

***Spoken and Written Discourse in Online Interactions: A Multimodal Approach***

**Maria Grazia Sindoni (2013)**

**New York: Routledge. Pp. xv, 240**

**ISBN: 978-0-415-52316-5**

*Reviewed by Martha C. Pennington*

*Spoken and Written Discourse in Online Interactions: A Multimodal Approach* is a recent title in the growing Routledge Studies in Multimodality series, which is edited by Kay L. O'Halloran, National University of Singapore. Author Maria Grazia Sindoni (University of Messina, Italy) has contributed a significant and original data-based work of scholarship to the series and to the larger body of research and theory on multimodality and online discourse. Sindoni's treatment of online discourse is based on several online corpora which she assembled and which she subjected to lexical analysis (primarily analysis of key words and lexical clusters) as well as other forms of analysis that elucidate the ways in which different modes and technological resources are employed in online contexts. The main content of the book is contained in four meaty chapters, which include numerous helpful summative lists, tables, and figures. The central chapters are framed by brief introductory and concluding chapters, and supported by 17 pages of references, in addition to appendices which include detailed multimodal transcripts of both online speech and videochat and an overview of the ICE: Great Britain Corpus Structure that the author analyzes for baseline comparisons to her data.

In the Introduction, Sindoni makes clear the overall aim and thrust of the book, underpinned by multimodal and intersemiotic theory, as '...to reconsider underlying semiotic frameworks of analysis of spoken and

---

**Affiliation**

University of London  
email: mp64@soas.ac.uk

written discourse in the digital age, in the light of the new paradigms of online communication' (p. 2). As she goes on to observe:

Common patterns of interactions are altered in the web galaxy and new patterns of communication emerge, challenging previous notions of what communication actually *is* in the contemporary age. Online configurations of interaction, such as videochats, blogging and social networking practices, demand a profound rethink of the categories of linguistic analysis, given the blurring of traditional distinctions between oral and written discourse in digital texts. Typical modes of online interaction encompass speech, writing, gesture, movement, gaze and social distance in spontaneous web-based interactions. This is nothing new, but this volume claims that all these modes are integrated in unprecedented ways, enacting new interactional patterns and new systems of interpretation among web users. (p. 2; emphasis in the original)

In the chapters that follow, Sindoni shows how the various modes of online interaction co-occur and are sequenced in mode-switching (i.e. between speech and writing) as well as in resource-switching, 'which is broader than mode-switching and includes all the semiotic alternations that can be found in multimodal web-based environments, for example [*sic*] videos, photos, pictures, tagging and hyperlinking, etc.' (p. 3).

Chapter 1 presents background on the nature of spoken and written textuality as these have been conceived both outside of and within digital environments. Much of the chapter is devoted to describing and updating the sorts of stereotypical differences between speaking and writing that other scholars have popularized. These differences are then re-examined for digital contexts, resulting in a table (Table 1.3, p. 44) which summarizes the specific features of spoken, written, and digital texts in terms of: *materiality, context, localization, authorship, degree of involvement, production, reception, degree of interaction, degree of control, storage, provenance (source), use of multimodal resources, and typical kind of multimodal resources.*

Chapter 2, the first of three chapters based on analysis of data, examines multi-party web-based videochats at a site called Camfrog. Camfrog allows users to instant-message each other on a private, one-to-one basis and also to connect to various chatrooms and to upload digital objects such as video and audiofiles for viewing or sharing with others. In videochat, mode-switching is possible, as participants may talk as well as write and send comments. Beyond having access to these different affordances of the videochat mode, Sindoni observes that communicating over the internet by way of video rather than face to face changes the context of talk in that movement is restricted, distance is fixed, and the communicators can see themselves on the screen as they talk. These characteristics

of the videochat context change the psychological state of the participants and their behavior, as the proxemics and kinesics of communication are inevitably altered. In the view of Sindoni, 'videochat systems erase all kinds of differences and favour a flat representation of social and cultural identities' (p. 57), as they also '[create] a *frozen yet living* image of users' (p. 59; emphasis in the original).

