

National Centre for Research Methods Working Paper

05/14

Multimodal analysis of aesthetics: Two versions of a food blog compared

Elisabetta Adami, MODE node, Institute of Education

Multimodal analysis of aesthetics: Two versions of a food blog compared

Elisabetta Adami, Università G. D'Annunzio Chieti-Pescara

Introduction

Even when considering visuals besides language, current research on digital texts generally focuses its analysis on contents. Investigations on the visuals of a website usually involve counting, coding and analysing the images that are published as contents along with the written texts (e.g., Carpenter 2010; Jeffres and Lin 2006; Kenix 2009; Kim, Coyle, and Gould 2009). Among qualitative studies in the social sciences, in communication and media studies especially, when it comes to analysing visuals, content analysis seems to be the preferred methodology (Kim and Weaver 2002; Peng, Zhang, Zhong, and Zhu 2013). Even more, irrespectively of the different theoretical perspectives or methodologies adopted (for an overview, van Leeuwen and Jewitt 2001), most visual analyses on texts focus solely on their contents, i.e., still/moving images (for an example in multimodal studies, Knox 2009).

Yet, websites have meaningful form beyond their posted texts; they make meaning well before we engage with their posted contents. Visual resources such as layout, colours and fonts concur to shape the overall aesthetics of a webpage beyond (and before) its contents. By “merely” looking at a webpage, we can have an idea not only of its genre, but also of its style, the kind of intended readers it addresses and the voice with which they are addressed, e.g., as peers, as a general public addressed by professionals, as learners by teachers, as children by adults etc. (cf. the notion of “implied reader” in Iser 1974 and “model reader” in Eco 1979, both limiting their scope only to written texts). The overall design of a webpage shapes its aesthetics, answering the question “what does this look like?” “what kind of user is this text for?” “what kind of author made this?”.

By drawing on data of an on-going project on *Methodologies for Multimodal and Narrative Analysis of UK Food Blogs*,¹ the present chapter discusses the role of visual modes in shaping the aesthetics of digital texts and provides methodological indications to its analysis.

Multimodality and digital texts

¹ The collaborative project (Principal Investigator: G. Kress) is funded by the UK National Centre for Research Methods (NCRM) of the Economic and Social Research Council (ESRC) and it involves two Nodes: MODE (Multimodal Methodologies for Researching Digital Data and Environments) and NOVELLA (Narratives of Varied Everyday Lives and Linked Approaches); cf. <http://mode.ioe.ac.uk/2012/09/16/multimodal-analysis-of-food-blogs/>

All communication is multimodal (Kress 2010); in all instances of communication we make meaning through signs produced in a variety of semiotic modes, i.e., organized sets of resources having distinct affordances, historically developed to make meaning. This holds not only for face-to-face communication, but also, and increasingly so, for “written” texts.

Thanks to contemporary technological advances (cf. digital technology, which treats all textual output equally as bytes) texts traditionally conceived as written are increasingly multimodal. Not only do they frequently include images, videos and audio files along with writing, but they also use colour, fonts, shapes and the overall layout of the page as meaningful resources.

Visual modes in webtexts

In digital texts, sign-making resources relying on visual perception include the modes of layout, colour, font, (still and dynamic) image and writing. Layout (Kress 2010) is the overall arrangement of the elements on a page; it makes meaning through framing and positioning; framing encompasses all devices that connect and separate elements; positioning refers to the relative position of the elements on a page. Colour (Kress and van Leeuwen 2002) traverses all other modes on a page; it is used in framing devices, in images and fonts; modal elements of colour include palette and effects, such as lighting, saturation and nuancing. Font (Bezemer and Kress 2008) is the material shape of writing and makes meaning through font type, size and effects such as bold, italics and capitalization. Image can be still or dynamic; it can be a photograph, a drawing, a symbol/shape, a diagram (often combining image and writing) or a video (which combines auditory resources too); image makes meaning through framing, modality, camera angle and others (cf. Kress and van Leeuwen 1996[2006]). Writing is verbal language relying on visual perception, thus using visual-specific resources, such as spelling, punctuation and paragraphing/pagination; its other resources are typical of verbal language, such as vocabulary and syntax (for a social semiotic take on writing, Kress 2003).

In multimodal studies (for an overview, Jewitt 2009), the combination of modes on a page has been largely investigated with reference to Halliday’s (1978) three metafunctions. The present chapter analyses modal resources for their role in shaping the aesthetics of a text, thus locating a text in terms of “taste”, as discussed in the next section.

