Understanding semiotic technology in university classrooms
a social semiotic approach to PowerPoint-assisted cultural studies lectures
Zhao, Sumin; van Leeuwen, Theo

Published in:
Classroom Discourse

DOI:
10.1080/19463014.2013.859848

Publication date:
2014

Citation for published version (APA):

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

Take down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.
Semiotic technology and practice:  
a multimodal social semiotic approach to PowerPoint

Abstract: The ubiquitous software PowerPoint has significant influence on evaluations of professional and academic success, and has attracted considerable attention from both social commentators and researchers in various fields. Existing research on PowerPoint considers the software, slideshows created with it, and PowerPoint-supported presentations in isolation from each other and is therefore unable to promote better understanding of the interaction between the software’s design and its use.

This article proposes a model for exploring this interaction. Specifically, it introduces a multimodal social semiotic approach to studying PowerPoint as a semiotic practice comprising three dimensions – the software’s design, the multimodal composition of slideshows, and their presentation – and two semiotic artefacts, the software and the slideshow. It discusses the challenges each dimension presents for discourse analysis and social semiotic research, focusing especially on the need to step away from the notion of text and to develop a holistic, non-logocentric, and adaptive multimodal approach to researching semiotic technologies. Using PowerPoint as a case study, this article takes a step toward developing a social semiotic multimodal theory of the relation between semiotic technologies, or technologies for making meaning, and semiotic practices.

Keywords: PowerPoint, semiotic technology, semiotic practice, software, slideshows, slideshow presentations

*Corresponding author: Sumin Zhao: Faculty of Arts and Social Sciences, University of Technology, Sydney, Australia. E-mail: Sumin.Zhao@uts.edu.au

Emilia Djonov: Faculty of Human Sciences, Macquarie University, Sydney, Australia.
E-mail: emilia.djonov@mq.edu.au

Theo van Leeuwen: Faculty of Arts and Social Sciences, University of Technology, Sydney, Australia & Institute of Language and Communication, University of Southern Denmark, Odense, Denmark. E-mail: theo.vanleeuwen@uts.edu.au
1 Orientation to research on PowerPoint

In July 2011, a newly launched political party in Switzerland captured the imagination of the international media. The Anti-PowerPoint Party (APPP) claimed “350 billion Euro” (Bindel 2011) could be saved globally each year if PowerPoint were to be abolished and presented itself as an “advocate for the approximately 250 million world citizens […] who, every month, are obliged to be present at boring presentations in companies, at universities or at schools” (Anti-PowerPoint Party 2012). Whether these figures are reliable or mere tongue-in-cheek, the impact of PowerPoint on public speaking, since the software’s first launch in 1987, cannot be overlooked.

PowerPoint presentations are complex and multimodal. To understand them, we need to consider the software’s design (the semiotic resources PowerPoint makes available and how these are presented through its interface and help menu), the ways in which they are co-deployed in the design of slideshows, the way the design of slideshows interacts with the speech and gestures of the actual presentations, and the various social norms that impinge on the software’s design and its use. By proposing a social semiotic model for analyzing PowerPoint as a complex multimodal practice, this article is envisaged as a step toward developing a social semiotic theory of the complex interaction between the design and the use of semiotic technologies (i.e., technologies for making meaning).

Over the past decade, the importance of PowerPoint for evaluating success in a growing number of work and educational settings has been widely recognized and discussed in the public media as well as in the professional and academic literature. Debate on PowerPoint tends to be strongly polarized, with some authors praising PowerPoint for increasing presenters’ confidence and eloquence (e.g., Gold 2002) and others condemning it for limiting users’ ability to achieve important social purposes (e.g., to present an argument or a complex narrative) and to explain technical information (e.g., Huang 2009; Tufte 2003; Turkle 2003). There is also a large body of prescriptive literature on PowerPoint providing decontextualized guidelines on elements of slide design (e.g., bullet points, headings, graphics) or physical aspects of presentations (e.g., visibility of speaker and slides) (e.g., Atkinson 2007 [2005]; Bozarth 2008; Doumont 2005). This literature typically relies on anecdotal evidence and personal observation rather than on systematic exploration of either slideshows or their use in presentations (Farkas 2006).

Most empirical research on PowerPoint-supported presentations comes from the field of higher education and focuses on learners’ and teachers’ perceptions of the effectiveness of PowerPoint slideshows versus transparencies in lectures and on the ability of presentation features such as graphics to support the reten-
tion and understanding of content (e.g., Apperson et al. 2006, 2008; Bartsch and Cobern 2003; James et al. 2006; Rockwell and Singleton 2007). Like the debate and prescriptive literature on PowerPoint, these studies, too, ignore the multimodal integration of semiotic resources in slideshows and presentations, the interaction between software, slideshow, and presentation, and the type and influence of various social norms over this interaction.