The chapter reports briefly on a sociolinguistic survey of a hundred high school and university students' self-perceptions of what they were doing during videochat – writing versus speaking – and what the purpose of their written chat was. Reasons given for switching from oral to written chat were *secrecy, intimacy, fun/kidding, preciseness*, and 'other' (p. 66). The bulk of Chapter 2 centers on a multimodal analysis of 300 hours of recorded multi-party videochat, followed by an appended representative multimodal transcript of the multi-party videochat interaction collected through screencasting. Screencasting makes it possible to save the video and any accompanying narration as a recording that can be replayed for review and analysis. The analysis includes brief overviews of: the use of punctuation and emoticons; quantification of total and average number of turns and of word types and tokens; and lexical analysis showing 'a wide range of topics [and] a high incidence of "metamessages"' (p. 82), that is, messages about the medium, and profiles of individual speakers. The analysis and appended transcript of Chapter 2 make clear how much individual variation there is in these and other variables of the communicative context and how many different modes and resources are in play during videochats. A main finding is that some users stuck mainly with written conversation and did not broadcast their chat in video mode; these users appeared to be less engaged than those who used a spoken mode of chat. Those who used the spoken mode were generally also those who broadcast a video of themselves and switched to written mode for various reasons, such as not wanting to wait for a turn at speaking.

Chapter 3 focuses on blogging, which, as Sindoni notes, was formerly defined by writing but is now multimodal. A blog is a web page regularly updated by its creator in individual entries ('posts' or 'postings'), which may be in the form of a written blog, audioblog, videoblog ('vlog'), photoblog, or any combination of these. A blog may include links to other websites that provide related content or sources for information. The most recent post shows first, with older entries listed, or 'stacked', underneath, and there is usually a place on the site where readers can add their own comments. Blogs vary greatly in terms of content and typically develop their own dedicated audiences based on their focus on a particular topic (e.g. health care), purpose (e.g. reporting or editorializing on current events), or individual blogger style or personality. Some bloggers identify

themselves, while others prefer to remain anonymous or use pseudonyms. Some blogs are open-access, while others require subscription and can only be accessed through a type of interface (an RSS aggregator) that allows a reader to selectively filter and feed blogs to an individual web page. Users can also highlight content by ‘tagging’ specific words, which can then be cross-associated with occurrences of the same word in other texts. Blog sites, the aggregation of these, and the tagging of keywords thus function to define like-minded internet user groups.

Because blogs are easy to set up, Sindoni notes that personal blogs are flourishing:

Since the early 2000s, everyone has been jumping on the blogging bandwagon. Actors, musicians, politicians, painters, novelists, sports people and newscasters, to name but a few, have been launching personal blogs for a variety of purposes, mainly commercial, but also driven by vanity and exhibitionism .... For this reason, a blog can also be a highly controversial piece of writing. It is rapidly written, published, read and commented on .... Anyone can easily start a blog and write about names and facts that may become public in a matter of seconds. (p. 119)

Sindoni reports (pp. 120–121) on findings by the Pew research group showing that, although interest in blogging is declining among adolescents and young adults, it is expanding among adults aged 25 to 34. She suggests that, rather than traditional blogging, or ‘macro-blogging’, young people engage in ‘micro-blogging’ to update their status on sites such as Facebook or Twitter. Thus social networking activities are not clearly distinguishable from blogging, which is in a state of rapid change and not easily definable as a unified genre. It is rather becoming, in Sindoni’s view, a ‘polymorphic’ set of online activities: ‘From a theoretical standpoint, polymorphic blogging is a definition that exempts us from an excessive trust on traditional notions of genres’ (p. 129). While their form varies, blogs have in common that they center on expression of opinion. Blog readers seem to ‘value the opinionated style of blog writing’ (p. 123) and even to find it ‘more trustworthy than the presumed “objective” style of mainstream media’ (*ibid.*).