A social semiotic perspective on aesthetics

Kress (2010) defines aesthetics as the “politics of style”, with style as the “politics of choice”. Aesthetics expresses and reveals power relations existing in society, naturalized as “taste”. It locates a

text socially; in its labelling (e.g., minimalist, kitsch etc.) it condenses a recollection of characteristics which reveal a set of identity features projected by the text onto its author and addressed audience. Bourdieu (1979[1986]) has shown that taste is a social product. Differences in taste, or aesthetic preferences, are regularly distributed among social classes and among sub-groups sharing specific agglomerates of cultural capital, encompassing life-style, class, gender, age, profession and education. In Bourdieu's terms, in classifying its object, taste classifies its classifier. Aesthetic judgment is the product of someone's affiliation and belonging to a given social group and its values. Hence an analysis of a text's aesthetics can reveal also the observer's social positioning with reference to a given set of values existing within society. Yet very little has been attempted to single out the elements concurring to shape a text's aesthetics as we perceive it, as well as the role played by each of them in doing so.

A multimodal analysis of aesthetics

Fig.1 shows the homepage of the UK food blog *The Diary of a Frugal Family* (www.frugalfamily.co.uk), as it appeared in October 2013. Fig.2 shows the July 2009 version of the same blog as it appeared when it was first archived in *The Wayback Machine* Internet Archive (archive.org.web; for its uses in Internet research Murphy, Hashim, and O'Connor, 2007). Both images combine the first two vertical screens as visible on a 13" screen (browser's viewing options as medium size).



Fig. 1. *The Diary of a Frugal Family* homepage (Retrieved Oct 23 2013).

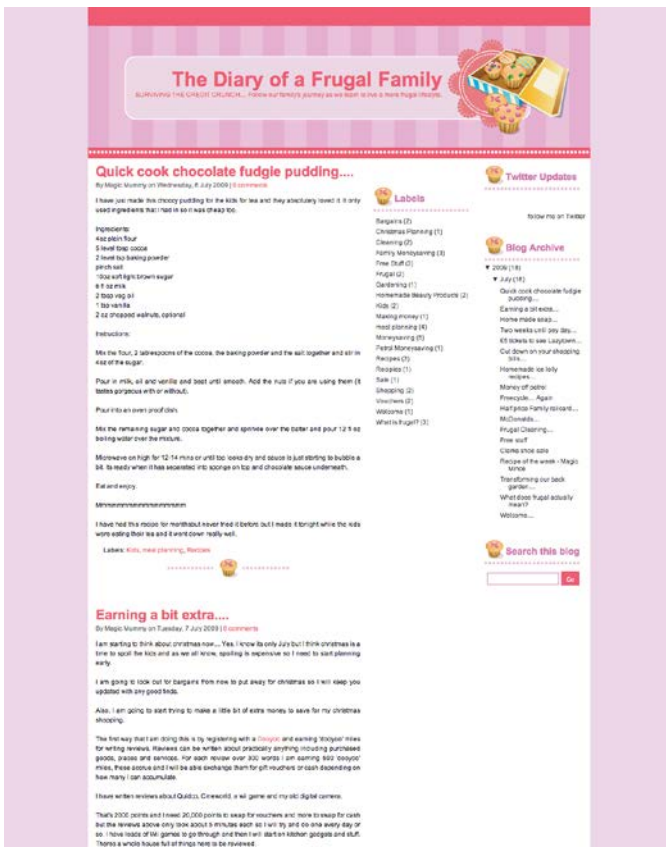


Fig. 2. *The Diary of a Frugal Family* homepage on Jul 10 2009 (Retrieved Oct 23 2013 from *The Wayback Machine*).

While the more recent version (Fig.1) is published via *Wordpress* (www.wordpress.com), the older version (Fig.2) was published on *Blogger* (www.blogger.com). Both blog providers offer several customizable templates that bloggers can apply to their blog. Notwithstanding the fact that in both cases the blogger has selected and customized a pre-given template rather than designing it from scratch, in a social semiotic perspective the different aesthetic values between the two is no less meaningful. Choice is indeed meaningful; since the advent of mass-production and consumption, sign-making through selection and assemblage is frequent in most semiotic domains, in our clothing style and home decoration no less than in digital text production. As any sign-maker chooses a given sign out of his/her available resources as the most apt to represent the criterial aspects of what s/he wants to communicate (Kress 2010), when choosing a template, bloggers select a series of combinations which shape their blog's aesthetics as the most apt for their sign-making interests.²