This paper addresses this gap by proposing a holistic social semiotic approach to studying PowerPoint as a complex multimodal practice that comprises three interrelated dimensions: the software’s design, the multimodal composition of slideshows, and the slideshow presentation. These dimensions are connected through two “semiotic artefacts” (Kress and Van Leeuwen 2001): the software and the slideshow. We shall argue that a model for exploring all three dimensions and their relationships is crucial for understanding PowerPoint as a semiotic technology.

Our model derives from an ongoing project that explores the interaction between the design of PowerPoint and its use in corporate and higher education settings by examining four data sets:

- video-recordings of 27 PowerPoint presentations (including 17 university lectures in engineering and cultural studies, and 10 business presentations);
- the slideshow files used in these presentations;
- follow-up interviews with each of the presenters.

In this paper we employ a small number of selected empirical findings of this project in order to outline our position on studying the relationship between semiotic technology and semiotic practice.

2 PowerPoint as a multidimensional multimodal semiotic practice: an overview

2.1 The challenges PowerPoint presents for discourse analysis and social semiotics

The term PowerPoint at various times refers to (i) the software designed by Microsoft, (ii) a slideshow created with it, or (iii) a presentation featuring such a slideshow. These three are interrelated aspects of the same phenomenon. Yet, in research on PowerPoint they tend to be considered in isolation from each
other. Until recently (Djonov and Van Leeuwen 2011, 2012, 2013), PowerPoint has received only limited attention in the fields of discourse analysis and social semiotic research. Most research has focused on the design of slideshows (e.g., Campagna 2009; Finn 2010; Rowley-Jolivet 2004). Rowley-Jolivet (2004), for instance, has argued that conference presentations in science vary in their visual communication style depending on their field within the discipline, so that each field displays different use of semiotic modes (e.g., language, images, and mathematical symbolism) and spatial or temporal composition principles. The assumption is that PowerPoint slideshows can be analyzed as “texts.” This often leads to a logocentric approach, in which the verbal text is the primary reference point (e.g., Rowley-Jolivet 2004; Campagna 2009). Other studies treat slideshows as products of visual and multimodal composition (e.g., Finn 2010) but evaluate their effectiveness independently of the ways in which slideshows are integrated in PowerPoint presentations. This has value for understanding the increasingly common practice of disseminating and reading slideshows as standalone documents (Farkas 2006; Yates and Orlikowski 2007), but obscures the complexity of PowerPoint as a multidimensional semiotic practice. Analysis of slideshows alone cannot show how the design of the software privileges certain ways of using the semiotic resources it makes available (e.g., layout, texture, color) and of composing and presenting slideshows (Djonov and Van Leeuwen 2011, 2012). And many slideshows are not designed to make sense outside the context of the actual presentations for which they are designed. An example from our data is a slideshow designed for a one-hour university lecture in cultural studies titled “Europe as a Region”. It contains 26 slides, including 1 title slide and 25 content slides. The content slides include 9 that feature a single noun or nominal group on a black background, 8 with a single image without any accompanying words, and 1 showing 5 maps with barely visible legends. When analyzed alone, the slide set appears to lack any internal cohesion or even coherence. While all the slides are to various extents semantically related to the key lexical message “Europe,” there are few semantic relations that can be established with certainty between individual slides.

To illustrate, consider the two consecutive slides from this slideshow shown in Figure 1. The first shows a map of the Greek-Phoenician trade routes, and the second a single image of a woman standing amidst building relics. As neither the visual participant nor the visual location (Kress and Van Leeuwen 2006 [1996]) shown in this image is identified, a “semantic bridge” (Martin 1992) between the two slides cannot be established without interpreting them as part of the presentation.

Their position in the overall presentation and the speech that accompanies them, however, reveal the semantic link between the two slides. The first slide
Fig. 1: Two consecutives slides from the lecture “Europe as a Region”: (a) Greek-Phoenician colonies and trade; (b) a woman standing amidst building relics in South Beirut
follows the lecturer’s retelling of the Ovidian myth of Europa, and an excerpt of
the speech that accompanies the slide reminds students that: “in this story [. . .] Europe was [. . .] actually from Phoenicia, what is now today south Lebanon”. As
the second slide appears, it is accompanied by the words “From this place, this
is um south Beirut in 2006 during the last war. Um and geographically speaking
this is exactly where Europa was from so that the beginnings of Europe are in
what today is called The Middle East.” The speech thus weaves the map and the
image into the same lexical string and reveals their relationship as one of loca-
tion, as illustrated in Figure 2. Without the lecturer’s words this relation cannot
be established, and the sequencing of the two slides appears incoherent.

This slideshow is an extreme example in that the relationships between the
slides in it are established solely through the presenter’s speech, and contrasts
strongly with slideshows designed to be read as standalone documents. Yet, our
data includes many instances where the meaning of individual slides and rela-
tionships between adjacent slides cannot be interpreted in isolation from their
interaction with the speech and gestures of the presenters. Such observations
have led us to develop an integrated multimodal social semiotic approach towards
studying PowerPoint, in which we focus on understanding PowerPoint as a multi-
dimensional multimodal semiotic practice, and away from the notion of text.

