different kinds of vectors and different configurations of constituents such as 'Actor', 'Process' and 'Goal', Kress and I were able to map a meaning potential, a range of meaningfully different ways in which visual composition can construct what we called 'narrative' (as opposed to 'conceptual') representations of reality, for instance either as transactions involving two participants, an 'Actor' and a 'Goal' or as non-transactional actions involving only an 'Actor' or only a 'Goal', or as either reactions involving an 'eyeline' vector or as actions, involving other kinds of vectors. These options we then represented as system networks of mostly binary choices. In our account of conceptual representations of reality (inspired by Halliday's 'relational', 'identifying', and 'existential' systems), and of systems realizing the other metafunctions, we followed the same approach.

Comparing these systems to Halliday's transitivity networks showed considerable overlap, and we concluded that many (but not all) of the options for constructing narrative representations of reality can be realized linguistically as well as visually - the visual, for instance, has less resources for representing different kinds of mental processes, while language does not have resources equivalent to the complex structures of visual diagrams. In short, like Mukařovský, and unlike Barthes, we saw semantics as belonging to culture, not to language. Cultures distribute the meanings that constitute their overall semantics among the different semiotic modes they use in specific ways, so that, for instance, some meanings can only be expressed visually, others only verbally, and again others both visually and verbally, and so on. In this they often privilege some modes over others, in ways that are subject to historical change, sometimes radically (iconoclasms). In the most recent edition of *Reading Images* (Kress and Van Leeuwen, 2021) we have, for instance, tried

to show how the semantic reach of the visual has increased over, roughly, the past century.

This approach was subsequently adopted in social semiotic accounts of other semiotic modes, for instance sound and music (Van Leeuwen, 1999) and space (Ravelli and McMurtrie, 2016), while others adapted other aspects of Halliday's grammar to the visual, for instance its hierarchical ranking of linguistic units (O'Toole, 1994, Boeriis and Holsanova, 2012). Nevertheless, over time I began to question some aspects of the close fit between language and other modes, and between linguistic and multimodal analysis, asking, for instance:

- Can all semiotic modes realize all three of the metafunctions?
- Is meaning only made at the functional level, or also at the level of the materiality and the embodiment of the signifier?
- Has the emphasis on systems led to a neglect of meaning-making as practice?
- Can the 'same' meanings expressed through different modes ever fully be the same?

(4)

Starting with the first of my questions: can all semiotic modes realize all metafunctions?

In *Reading Images* Kress and I drew a parallel between the linguistic system of mood and the system of the gaze in images. When depicted people (or, indeed, animals) look at the viewer, this realizes, we said, a 'Demand' analogous to Halliday's basic speech function of 'Demand of goods and services', which, in English, is prototypically realized by the imperative mood. When they do not look at the viewer, this realizes an 'offer of information', in English prototypically realized by the indicative mood. But, I realized, to express such a visual demand, someone must be *represented* as looking at the viewer. The interpersonal must piggyback on the ideational, so to speak.

More generally, some modes, for instance the visual, have developed more resources for representation, while others, for instance music, have developed more resources for interaction, for the 'interpersonal'. Such divisions of labour may change over time, as new demands are placed on modes. The beginnings of opera, for instance, led to the deliberate development of representational resources in music, the so-called *stile rappresentativo* (McClary, 1991), which

are still in evidence in contemporary film music. Ravelli and I (Ravelli and Van Leeuwen, 2018) tried to show how the semantic reach of visual communication has extended during the 30 years since the first version of *Reading Images* was published.

In short, as I put it in an afterword to *Speech, Music, Sound (*1999: 190):

> Different semiotic modes have different metafunctional configurations, and these metafunctional configurations are neither universal, nor a function of the intrinsic nature of the medium, but cultural, the result of the uses to which the semiotic modes have been put and the values that have been attached to them.