When looking at the pages reproduced in Fig.1 and 2, even when not reading (or understanding) the written text or the content of the images, we can immediately make inferences on the text type and style, thus deriving assumptions on the author, her "taste" and design skills. We can also make assumptions on the type of audience addressed by the text, i.e., those who align with its aesthetics. Depending on the observer's social positioning, assumptions might differ, thus revealing the observer's taste and social positioning too. Although we might not always be aware of the distinction, aesthetic assumptions (which we all make when engaging with a text) refer to what viewers infer, not to what the text really is or is meant to communicate, let alone to what kind of identity, skills or taste its author really has. In this regard the above mentioned notion of "model/intended" author/reader, distinguished from the "real" one is crucial. As an example, a text that looks "amateur" can be designed professionally to look precisely as such; so we might interpret its "model" author as amateur despite its "real" author being a professional, as in the case of many TV commercials professionally designed to look amateur to communicate "authenticity/lack of sophistication", and hence to seem more sincere and honest, in their shaping a relation with viewers as peers rather than as professional seller to prospective buyers. Then, when it comes to the observer, an amateur aesthetics might be either positively or negatively evaluated on the basis of the observer's affiliation with – say – more or less horizontal systems of power relations (in the case of the observer trusting and belonging to a peer-to-peer model or to an expert-to-general public one), for example. The crucial issue is that visual design conventions are naturalized so we might not be aware of the underlying mechanisms of the aesthetic meanings which we make. Analysis can help making these conventions explicit.

² Clearly, further ethnographic research could then investigate the reasons that triggered the blogger to abandon one template and adopt another.

Methodological caveats

As discussed in detail in Adami (2014), in order to determine analytically the role of each modal element in shaping a text's aesthetics, its configuration needs to be assigned a qualifier. The assignment of a qualitative descriptor (an adjective, in verbal language) to a modal resource answers the question "what does it look like?" rather than "what is it?", hence describes its aesthetic value (rather than ideational meaning). Following the framework in Adami (2014), descriptors assigned will be first as objective and measurable as possible; then, a metaphorical/analogical process will derive more socially-specific evaluative values, thus hypothesising a range of identity features possibly expressed by the modal element in the text. The descriptors assigned analytically to each modal element need to be combined in a later synthesis, to derive the aesthetic meaning potential of the overall text.

The evaluative descriptors derived from the objectively observable ones refer to *designed* identities projected by texts onto their *model* authors and audience. They might be stereotyped identity features. They might be unintended or might have been carefully selected and designed for communicative purposes (as mentioned earlier, an amateur aesthetics might be designed to fulfil certain communicative needs, e.g. to communicate authenticity).

As taste indicators, evaluative descriptors are socially and culturally specific. As an example, the here-below reading of a UK food blog is seen with Italian eyes (just to mention nationality); it clearly might differ considerably to a Brazilian eye. This could provide insights into culture- and society-specific differences in taste and social judgments, in terms of alignments and affiliations with more or less globalized Anglo-Saxon aesthetic sign-making practices. Even more, evaluative descriptors might differ among readers; the results of the analytical process could provide interesting outcomes when tested against readers' perception or web designers' goals. In this light, the subjective component inevitably implied in the qualifying process, when acknowledged and carefully considered, might constitute one of its strengths and analytical tools.

Analysis: Two page versions compared

Colour

In both versions, colour is used to frame and identify specific areas of the page (masthead, content page, background) and of the verbal text (titles and sub-titles in different sections). In Fig.1, the colour palette is highly diversified, encompassing green, pink, light-blue, light and dark brown, white

(for the page) and black (for the body text), with other colours used in minor proportions in images and drawings. The palette is wide, using unrelated colours, with pastel as their only shared feature. So, “unrelated”, “colourful” (for a wide palette) and “pastel” can be assigned as observable descriptors. By analogy, “unrelated” can be socially evaluated as “non-cohesive”, while “colourful” as “playful”. Given the provenance (Kress and van Leeuwen 2001) of pastel colours as used in objects for infants, “infant” might be derived as a social value for the “pastel” descriptor.