2.2 The three dimensions of PowerPoint

In our approach, overviewed in Figure 3, the semiotic practice of PowerPoint
is modeled and studied as encompassing three interdependent dimensions:
(i) the software’s design, (ii) the multimodal composition of slideshows, and
(iii) slideshow-supported presentations.
From a social semiotic perspective, the first dimension involves software designers selecting from potentially unlimited meaning-making resources that are recognized in the culture (such as language, images, colors, layouts and texture, etc.), making the selected resources available to users through the software’s interface, and using both the interface and the help menu to present, and potentially impose on users, their ideas on how the resources PowerPoint provides should be (co)deployed in the composition and presentation of slideshows. The process of software design is itself regulated by various normative discourses, and reflects both professional design conventions as well as the designers’ understandings of the semiotic practices which the software is designed to support (Djonov and Van Leeuwen 2012, 2013). The design practice produces the software, which is employed in the second dimension of PowerPoint, the multimodal composition of slideshows. Following Kress and Van Leeuwen (2001) and Van Leeuwen (2005), we view the software as a “a semiotic artefact”: it is a semiotic resource that has material form and incorporates selections from different semiotic modes (e.g., layout, texture, color, sound) and media (e.g., visual, aural, print, electronic), and is deployed for making meaning alongside other types of semiotic resources (i.e., modes, media and abstract semiotic principles such as genre and rhythm).

The composition of slideshows represents a type of “new writing,” or writing that relies on visual resources and multimodal interaction, rather than on verbal resources alone, to establish cohesion and coherence (Van Leeuwen 2008). This dimension of PowerPoint is where the interaction between the software’s design and use occurs, and – as we illustrate in Section 4 – can reveal discrepancies between the communicative purposes of the software’s designers and its users.

**Fig. 3:** An integrated multimodal social semiotic approach to PowerPoint
When composing a slideshow, users select from the resources available within the software (e.g., font types, texture, layout, animation, and sound effects) at the same time as using other semiotic modes and artifacts that the software allows them to incorporate into slideshows (e.g., written language, images, video, and sound clips). The dimension of software design then both constrains and enables the composition of slideshows (Doumont 2005; Stark and Paravel 2008; Yates and Orlikowski 2007). The multimodal composition of slideshows, like the dimension of software design, is also subject to various norms. However, what influences "new writing" practices is far from straightforward. Existing literature and our own explorations suggest that there is a combination of factors, including (sub)disciplinary differences (Rowley-Jolivet 2004), familiarity with the software, with semiotic resources in general and semiotic practices not supported by the software, as well as with conventions for the composition and presentations of writing, visual and multimodal documents in different professional fields, including graphic design, teaching in higher education, sales and business presentations, and so on (e.g., Craig and Amernic 2006; Djonov and Van Leeuwen 2013; Yates and Orlikowski 2007). The dimension of multimodal slideshow composition, like that of design, also yields a semiotic artifact – the slideshow. This artifact is employed and reinterpreted in the dimension of PowerPoint-supported presentations.

In the third dimension of PowerPoint, the presenter, who may or may not be the author of the slides, incorporates the slideshow as a semiotic artifact into a presentation, combining the interaction of various semiotic resources with the writing and any nonverbal or multimodal objects embedded in the slideshow with speech and gestures in order to achieve a given communicative purpose (e.g., to promote a new product or explain a complex phenomenon). Successfully combining different semiotic resources and objects during a presentation requires intensive temporal coordination; that is, the presenter needs to coordinate the meaningful interactions between various (multi)semiotic units (e.g., image, speech, and gesture) while presenting. Adopting Goffman’s (1981) framework for defining participant roles in the production of talk, it is important to note that while presenters may or may not fulfill the role of “principal,” of the person or institution “whose beliefs are told” (1981: 144), their participation in PowerPoint as a semiotic practice always conflates Goffman’s other two roles: the presenter is never merely an “animator,” never just someone who physically executes a script, but is also an “author,” someone who makes decisions about how to combine the resources employed in the composition of slides with semiotic resources available for interpreting and explaining the ideas represented on the slides and for relating to the audience. In our model, this conflation of roles is acknowledged by conceptualizing the slideshow as a semiotic artifact, or a resource for making meaning, rather than a text.
As in the dimension of slideshow composition, authorial power in presentations is not unlimited; a presentation is simultaneously constrained and facilitated by the software’s design, the slideshow’s composition, contingencies of the actual presentation event, and various communicative conventions that presenters are familiar with and observe. The strength of conventions for the design and presentation of slideshows, moreover, varies across different settings. In our data, corporate slideshows and presentations are strongly regulated, while cultural studies lectures display considerable freedom in slideshow design and presentation.