Metafunctional divisions of labour between modes can also occur in actual instances of communication, as became evident in my work with Emilia Djonov and others on PowerPoint (e.g. Djonov and Van Leeuwen, 2014; Zhao, Djonov and Van Leeuwen, 2014). The dot points in many Powerpoint slides often take the form of complex nominal groups and therefore lack mood, modality and direct address. Language here loses most of its interpersonal dimension, as it is mood and modality which, in Halliday's words (1985: 70) make language "something that can be argued about – something that can be affirmed or denied, and also doubted, contradicted, insisted on, accepted with reservation, qualified, tempered, regretted and so on". But PowerPoint slides form part of a larger, multimodal presentation, in which the presenter will supply the interpersonal elements, attuning the slides to the specifics of the occasion where, and the audience to whom, they are presented, and in which the ideational function therefore takes a backseat, becoming, for the most part, a matter of deictic reference to the information on the slides. In other words, there is a division of labour between modes – writing does the ideational work, speech the interpersonal work.

There is also the broader question of whether Halliday's three metafunctions are, in the end, adequate. Like other multimodal researchers, and like Jakobson a long time ago (Jakobson,1960), I have noted the increasing importance of aesthetics, even in non-artistic texts and other semiotic artefacts (Van Leeuwen, 2015). Although aesthetics has often been regarded as non-functional, it has come to play an important role in many forms of functional communication, not only in advertisements, but also, for instance, in news (Caple, 2013), and of course in the templates Microsoft Word offers for everyday texts such as business letters, invoices and invitations, which all must 'look good'. Even as I write this, Microsoft adhorts me to "add some flair to your text by applying a text effect such as shadow or glow". The function of aesthetics, then, as

propagated by Microsoft, is to imprint identity on the text and to enhance the communicator's affective investment in that identity, all this in terms that have been pre-packaged by Microsoft.

(5)

Is meaning only made at the functional level, or also at the level of the materiality and the embodiment of the signifier?

Just as linguistic analyses can be applied to speech as well as writing, two very different media, the analytical methods Kress and I developed in *Reading Images* can be applied to materially quite different media – paintings, etchings, relief sculptures, photographs and so on (see 6 below for our distinction between 'modes' and 'media'). This works at the level of the functional design of texts and other semiotic artefacts, but, like others, we increasingly realized the need for an approach that can explain how and why a glass bottle means something different from a plastic bottle, a velvet jacket something different from a denim jacket – and a text with "glow" and "shadow" effects something different from a text in which typography is self-effacing and aims only at legibility and consistency.

This led us to develop two 'non-lexicogrammatical' approaches to multimodal analysis. The first focused on the *parameters* of materiality – graphic shape (including typography), colour and texture, and, in time-based media, movement and timbre. In analysing colour (Kress and Van Leeuwen, 2002), we took our cues from Jakobson and Halle's phonology (Jakobson and Halle, 1956) and interpreted colours as bundles of distinctive features: just as a [p] is plosive *and* frontal *and* bilabial, we argued, so a colour can be dark *and* saturated *and* plain *and* several other features. Later I extended this also to written language: a font can be bold *and* angular *and* elongated *and* several other features (Van Leeuwen, 2006). Such features are all graded, so that a colour, for instance, can be *quite* dark and *somewhat* saturated and *maximally* plain and so on, and result from all these features in their different proportions, just as the taste of a dish results from all its ingredients in their different proportions. But we differed from Jakobson and Halle in seeing these features, not as purely distinctive, serving only to make lexicogrammatical meanings perceivable, but as themselves carrying meaning potential. In this we were inspired by Lakoff and Johnson's metaphor theory, in which meanings derive from metaphors, and metaphors from concrete experiences: "No metaphor can ever be comprehended or even adequately represented independently of

its experiential basis" (Lakoff and Johnson, 1981:17). And although Lakoff and Johnson work primarily with linguistic evidence, they do mention that the semantic structures which metaphors set up are "not merely a matter of the intellect" but "involve all the material dimensions of our experience, including aspects of our sense experiences: colour, shape, texture, sound, etc." (ibid: 235). Thus, I argued (Van Leeuwen, 1999: 130-131), the meaning potential of vocal (and instrumental) tension, one of the distinctive features of timbre, rests on our common experience of *what* happens when the voice tenses (it becomes higher, sharper and brighter) and *when* that happens, for instance when we are fearful, nervous – or excited, to mention just some possibilities. This range of experiences then creates a meaning potential, and how that potential will be actualized and narrowed down will depend on the context.