In contrast, the colour palette of the older version of the blog (Fig.2) is rather “attuned/nuanced”, restricted to different nuances of pink (plus the light brown of the cupcakes, and black/white for the written text/background). In more evaluative terms, “nuanced” can be assigned “cohesive”, while socially, pink has a “feminine” association, with pastel pink additionally connoted as “romantic” in Western cultures, while the more saturated nuance having a more “determined” character. In other words, while Fig.1’s use of colours might shape aesthetically the blog as “non-cohesive”, “playful” and “infant”, Fig.2’s use of colours might shape it as “cohesive”, “romantic but determined”, and “feminine”.

Layout

In both blog versions, besides the horizontally-oriented masthead, all other elements are arranged in three vertically-oriented columns, suggesting a linear (top to bottom) reading path. The two versions differ in the functions played by each column, with the central one devoted to blog posts in Fig.1 vs. the left one in Fig.2.

In Fig.1, within each column no lines or regular shapes are used for framing elements together and separating them from others, while their central-alignment within each column, contrasting with the left-alignment of typed titles and body text, gives the white page an irregular outline. In contrast, Fig.2 uses left-alignment consistently for all elements in the three columns, and an horizontal dotted line with cup-cake drawing as framing device to separate posts, cohesively tuning with the blog colour palette and the cupcakes used as signposts in the other columns. This confers the page a high level of regularity, orderliness and cohesion. Finally, the page in Fig.1 is rather densely filled with elements, while the one in Fig.2 has wide areas left unoccupied by signs. Hence layout can be assigned “linear”, “central”, “irregular” and “dense” as measurable descriptors for Fig.1 vs. “linear”, “left”, “regular” and “spaced” for Fig.2.

As socially derived descriptors, a linear orientation recalls by analogy “old” types of written text, when compared to the modular orientation of contemporary web designs (Bezemer and Kress 2009). Within it, a left-right orientation recalls traditional writing patterns (left-to-right in western conventions)

more closely than the centre-margins one, thus, comparing the two versions, orientation shapes Fig.2 as “older” than Fig.1. Clearly, the layout orientation of a webpage does not prevent readers from choosing alternative reading paths; yet it does shape an aesthetics that recalls more traditional forms of writing.

“Irregularity” might be indicative of “inconsistency” but also “movement”, while a “dense” page is an indication of “economy”, within societies where space is a valuable commodity, thus it might be an indication of class (cf., by contrast, “minimalism” associated to upper classes). Instead, layout in Fig.2 is “regular->consistent/orderly”, but also “rigid, fixed”, and “spaced->wealthy”. As anticipated, any negative or positive connotations attributed to descriptors will depend on the observer’s social affiliation and positioning to the values of a given social group.

Font

Fig.1 combines different fonts, both serif and sans-serif, in different sizes and colours, in the masthead, titles, body-text and images. Its highly differentiated use of fonts contrasts sharply with the rigorously homogeneous use of the same sans-serif font in Fig.2, differentiated only in colour and size for (functionally different) titles. In line with colour descriptors, “differentiated” might derive “non-cohesive” or “playful” and could be socially associated to texts for children, often highly diversified in both fonts and colours. Furthermore, against a “less is more” minimalist and strictly function-related use of font in the ongoing conventions of professional web design practices (e.g., the recommendations in Williams, 2012), a highly diversified use of font features recalls the early attempts at familiarizing with the mode, typical of beginners, of non-expert/amateur web designers, who explore all newly-discovered affordances and potentials. Thus while Fig.1’s aesthetic value for font is “differentiated->non-cohesive/playful ->amateur/beginner/children”, Fig.2’s is “unified->cohesive->professional/expert”.

Images

Both versions use drawings; the newer one (Fig.1) uses also photos, which are completely absent in Fig.2. Fig.1’s more frequent use of drawings (the cupcakes) and drawing-like symbols (cf. the social networking site buttons in the right column) than photos reinforces the “infant/child-like” descriptor of colour and font.

The photos in the blog post in Fig.1 do not deploy the modality of professional photography; rather, the fuzzy effects, the flashed eye of the dog in the second photo, and the seemingly careless frames, as if shot in haste, in a casual and unplanned way, reinforce the “amateur/unprofessional” descriptor of font.

Fig.2's exclusive presence of drawings might give a more "childish" aesthetics to the page and a sense of "non-serious/-professional/-adult" content of the page (further associated stereotypically with the "romantic" and "feminine" descriptors of colour), with the absence of photos possibly reinforcing the "older" descriptor of its left-oriented layout.

Writing

Both versions have one-plus-triple-dot punctuation in their post titles, suggesting "informality" and "immediacy", while the standard use of capitalization denotes a non-internet-specific usage, thus a "non-insider" status in the internet community. The first-person writing style confers a "personal" character to both versions.