Finally, although the presentation does not produce a semiotic artifact directly, it has the potential to feed back into the revision of existing slideshows (i.e., sets of slides) and the composition of new slideshows as well as into the software’s design (as suggested by the dashed arrows that connect dimensions in Figure 3). The key challenge presented by this dimension is how to model the interaction between different meaning-making units that characterizes different points in the unfolding of a PowerPoint-supported presentation, and is discussed further in Section 5.

Having presented the motivation behind and an overview of our approach to studying PowerPoint as a multidimensional multimodal semiotic practice, we shall now discuss how we explore each of PowerPoint’s dimensions, the challenges they present, and their relationships.

3 PowerPoint as software: semiotic resources and design norms

The first dimension of PowerPoint we investigate is the software’s design. In applied linguistics and discourse analysis, software is studied as a tool for facilitating and researching language learning (e.g., Chapelle 2003) and for analyzing verbal and multimodal texts, especially large corpora (e.g., Baldry 2004; McEnery et al. 2006), but is not considered an object of study on its own right. One possible reason for this is that a software does not represent what discourse analysts

---

1 The process of the presentation can be turned into an artifact through video-recording, and can then be employed in practices such as analyzing or presenting the video. (See Jewitt 2011 on video-based social research.)

2 By contrast, studies in computer semiotics (Andersen 1997), human–computer interaction (De Souza 2005), and the semiotics of programming (Tanaka-Ishii 2010) propose abstract classifications of different signs in computing and recognize that software incorporates verbal, visual, tactile, and aural resources but offer no model for analyzing their multimodal interaction.
traditionally conceive of as “text” (that is, as their traditional object of study), whether we adopt a broader social semiotic definition of text “as a message or message traces with socially ascribed unity” (Hodge and Kress 1988: 6) or a more technical one of text as a unit displaying the quality of “texture,” or having both internal cohesion (which in language is achieved through choices from the systems of REFERENCE, SUBSTITUTION, ELLIPSIS, CONJUNCTION, and LEXICAL COHESION) and coherence in relation to its context of situation (Halliday and Hasan 1985).3

Even though software does not fit into the category of “text,” it is undeniably “meaningful.” In the most direct sense, software is meaningful as its interface incorporates various verbal, visual, and aural resources (Andersen 1997). To produce a paper in Word, a user needs to read and understand the meaning of options presented through verbal labels such as “blank document” and visual-verbal icons such as the print button. Software is also meaningful in the sense that it is not “neutral” but has the potential to reflect social power relations and the interests and values of its designers and users, as has been convincingly argued in Arola (2010), Fuller (2008), Selfe and Selfe (1994), and Sorapure (2006) from the perspectives of cultural, media, and composition studies. Finally, software such as PowerPoint provides users with various resources for making meaning, which is why we see it as a “semiotic technology.” In our approach to PowerPoint communication, we analyze the software with the purpose of mapping out the semiotic resources it makes available to users and how these are presented through the interface created by the software’s designers. Considering the interface (alongside the help menu) offers insights into the designers’ views on how these resources should be (co)deployed and combined with written text, images, and media files in the composition of slideshows and, to a lesser extent, on how slideshows should then be presented.

The first focus of our analysis is to map out what resources are available within the software and their meaning-making potential. We are especially interested in resources that can help shed light on the composition of PowerPoint slideshows as a type of “new writing” practice (Van Leeuwen 2008), including visual texture (Djonov and Van Leeuwen 2011), layout (Djonov and Van Leeuwen 2013), and animation/kinetic typography (Van Leeuwen 2011). Yet, these resources are not unique to PowerPoint as their semiotic potential has been and continues to be explored and exploited in a range of other semiotic practices and media (cf. Djonov and Van Leeuwen 2014). To understand how PowerPoint constrains or enables

3 On a more intuitive level, perhaps discourse analysts and linguists do not recognize software as a legitimate object of study as it is not easily captured in a medium they are familiar with (e.g., print or video).
the use of these resources, therefore, we first model their meaning-making potential by researching the ways they are employed, discussed, and theorized across a range of professional and academic fields such as graphic and fashion design, sculpture, painting, art history and theory, and research in semiotics, cognitive psychology, cultural and media studies. This strategy enables us to model the meaning-making potential of both semiotic modes such as layout and material resources such as texture, and to explore the relationship between the software’s design and other semiotic practices as well as changes in the software’s design over time.