The same approach can be used in analysing material objects rather than embodied performances such as speech or singing. Johannessen and I explored the distinctive feature of irregularity in writing and other forms of what we called 'trace making' (Johannessen and Van Leeuwen, 2018: 186). Experience tells us, we argued, that writing can become irregular for a range of reasons – because we lack the skills needed to produce regular writing, as in the case of young children; because we refuse to produce neat writing for one reason or another; because we use tools and materials that make it difficult or impossible to produce neat writing; or because of infirmity or intoxication, and we also know that even the neatest handwriting is always more irregular and less mechanical than printing. When irregularity is deliberately produced, or part of the habitus of individuals or social groups, as it usually is, we understand what it means on the basis of such experiences and of the context. On an invitation to a children's party irregularity may mean playfulness, on the cover of a heavy metal album rebellion, on the 'handwritten' menu of an expensive restaurant unique quality and personal service, for example. In all such instances, materiality and embodiment create the *styles* that express the values which make up identities, whether the identities of individuals or social groups, or even nations.

Styles can also be recognized holistically, on the basis of 'where they come from' – and that 'place' may be an actual place but could also be a historical period or a particular social group. Their meaning potential then derives from the associations people have with that 'place' within a given context. This concept we derived from Roland Barthes' analysis of the role of 'connotation' and 'myth' in popular culture. His now famous example was an advertisement for Panzani pasta, sauce and parmesan. The name Panzani and the colour scheme

of the advertisement, based on the Italian flag, signify, said Barthes, not 'Italy' but 'Italianicity', Italian identity ("the condensed essence of everything that could be Italian") (1977: 48), as understood in a particular context, in Barthes' case France. Analysing such provenances therefore rests, not on common experience, but on recognizing cultural references, whether unreflectively, on the basis of "cultural lessons half-learnt" as John Berger put it (1972: 140) or consciously and deliberately, as for instance in the case of designers.

(6)

Has the emphasis on systems led to a neglect of meaning-making as practice?

Exploring materiality and embodiment led to our distinction between 'modes' and 'media' (Kress and Van Leeuwen, 2001). Modes we defined as resources for the functional design of texts and other semiotic artefacts. Their key characteristic is that they can be materially realized in different ways, in the way that language can be realized as speech or as writing, or in the way that narratives can be told in many different media. The design principles of visual communication we described in *Reading Images* also form a mode as they can be realized in materially different ways – as drawings, photographs, paintings and so on. *Media* are then the resources which materialize functional designs and, in the process produce the *styles* in which functional designs are executed, and the meanings and values which these styles express. The distinction between 'functional design' and 'identity design' is discussed in detail in my *Multimodality and Identity* (van Leeuwen, 2021).

Our distinction between 'modes' and 'media' led Kress and me to rethink the concept of stratification. In systemic-functional linguistics, stratification posits a hierarchical layering of phonology, lexicogrammar and discourse semantics in which each layer realizes the next: phonology or graphology realizes lexicogrammar and lexicogrammar realizes discourse semantics. Our rethinking of this concept was inspired by Goffman's theory of the 'production format' of talk, which conceives talk, not as an 'object' with layers ('strata') but in terms of different roles. The 'principal' is the person (or institution) "whose beliefs are told", "whose position is established", the 'author' the person who "selects the sentiments that are being expressed and the words in which they are encoded", and the 'animator', the "sounding box in use" (Goffman, 1981: 44). These roles can combine in one person or lead to divisions of labour. We recontextualized them in terms of the stages of communicative practices and the semiotic resources that are used in each stage. *Design* draws on discourses,

different ways of understanding aspects of reality, to select what will be communicated, and on modes to decide how, to what end and to whom it will be communicated. *Production* draws on the experiential meaning potentials and provenances of material resources and aspects of embodiment to create the styles that express the values which carry the identity of the communicator, whether that is an individual or an institution. *Distribution* draws on technologies for recording and distributing semiotic artefacts and performances which not only embody values themselves, but can also become resources for production. With Johannes Mulder (Mulder and Van Leeuwen, 2019) I explored, for instance, how contemporary amplification and recording technologies not only 'distribute' and 'record' but also modify singing voices, so contributing to the creation of new vocal styles. The relation between these stages is not hierarchical but complementary – each contributes specific kinds of meaning to the whole. And by focussing on 'stages' rather than 'layers' we foregrounded practices (which then use specific kinds of semiotic resources) rather than foregrounding resources (which are then 'instantiated' in practices).

(7)

Can the 'same' meanings, expressed through different modes, ever fully be the same?