In the blog post in Fig.1, sentences are quite long, with a syntax recalling spoken language (normalized for writing, without any semantic/syntactic plan changes, hesitations or repetitions), again communicating immediacy and "speed/haste" in production, especially for a prevalence of coordination over subordination and lack of punctuation; cf. the second and third sentence:

But by Sunday night we needed to get out and do something so we put our willies and coats on since it was looking a bit grey outside and headed to the park for a conker hunt. The park where all the good conkers can be found is about 15 minutes drive away and I'm not exaggerating when I say that by the time we go there the rain was bouncing and the sky was so dark you'd think the end of the world was about to happen.

"Informality" is reinforced by the use of emoticons in the post opening and closing, by vocabulary (e.g. *a do nothing weekend, wellies, get out, outdoorsy, we were going to be going home, stay tuned*), phono-morphology (Subj+Aux contractions) and syntax (sentences starting with *but*, lack of tense agreement in *by the time we go there the rain was bouncing*), with occasional typos (*the best thing about it was the hot chocolate and marshmallows when we had when we got home!*) further confirming the "speed/haste" descriptor.

Also the two posts in Fig.2 give an impression of "informality" (e.g. *choccy* in the first sentence, *Mmmmmmmmmmmmmmmmmmmmm* at the end of the first post), "haste/speed" and "immediacy". Although the first post is carefully paragraphed as a recipe, "immediacy" is conferred by the occasional typos (non-spaced *4oz plain flour* and *10oz soft* against the spacing between number and unit of measurement in the rest of the list of ingredients; *Its ready when it has separated*; non-spaced *monthsbut*); intra-

sentence punctuation is virtually absent, along with a speech-like syntax, e.g. in the last sentence (cf. also the double coordination with *but*):

I have had this recipe for monthsbut never tried it before but I made it tonight while the kids were eating their tea and it went down really well.

In the second post in Fig.2, a speech-like syntax is present since the very beginning (*Yes, I know*), with informality closing it (*Have a go....*); then a missing *to* (in *I will be able exchange them for*) pairs with apostrophe dropping occurring twice (*Yes, I know its only July*, and *Theres a whole house full of things here*). Both occurrences of *christmas* have no capitalized initials; this and apostrophe dropping comply with ongoing internet spelling conventions and might contrast (or mitigate) the “non-insider” descriptor assigned earlier to standard capitalization.

The whole picture

Table 1 assembles the list of observable and evaluative descriptors assigned for each modal element to the two versions of the page. A stage of synthesis can then assign different weight to each descriptor in relation to the others, with conflicting descriptors serving to mitigate, nuance differently or create some dissonance within the overall aesthetic value.

Mode	Modal element	Fig. 1 (more recent version)		Fig. 2 (older version)	
		Observable Descriptor	Evaluative Descriptor	Observable descriptor	Evaluative Descriptor
Colour	Palette	Diversified/colourful	Playful/non-cohesive	Nuanced / attuned	Cohesive
	Effect	Pastel	Infant	Pink (pastel-and-saturated)	Feminine (romantic-and-determined)
Layout	Position	Linear (Central)	Old (less)	Linear (left)	Old (more)
	Framing	Dense Irregular	Economy / lower class Inconsistency / movement	Spaced Regular and cohesive	Wealthy / upper class Consistent / orderly /polished /rigid/fixed
Font	Type and effect	Differentiated	Playful/non-cohesive	Homogeneous	Cohesive / polished
		Varied	Amateur/beginner (children)	Unified	Professional/expert (adults)
Image	Type	+ drawings	Infant	Drawings only	Childish (romantic?)
	Modality	Fuzzy, imprecise	Unplanned, casual Amateur	Less multimodal (no photos)	Old Less adult, pro and serious
Writing	Visual and language	Informal	Down-to-earth	Informal	Down-to-earth
		Immediate	hectic	Immediate	hectic

resources	Standard (capit.) Personal	non-insider Sincere and Authentic	Standard and internet specific (capit.) Personal	(less) non-insider Sincere and Authentic
-----------	-------------------------------	--------------------------------------	---	---

Table 1. Synoptic view of aesthetic values assigned.

Conclusions

The comparative analysis of the aesthetic meaning potential of two versions of a webpage has attempted at making explicit the naturalized social conventions that underlie aesthetic judgement. It has shown that a text's visual meaning is expressed by the overall multimodal orchestration of a page rather than merely by images.