Our exploration of texture (Djonov and Van Leeuwen 2011), for example, involved modeling the meaning potential of both surface tactile and visual texture and studying the relationship between the two. Like the material resources of color (Kress and Van Leeuwen 2002), voice quality (Van Leeuwen 2009), and typography (Van Leeuwen 2006), we modeled the meaning-making potential of texture as the result of (i) the provenance of specific textures (i.e., their association with particular objects such as cashmere or denim, and the significance of these objects in particular sociohistorical and cultural contexts) as well as (ii) the interaction of a texture’s physical properties and their association with bodily experience. These properties were modeled paradigmatically and presented as a parametric system. The system can be used to describe a surface tactile texture in terms of several parameters (liquidity, viscosity, temperature, relief, density, and rigidity), that is, several simultaneously available and gradable, rather than alternative and binary, choices (e.g., the parameter of rigidity is a scale from soft to hard), which not only differentiate one texture from another but add layers of meaning to it. We then examined the applicability of these parameters to visual texture through examples from painting and graphic design before considering how visual texture is presented in PowerPoint.

The analysis of texture in PowerPoint explored changes in the presentation of this visual semiotic resource across all versions of the software for Windows from 1992 to 2007. This diachronic perspective illuminates the role that semiotic technologies like PowerPoint, not unlike semiotic theories, play in reflecting as well as shaping the meaning-making potential of different semiotic resources. The comparison of texture options across different versions of PowerPoint, for example, revealed a shift “from the logic of the ‘lexicon’ to the logic of the ‘grammar’ whose rules allow a multitude of complex signs to be constructed with a toolbox consisting of a restricted set of lexical choices and grammatical operations” (Djonov and Van Leeuwen 2011: 560). To elaborate, in earlier versions users could merely select from a number of available textures (e.g., “green marble” or “walnut”) which derived their meaning from provenance, while in PowerPoint 2007 they could also modify the appearance of available textures by changing parameters.
such as a texture’s transparency and regularity. This change in PowerPoint 2007 is thus a step towards expanding the meaning-making potential of visual texture within the software.

A second focus of our analysis is to reveal how different semiotic resources are presented and accessed through the software’s interface. This enables us to understand the potential of PowerPoint’s design to impose certain norms on the composition of slideshows. A comparison of the layout templates available in two versions of PowerPoint, 2003 and 2007, for example, reveals that the templates “guide” the use of layout in slideshows by pre-selecting the relative salience and framing of different types of placeholders and design elements such as typography, color, and media objects. The labels of layout templates in PowerPoint 2007 also reflect a shift from a focus on the types of objects (e.g., text, media clip, table) that can occupy different positions on a slide to a stronger focus on the rhetorical relations that templates could be used to signal between the elements that

Fig. 4: Default layout options in PowerPoint 2007 (used with permission from Microsoft)
occupy them (e.g., “Comparison”) or between slides (e.g., “Section Header”) (see Figure 4). Additionally, the interface makes certain choices easier than others: for example, when a new PowerPoint file is created, the title slide template (which is also presented first in the slide layout menu, as shown in Figure 4) appears automatically, which strongly suggests that a slideshow must start with a title slide. The software’s interface thus sets up normative expectations about the ways in which layout should be used in slideshows, not only through the resources it makes available to users and the ways in which these resources (e.g., layout templates) are labeled or represented visually but also through its spatio-temporal, or syntagmatic, organization. (See further Djonov and Van Leeuwen 2012, 2013.)

In essence, exploring the software’s design from a social semiotic perspective offers insights into its designers’ perspectives on meaning making in general and PowerPoint communication in particular. Such insights, however, need to be considered in relation to the ways in which users employ the resources PowerPoint makes available and the extent to which they are aware of and follow the norms built into the software’s design, as we discuss next.

4  PowerPoint slideshows – where software design meets software use

The second dimension of PowerPoint we focus on is the multimodal composition of slideshows, including both the slideshows themselves and the discourses that surround them. We explored the latter through interviews with the authors of the slideshows we collected. These interviews allowed us to gain insight into the authoring process itself and the types and power of normative discourses that impinge on this process – for example, about authors’ degree of familiarity with the software and with slideshow design and presentation conventions deriving from different institutional and professional contexts. The analysis of the interviews suggested that individual slides and/or slideshows are often (re)designed or repurposed for different presentations, shared by different presenters, and distributed through different channels. In higher education settings, lecture slides are often available for students to download, and may or may not be accompanied by an audio/video recording of the lecture, while in business settings, many PowerPoint slideshows are distributed as standalone documents, exemplifying what Yates and Orlikowski (2007: 89) term a “corollary genre,” a genre “decoupl[ing] texts from the particular recurrent situation around which they emerged.” The interviews thus reinforced the view that the “PowerPoint presentation [...] both reflects and shapes organizational practices while also enabling and constrain-
ing range of social actions and outcomes” (2007: 87), and offered insights that
enriched our interpretation of slideshows as part of a complex semiotic practice,
and ultimately helped us avoid viewing slideshows as standalone texts divorced
from their contexts of production and reception. The interviews also showed that
slide show authors reinterpret the software’s meaning-making potential on the
basis of different communication norms with which they may be familiar and
which may or may not be built into the software.