As mentioned before, in *Reading Images* we assumed that cultures are constituted both by the ways in which they make meaning of the world and of people's place in it, and by the ways they use modes and media to express and communicate these meanings, so that, for instance, some meanings can be expressed verbally and visually and musically, others only verbally, and yet others only visually, or only musically, and so on. To some extent this continues to be true. As I argued in *Introducing Social Semiotics* (Van Leeuwen, 2005: 112), the 'special operations' discourse of war, in which small teams of highly trained commandos execute quick and effective operations such as extracting hostages or 'taking out' terrorist leaders, can take the form of newspaper reports, novels, movies, computer games, etc., all using different combinations of semiotic resources and serving different communicative functions. Yet Machin and I noted differences between, for instance, films and computer games realizing this discourse (Machin and Van Leeuwen, 2007: 74 ff). The same applies to the 'lexicogrammatical' resources of different modes. We had, for instance, assumed that all semiotic modes have resources for expressing modality, but that they do so in different ways – language by

means of modal auxiliaries and the lexis of probability and frequency; images through the degree to which they represent visual detail, depth, gradations of colour, light and shade, and so on; other modes in yet other ways. But visual modality differs from linguistic modality not only in the way it is expressed, but also in the *kind* of truth criteria these expressions realize. Visual truth is based on the concept of realism, and realism is not quite the same thing as probability or frequency. Concepts of this kind have cultural histories. Ever since the Renaissance, visual modality has been based on perceptual criteria – on how much an image looks like what we think we would see if we saw what it depicts in reality. Hence the importance of visual detail, depth, gradations of colour, light and shades and so on in its expression – aspects of visual representation which Renaissance artists like Da Vinci developed, and which later became essential in photography.

Visual modality is therefore also closely tied up with technology. The introduction of Photoshop in 1989 led to a debate about the truth criteria of photography which had been based on its 'indexicality', on the chemical, hence causal relation between what is in front of the camera and what will appear on the photograph. People now began to call for "a truer truth, one closer to conceptual adequacy", which does not rely on the "illusionistic surface of (and even definition of) reality" (Rosler, 1991: 58).

This can and should lead to considering the differences between modality in speech and in writing. More than 40 years ago, Kress and Hodge (1979: 129) already drew attention to embodied expressions of modality in speech: "There are a large number of ways of realising modality: non-verbal and verbal, through non-deliberate features (hesitations, *ums*, *ers*, etc.) and deliberate systematic features", and I have already mentioned the close integration between speech and writing in PowerPoint presentations.

(8)

Clearly, in my practice as a 'multimodalist', I draw not only on the analytical concepts and methods of linguistics, but also on the concepts and methods of disciplines that have a long experience in analysing the semiotic modes I write about and in documenting their cultural histories.

When editing a collection of papers on different approaches to visual analysis, Jewitt and I (van Leeuwen and Jewitt, 2001) commissioned a chapter from two cultural studies scholars. In contrast to contributors from fields such as social psychology, linguistically inspired semiotics and anthropology,

they confessed to a problem with the brief we had given them, which asked contributors to outline and exemplify their discipline's analytical approach. They saw their field as defined, not by a method of analysis, but by a particular set of questions, each requiring different methods, "a compound field rather than a discrete discipline, appropriating and re-purposing elements of theoretical frameworks and methodologies whenever that seems productive" (Lister and Wells, 2001: 63). This needs to be the approach in multimodality studies, too, although it can be formulated in a less ad hoc manner, as a new kind of quadrivium (literally, 'the place where four roads meet') comprising social theory, semiotics, ethnography and cultural history.

It includes social theory, because, as Cicourel (1964) already argued, any form of empirical work must first of all link to social theory, both in the interest of the self-reflexivity that is vital for researchers of social actions and its products, and because theory inevitably plays a role in the interpretation of empirical data.

It includes semiotics because semiotics can offer systematic approaches to the analysis of texts and other semiotic artefacts or practices – and it is here that the experience of linguistics and rhetoric in analysing and understanding speech and writing remains a vital resource, and one that therefore needs ongoing development and refinement in its own right.

It includes ethnography, because semiotic practices must be understood in their contexts. Text and talk can tell us much, but they do not necessarily reveal all aspects of the practices of which they are part, or tell us how these practices are understood by the participants, taught, regulated, critiqued, changed, and so on.