Far from having any normative intent, the methodology exemplified in the analysis can be useful in revealing existing power relations and dominant sets of evaluations (what is deemed as appropriate/trendy with reference to specific social groups). This can contribute to raise sign- and meaning-makers awareness of internalized conventions, social evaluations and (pre-)judgments that are often naturalized in visual sign-making practices.

Sociolinguists and critical discourse analysts have long pursued this line of research for language; in an increasingly visual world of representation and with today's exponential growth in the number of sign-makers having access to digital text design, production and dissemination, the development of analytical tools that make aesthetic sign-making conventions explicit can hardly wait further to enter a visual literacy agenda.

References:

- ADAMI, E. (2014). "Aesthetics and identity in digital texts beyond writing: A social semiotic multimodal framework", in: A. ARCHER and E. BREUER (Eds.), *Multimodality in Writing. The state of the art in theory, methodology and pedagogy*. Bingley: Emerald Group Publishing Limited.
- BEZEMER, J., and KRESS, G. (2008). "Writing in Multimodal Texts: A Social Semiotic Account of Designs for Learning". *Written Communication*, vol. 25, pp. 166–195.
- BEZEMER, J., and KRESS, G. (2009). "Visualizing English: a social semiotic history of a school subject". *Visual Communication*, vol. 8 n.º 3, pp. 247–262.

- BOURDIEU, P. (1979[1986]). *Distinction: A Social Critique of the Judgement of Taste*. London: Routledge.
- CARPENTER, S. (2010). A Study of Content Diversity in Online Citizen Journalism and Online Newspaper Articles. *New Media and Society*, OnlineFirst (February 9), pp. 1–21. doi:10.1177/1461444809348772
- ECO, U. (1979). *Lector in fabula. La cooperazione interpretativa nei testi narrativi*. Milano: Bompiani.
- HALLIDAY, M. A. K. (1978). *Language as Social Semiotic: The Social Interpretation of Language and Meaning*. London: Arnold.
- ISER, W. (1974). *The Implied Reader: Patterns of Communication in Prose Fiction from Bunyan to Beckett*. Baltimore MD: Johns Hopkins UP.
- JEFFRES, L. W., and LIN, C. a. (2006). "Metropolitan Websites as Urban Communication". *Journal of Computer-Mediated Communication*, vol. 11, n.° 4, pp. 957–980.
- JEWITT, C. (2009). *The Routledge Handbook of Multimodal Analysis*. London: Routledge.
- KENIX, L. J. (2009). "Blogs as Alternative". *Journal of Computer-Mediated Communication*, vol. 14, pp. 790–822.
- KIM, H., COYLE, J. R., and GOULD, S. J. (2009). "Collectivist and Individualist Influences on Website Design in South Korea and the U.S.: A Cross-Cultural Content Analysis". *Journal of Computer-Mediated Communication*, vol. 14, n.° 3, pp. 581–601.
- KIM, T. S., and WEAVER, D. (2002). "Communication Research About the Internet: a Thematic Meta-analysis". *New Media and Society*, vol. 4, n.° 4, pp. 518–538.
- KNOX, J. S. (2009). "Punctuating the home page: image as language in an online newspaper". *Discourse and Communication*, vol. 3, n.° 2, pp. 145–172.
- KRESS, G. (2003). *Literacy in the New Media Age*. London: Routledge.
- KRESS, G. (2010). *Multimodality. A social semiotic approach to contemporary communication*. London: Routledge.
- 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*, vol. 1, n.° 3, pp. 343–369.
- KRESS, G., and VAN LEEUWEN, T. (1996[2006]). *Reading Images. The Grammar of Visual Design* (2nd ed.). London: Routledge.
- MURPHY, J., HASHIM, N. H., and O'CONNOR, P. (2007). "Take Me Back: Validating the Wayback Machine". *Journal of Computer-Mediated Communication*, vol. 13, n.° 1, pp. 60–75.

PENG, T.-Q., ZHANG, L., ZHONG, Z.-J., and ZHU, J. J. (2013). "Mapping the landscape of Internet Studies: Text mining of social science journal articles 2000–2009". *New Media and Society*, vol. 15, n.º 5, pp. 644–664.

VAN LEEUWEN, T., and JEWITT, C. (2001). *Handbook of Visual Analysis*. London: Routledge.

WILLIAMS, R. (2012). *The Non-Designer's InDesign Book*. Berkeley: Peachpit Press.