This approach can also expose gaps between the software’s design and its
use. Consider, for example, the two slides in Figure 5, which belong to a slide-
show designed for a new staff induction session at an Australian university. Both
slides use the university’s design theme for corporate presentations and the same
layout template (“Two Content” from PowerPoint 2007; see Figure 4), where the
slide title is followed by two horizontally distributed images. In the first slide,
titled “Senior Executive”, both images have captions. In the second slide, only
the image on the left is accompanied by verbal text, a bullet list of four locations.
The layout template suggests that there are at least three sets of semantic rela-
tions: (i) between each image and its caption, (ii) between the title and body of
the slide, and (iii) between the two images on each slide. In the first slide, these
relations are straightforward: (i) the captions name and identify the two visual
participants, (ii) each participant is part of the university’s Senior Executive
team (meronymy), and (iii) both are senior executive members (i.e., related by
co-meronymy). As this slide successfully translates the layout structure into se-
monic relations to establish cohesion (i.e., links between the slide elements) and
coherence (i.e., links to the purpose of this presentation, which is to introduce
and promote the university to new staff members), it lends itself to being read as
a “text.”

The layout structure of the second slide, by contrast, fails to achieve this. For
example, positioning the bullet list as a caption suggests a close relation between
the list and the image above it. The list mentions six buildings, whereas the image
shows only one, and this building is not identified. In fact, the slide has been
designed to introduce the university’s two campuses: a city campus (comprising
all the buildings in the bullet list and the building shown in the image) and the
campus represented by the other image. A new staff member reading this slide on
its own, without the presentation it was part of and without familiarity with the
university, would not be able to recover the actual semantic relations built into it
based on layout alone.

For us, this mismatch between the meaning-making potential of layout tem-
plates available in PowerPoint and the semantic relations they are employed to
signal in actual slideshows and slideshow presentations exposes a gap between
the design and the use of PowerPoint as a ubiquitous semiotic technology with
Senior Executive

Professor Ross Milbourne
Vice-Chancellor and President

Professor Peter Booth
Senior Deputy Vice-Chancellor and Senior Vice-President

Campuses

- Broadway (Buildings 1, 2)
- Haymarket (Building 5)
- Jones St (Building 10)
- Harris St (Buildings 3, 4, 6)

Fig. 5: Two slides using the same layout template from PowerPoint 2007
profound implications for understanding the influence of semiotic technologies on new writing practices. It suggests that slideshow authors need to be aware of the meaning-making potential of layout as a key semiotic resource in new writing practices and should consider critically both how it is presented within the software’s interface and help menu and how it functions in various semiotic practices within the broader culture. They also need to keep in mind that digital technologies and organizational practices may lead to semiotic artifacts being distributed and re-appropriated in contexts they were not originally designed for (Djonov and Van Leeuwen 2013; Yates and Orlikowski 2007). Finally, this gap (like the example discussed in Section 2) highlights the importance of analyzing slide-shows in relation to their presentation, rather than as standalone texts, especially in cases where they are not designed to be cohesive and coherent on their own.

5 PowerPoint presentations as multimodal events

The final dimension of PowerPoint we investigate is the presentation. We can treat PowerPoint-supported presentations as texts because they are typically events with a beginning and an end, characterized by relative cohesion and coherence. But an analysis of this kind must treat the slideshow as a semiotic artifact whose meaning-making potential is both actualized and reconfigured during the presentation. It must also allow us to assign nonverbal resources such as gesture the status they deserve depending on their role within specific segments in the presentation, rather than regarding them a priori as paralinguistic resources, subservient to language (cf. Matthiessen 2007). Finally, it must have the capacity to capture the logogenetic patterning of semiotic resources and semantic relations during the presentation; that is, it should reflect how labor is (re)distributed among different semiotic resources and how different semantic relations are established between them and then reconfigured in the spatio-temporal unfolding of the presentation (Zhao 2010). By taking a step toward developing such a model, our research contributes to efforts directed at understanding the complex interactions between various semiotic resources in time-based data such as embodied semiotic events, music, and film (Lemke 2005; O’Halloran 2004).

Let us illustrate the complexity of PowerPoint presentations by considering another example from the new staff induction discussed in Section 4. We refer to this example as “the map sequence” as it involves the presenter using a map to explain her university’s renovation and construction work. A partial overview of the presentation’s physical setting (the presenter’s placement in relation to the audience, the screen, and the lectern) and the slide (titled “City Campus Master Plan”) used by the presenter in this sequence are shown in Figure 6.
Fig. 6: Map sequence: video snapshot and slide
This presentation is a standard introduction to the university given to new staff twice a year by the vice chancellor. On this occasion, however, a different member of the university’s senior executive team gave the presentation. As the slideshow was not designed by or for her, the presenter, although familiar with its format and content, had to reinterpret it during the presentation. She seemed particularly excited about explaining the city campus master plan because of her belief that “people[,] whether they love it or hate it[,] want to know what’s going to impact them and their environment, not just when it’s finished but also along the way”. So in this particular sequence, the presenter made considerable effort to draw the audience into the presentation through intensive coordination of her speech and gesture with the image on the slide as well as the physical location of the presentation. This is evident, for example, in the extensive use of exophoric reference to the physical setting in her speech (e.g., “here”, “there”) and in her pointing to various buildings on the map. Such intensive multimodal coordination was in part necessary because the verbal labels identifying the building sites on the master plan could not be read from the slide, which meant that semantic relations between these labels and the image were “in retreat” in the context of the presentation.