And it includes cultural history to help us understand how and why semiotic practices come into being, and why the semiotic resources they use take the form they do.

In addition, it is important to understand the material basis of semiotic media, and, today, the digital technologies that play an increasingly important role in semiotic practices.

This makes disciplines complementary. No discipline can adequately address any given problem on its own, and hence disciplines can no longer function as traditional professions, with the autonomy to define what will count as a research problem and how it will be addressed. Disciplines now become resources for research, contributing theoretical insights, analytical skills or observational experience. Which does not mean that they should not

work on further developing these resources – but with accountability to the multi-disciplinary alliances in which they partake.

All this is not a new insight in semiotics. In 1934 Mukařovský criticized Victor Sklovskij's *Theory of Prose*, in which Sklovskij used weaving as an analogy and argued that textual analysis should merely be interested in the type of yarn and the techniques of weaving and not in the situation on the international wool market or in the politics of trusts. To which Mukařovský replied (as recounted in Matejka, 1976: 270) that the techniques of weaving "necessarily reflect the needs and pressures of the market in accordance with its law of supply and demand, and that the techniques of art cannot be studied in a vacuum, artificially isolated from the related systems and, implicitly, from time and space." Language ought to be analysed, according to Mukařovský, "not only with regard to its autonomous properties but also in its relationship to […] structures such as science, politics, economy, social stratification, language, ethics, religion, each having its own immanent development".

Notes

1. Arnheim defines volumes as the "masses" with a distinct "weight" and "gravitational pull" which represent the main people, places and things (including abstract things) in images. Today they can be identified automatically through techniques such as 'edge detection'. Vectors are (usually diagonal) depicted elements such as outstretched arms or other elements with a sense of direction (e.g. arrows) which connect volumes . As Arnheim summarized their function in images: "We shall distinguish between volumes and vectors, between being and acting" (1982: 154)

References

Arnheim, R. (1982) *The Power of the Center.* Berkeley and Los Angeles, Ca: University of California Press

Barthes, R. (1967) *Elements of Semiology*. New York: Hill and Wang.

Barthes, R. (1973) *Mythologies.* London: Paladin

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

Berger, J. (1972) *Ways of Seeing*. Harmondsworth: Penguin

Birdwhistell, R.I. (1970) *Kinesics in Context: Essays on Body Motion Communication*, Philadelphia: Univ. of Pennsylvania Press

Boeriis, M. and Holsanova, J. (2012) Tracking visual segmentation: connecting semiotic and cognitive perspectives, *Visual Communication* 11(3): 259-281

Bühler, K. (1965 [1934]) *Sprachtheorie – Die Darstellungsfunktion der Sprache*. Stuttgart: Gustav Fischer Verlag

Caple, H. (2013) *Photojournalism – A Social-Semiotic Approach*. London: Palgrave Macmillan

Cicourel, A. (1964) *Method and Measurement in Sociology*. New York: The Free Press.

Classen, C. (2001) Foundations for an anthropology of the senses. *International Social Science Journal* 153: 401-412

Djonov, E. and Van Leeuwen, T. (2014) Bullet points, new writing and the marketization of public discourse. In E. Djonov and S. Zhao, eds. *Critical Multimodal Studies of Popular Discourse*. London: Routledge. pp. 232-250.

Elleström, P. (2019) *Transmedial Narration: Narratives and Stories in Different Media*. London: Palgrave Macmillan

Goffman, E. (1981) *Forms of Talk*. Oxford: Blackwell

Halliday, M.A.K. (1985) *An Introduction to Functional Grammar*. London: Arnold.

Hassler-Forest, D. and Nicklas, P. eds (2015) *The Politics of Adaptation: Media Convergence and Ideology.* London: Palgrave Macmillan.

Honzl, J. (1976 [1941]) Dynamics of the Sign in the Theatre, in L. Matejka and I.R Titunik, eds, *Semiotics of Art – Prague School Contributions*. Cambridge, Mass: MIT Press. pp 74-93

Jakobson, R. (1960) Closing Statement: Linguistics and Poetics. In T.A. Sebeok, ed. *Style in Language*. Cambridge, Mass: MIT Press, pp; 350-377

Jakobson, R. and Halle, M. (1956) *Fundamentals of Language*. Mouton: The Hague.

Jenkins, H. (2008) *Convergence Culture: Where Old and New Media Collide*. New York: New York University Press.

Johannessen, C.M. and Van Leeuwen, T. (2018) (Ir)Regularity. in C.M. Johannessen and T. van Leeuwen, eds. *The Materiality of Writing – A Trace-Making Perspective*. London: Routledge. pp. 175-192

Kress, G. (2011) *Multimodality – A social semiotic approach to contemporary communication*. London: Routledge.

Kress, G. and Hodge, R. (1979) *Language as Ideology*. London: Routledge and Kegan Paul.

Kress, G. and Van Leeuwen, T. (1990) *Reading Images*. Geelong: Deakin University Press.

Kress, G. and Van Leeuwen, T. (2001) *Multimodal discourse – The modes and media of contemporary communication.* London: Arnold.

Kress, G. and Van Leeuwen, T. (2002) Colour as a semiotic mode: Notes for a grammar of colour: *Visual Communication* 1(3): 343-368

Kress, G, and Van Leeuwen, T. (2021) *Reading Images – The Grammar of Visual Design*. 3rd Edition. London: Routledge

Lakoff, G. and Johnson. M. (1980) *Metaphors We Live By.* Chicago: Chicago University Press

Lister, M. and Wells, L. (2001) Seeing beyond belief: Cultural studies as an approach to analysing the visual. In Van Leeuwen, T. and Jewitt, C. eds. *Handbook of Visual Analysis*. London: Sage. pp. 61-91

Machin, D. and Van Leeuwen, T. (2007) *Global Media Discourse – A Critical Introduction*, London: Routledge.

Matejka, L. (1976) Postscript: Prague School Semiotics. In L. Matejka and I.R. Titunik, eds. *Semiotics of Art - Prague School Contributions*. Cambridge, Mass: MIT Press. pp. 265-290

Matejka, L. and Titunik, I,R, eds (1976) *Semiotics of Art – Prague School Contributions*. Cambridge, Mass: MIT Press.

McClary, S. (1991) *Feminine Endings – Music, Gender and Sexuality*. Minnesota: University of Minnesota Press

Metz, C. (1974) *Film Language – A Semiotics of the Cinema* New York: Oxford University Press.

Mukařovský, J. (1976 [1936]) Poetic reference. In L. Matejka and I.R. Titunik, eds *Semiotics of Art- Prague School Contributions*. Cambridge, Mass: MIT Press. pp. 155-163

Mulder, J. and Van Leeuwen, T. (2019) Speech, sound, technology. In Grimshaw-Aagaard, M., Walther-Hansen, M. and Knakkergaard, M. eds. *The Oxford Handbook of Sound and Imagination*. Vol I, Oxford: Oxford University Press. pp. 474-492.

Norris, S. ed. (2016) *Multimodality.* Vol I-IV. London: Routledge

O'Toole, M. (1994) *The Language of Displayed Art*. Leicester: Leicester University Press.

Ravelli, L. and McMurtrie, R.J. (2016) *Multimodality in the Built Environment – Spatial Discourse Analysis*. London: Routledge.

Ravelli. L. and Van Leeuwen, T. (2018) Modality in the digital age. *Visual Communication* 17(3): 277- 298

Rosler, M. (1991) Image simulations, computer manipulations, some considerations. *Ten-8* 2(2): 52-63

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

Van Leeuwen, T. (2005) *Introducing Social Semiotics*. London: Routledge

Van Leeuwen, T. (2006) Towards a semiotics of typography. *Information Design Journal* 14(2): 139-155

Van Leeuwen, T. (2011) *The Language of Colour – An Introduction*. London: Routledge.

Van Leeuwen, T. (2015) Looking good: Aesthetics, multimodality and literacy studies. In J. Rowsell and K. Pahl, eds. *The Routledge Handbook of Literacy Studies*. London: Routledge. pp. 426-239

Van Leeuwen, T. (2021) *Multimodality and Identity*. London: Routledge

Van Leeuwen, T. and Jewitt, C. (2001) *Handbook of Visual Analysis*. London: Sage

Zhao, S. Djonov, E. and Van Leeuwen, T. (2014) Semiotic Technology and Practice: A Multimodal Social Semiotic approach to Powerpoint. *Text & Talk* 34(3): 349-375