So, how should we model such complex multimodal events? Research inspired by linguistic theories tends to take either verbal grammar or phonology as a starting point, first parsing the speech into units such as clauses or tone groups and then mapping onto these any nonverbal, for example gestural, units (Zappavigna et al. 2010). This approach is useful for some presentations in our data where presenters read from lecture or presentation scripts and barely use gestures (hand, head, or body movements) to relate their speech to the projected slides or to the audience.

This strategy, however, is not suitable for segments like the map sequence where image, gesture, and speech play equally important roles. This sequence illustrates the incommensurability of image and speech as semiotic resources (Lemke 1998) – the map provides topological information about the spatial distribution of buildings across the campus, whereas the speech offers typological information about these buildings by labeling them or describing selected characteristics (e.g., color, function, cost) of individual buildings. In this sequence the image and the speech form one semantic unit, and neither can be understood without the other. This interdependency would be obscured if language were used as a starting point for modeling intersemiotic interaction, as would be the crucial role gesture plays in coordinating speech and image in this segment.

To overcome these problems, we propose that the analysis of complex multimodal events such as PowerPoint presentations should focus on understanding how different semiotic resources are coordinated to fulfil different semantic
functions at various points during each event, and not exclusively on identifying different types of intersemiotic relations. In other words, studying the dynamic and complex patterning of intersemiotic relations in multimodal events such as PowerPoint presentations requires an adaptive approach that can reflect subtle shifts in the distribution of labor among different semiotic resources and in their semantic relations. Such is the approach illustrated by the combined transcription and analysis of a short segment from the map sequence presented in Table 1.

The transcript has two main vertical columns: the first indicates the semiotic resource responsible for the coordination of semiotic resources at a given point in the presentation, while the second shows what resources are being coordinated. In this case, these resources include a visual unit (one of the buildings on the map) and speech, which has been segmented into rhythmic units (demarcated by square brackets in the transcript) following Van Leeuwen (1999), while the coordinating resource is gesture. Gesture has been segmented into units and phases following Kendon (2004). In this segment, there is one gesture unit comprising four phases. The unit is a right forearm gesture and starts with a preparation (i.e., right hand raised with index finger in a pointing position), during which the presenter searches the map for the building she is about to mention. The forearm arrives at the stroke when her finger finds the building, and then the presenter holds this pointing gesture while providing information about the building. The post-stroke hold contains two embedded phases: a head gesture phase, which occurs once the presenter has found the building on the map and turns her body toward the audience, and a left forearm gesture phase, which links the building on the image with the location of the presentation. The body recovers to starting position once the explanation of the building is completed. The two forearm gestures are primarily deictic, or pointing, gestures, while the head gesture serves a different discursive function – to signal interpersonal engagement with the audience.

In addition to fulfilling deictic and interpersonal functions, the gestures synchronize other semiotic resources in this segment. The right forearm gesture, for instance, is held until the presenter completes her explanation about the building. The head and left forearm gestures that take place during this post-stroke hold bring the resource of three-dimensional space into the explanation as well. The right forearm gesture phase also marks the boundaries of the “inter-semiotic coupling” (Zhao 2010, 2011) between the image of the building and the words “this big construction site in the middle of the building over here that’s the multipurpose sports hall.”

Taking an adaptive approach to analyzing intersemiotic relations recognizing that presentations and presentation segments may employ different coordinating semiotic resources and that intersemiotic semantic relations
Table 1: Temporal coordination in the map sequence

<table>
<thead>
<tr>
<th>Phase</th>
<th>Resource for coordination</th>
<th>Semantic integration</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><strong>Gesture units</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Right forearm</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>[Preparation]</td>
<td>[what's UNDERWAY]</td>
</tr>
<tr>
<td></td>
<td>Head</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Stroke</td>
<td>[THIS big construction SITE]</td>
</tr>
<tr>
<td></td>
<td>Left forearm</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Post-stroke hold</td>
<td>[in the middle of /THE ^]</td>
</tr>
<tr>
<td></td>
<td>Stroke</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Post-stroke hold</td>
<td></td>
</tr>
</tbody>
</table>


<table>
<thead>
<tr>
<th>Phase</th>
<th>Resource for coordination</th>
<th>Gesture units</th>
<th>Visual units</th>
<th>Speech units</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stroke</td>
<td>Left forearm</td>
<td><strong>Building over here</strong></td>
<td></td>
<td><strong>That's the multipurpose/sports mall</strong></td>
</tr>
<tr>
<td>Recovery</td>
<td>Right forearm</td>
<td></td>
<td></td>
<td><strong>Show you pictures in a minute</strong></td>
</tr>
</tbody>
</table>

---

**Table 1 (cont.)**

| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 | 26 | 27 | 28 | 29 | 30 | 31 | 32 | 33 | 34 | 35 | 36 | 37 | 38 | 39 | 40 |
established at one point of a multimodal event can be reconfigured at another. In the segment from the lecture “Europe as a Region” discussed in Section 2, for example, speech is the resource coordinating the map of the Greek-Phoenician trade routes with the image of the woman as visual units. By contrast, another presentation employs kinetic typography options to coordinate and establish relations among other semiotic resources. In “Europe as a Region”, speech plays a significant role first in establishing the semantic relation between the map and the image of the woman as one of location, and later in building on and reconfiguring this relation to construct a more abstract, metaphorical idea that brings together different images shown previously in the presentation and the notions of “continents,” “women,” “rape,” and “war” they evoke:

Continents and countries are often represented as women; they are defenseless and passive. If you think about metaphors of war, there are a set of metaphors that, for instance, equate war to rape of women. In the war on terror, the metaphor of rape was used several times as in “we raped that country”.

As the examples discussed in this section suggest, an adaptive approach to analyzing PowerPoint presentations as multimodal events is essential for understanding and evaluating their ability to enhance multimodal learning in education settings and fulfill communicative purposes such as promoting a company in corporate ones.

6 Conclusion

This paper presented a social semiotic model for analyzing PowerPoint communication as a complex multimodal practice comprising three interrelated dimensions – the software’s design, the multimodal composition of slideshows, and their presentation – and two semiotic artifacts, the software and the slideshow. We argued that studies of PowerPoint should also consider the ways and extent to which each dimension is influenced by various normative discourses, and discussed the challenges each dimension presents for discourse analysis and social semiotic research, focusing especially on the need to step away from the notion of text and to develop a holistic, non-logocentric, and adaptive multimodal approach to researching PowerPoint.

Viewed from a more general perspective, the approach presented here constitutes a step toward building a social semiotic theory of semiotic technology and semiotic practice, especially practices involving software for making mean-
ing such as PowerPoint. Such a theory must have the capacity to account for the following dimensions and the connections between them:

− the technology/software itself, including the processes of its (re)design, the semiotic resources it makes available to users, and the ways these resources are presented through its interface;
− the multimodal composition of semiotic artifacts produced with such technologies and the factors that influence it;
− the incorporation of the technology and semiotic artifacts produced with it alongside other semiotic resources in specific communication events.

Explorations of semiotic practices and semiotic technology must also consider changes in each dimension as well as the ways semiotic resources and normative discourses associated with each dimension relate to semiotic resources, normative discourses, and semiotic practices in the culture in general. As Van Leeuwen and Djonov (2013) argue, investigating the relations between a semiotic practice associated with a given semiotic technology and other, past and present, semiotic practices is key to establishing the extent to which a semiotic technology privileges some semiotic resources over others, expands or constrains the meaning-making potential of different semiotic resources, draws on and influences certain semiotic practices (e.g., Photoshop incorporates principles from visual arts, graphic design, and photography), and ultimately contributes to changes in the semiotic landscape.

Acknowledgments: The paper is part of a larger project, “Towards a social theory of semiotic technology: Exploring PowerPoint’s design and its use in higher education and corporate settings,” which is supported through an Australian Research Council Discovery Grant.

References


Arola, Kristin. 2010. The design of Web 2.0: The rise of the template, the fall of design. *Computers and Composition* 27. 4–14.


Bionotes

Sumin Zhao is Chancellor’s Postdoctoral Research Fellow at the University of Technology, Sydney. Her current project investigates preschool children’s interactions with mobile technologies in home-settings by combining multimodal discourse analysis with other research methodologies. She publishes in the research areas of digital literacies, systemic functional linguistic theory, and multimodal discourse analysis. Her latest book (co-edited with Emilia Djonov), Critical Multimodal Studies of Popular Discourse, has just been published by Routledge.

Emilia Djonov is Lecturer in multiliteracies at the Institute of Early Childhood, Macquarie University, Australia, and an honorary postdoctoral research fellow at the University of Technology, Sydney, where she is researching the interaction between PowerPoint’s design and use in higher education and corporate settings. Her research interests and publications are in the areas of (critical) multimodal and hypermedia discourse analysis, visual communication, social semiotics, systemic functional theory, and multiliteracies.
Theo van Leeuwen is Emeritus Professor at the University of Technology, Sydney, and Professor of Language and Communication at the University of Southern Denmark. He has published widely on critical discourse analysis, multimodality, and visual semiotics. His books include *Reading Images* (with Gunther Kress) and *Discourse and Practice*. He is a founding editor of the journal *Visual Communication*. 