The blog corpus which Sindoni created and analyzed, labeled ‘blogEng’, was compiled over one month (April 2012) from more than 3,200 entries by the LiveJournal blogging community. These included blogs and comments on the blogs in the following categories: *fandom and gossip, fanfiction, sport and leisure, tutorials and icons, self-help, personal life, politics, pictures and images, collectors, ask for general information (bleg)* (Table 3.1, p. 134). As the types of blogs included in this corpus make clear, the topics and activities in the LiveJournal blogging community are similar to those of social networking sites, tending to ‘coalesce around notions such as sharing, meshing up and socializing contents .... Common interests are at

the core of blogging communities' (p. 136). Sindoni grouped the semiotic resources found in her blog corpus into *language* (mainly written but with some spoken language), *images*, and *videos* (vlogs are excluded), with the order showing their degree of predominance in the blogging community. Her analysis of this blog corpus found 'a negligible presence of non-verbal resources' (p. 144) overall, as writing is the dominant mode. However, when images are used, 'writing plays a secondary role ... [and] is largely subservient to images, as it is employed to describe them, mainly in the form of captions' (p. 148).

The analysis of keywords in the blog corpus in comparison to those in the International Corpus of English, Great Britain (ICE GB corpus) '... [suggests] that some items commonly associated with written prose are less prominent than could be expected in a corpus based on written data' (p. 153) while also showing a comparatively low occurrence of keywords (interjections and contractions) that are common in spoken conversation. The blog corpus is therefore described as lying 'somewhat in-between the spoken and written genre' (p. 154) and as including:

highly hybridized and polymorphic texts. They are all written entries, but show some features that may be associated with speech. However, claiming that they are 'spoken-like' texts would be far-fetched, in so far as they also lack typical markers of spoken interaction, such as interjections or discourse markers. However, they also lack significant features related to typically written genres. (p. 155)

The analysis of three-word lexical bundles found a wide variety of functions being performed, including (in order of frequency): (1) evaluation and attribution (*stance bundles*); (2) specification (*referential bundles*); and (3) discourse marking and organization (*discourse bundles*). Overall, the analysis revealed a 'high incidence of lexical bundles expressing a blogger's stance, attitude and position' (p. 166), consistent with the opinion focus of blogging. This finding reinforces the 'in-between,' hybridized, and polymorphic status of the blog, which excludes the most typical features of both of spoken and written language and which keeps 'fragmenting into a galaxy of different constellations' (p. 169) within the blogosphere.

In Chapter 4, Sindoni turns to the online context of media-sharing communities where participants post and write comments on videos, and where interaction and language mode variation are defined by the 'conflicting dynamics' (p. 171) of: *interactivity* vs. *isolation*, *authorship* vs. *anonymity*, *verbal exchanges* vs. *video conversations*, *content evanescence* vs. *content stability*, *performance* vs. *informativity*, *repetition* vs. *creativity*, and *self-representation* vs. *self-disguise*. The chapter aims to make a contribution to understanding '[t]he strategies of personal expression

and performance within the textual web galaxy' (p. 171), a context where '[t]he opportunity to publish original and self-produced texts on the web amplifies the overexposure of personal identities, strengthened by the relationship between videos and related comments ...' (p. 172). As Sindoni further observes:

Exhibiting oneself through complex socio-semiotic strategies is within easy reach for web users, who have the unprecedented opportunity to harness linguistic, visual and multimodal strategies for self-representation .... The purpose [in this research] is to analyse discourse practices used for the *mise-en-scène* of the self, which fluctuates between attempts at authentic self-representation, performance and fiction. (*ibid.*; emphasis in the original)

The chapter continues on from the discussion of blogging in Chapter 3 by focusing on 'the dialogic dynamic of *self-representation* and *authenticity* that are two facets of spoken/written variation' (*ibid.*; emphasis in the original).

The focus in Chapter 4 is on the YouTube video sharing social networking site, where one can see '[t]he radical transformation of the once-only viewer into a producer, screenwriter, actor and broadcaster of their own video contents [that] is an aspect of the more general occurrence of mass media digitalization' (p. 172). On this site, one can see videos exhibiting:

[a]uthenticity and disguise, confession and falsity, exhibition and *mise-en-scène*. All these forms may be intertwined in multimodal texts, following plots whose textuality is made even more complex by the system of comments, which make the *master-text* (i.e., the video) as a mould for further virtually infinite discussions and debates. Such comments are multi-shaped 'adjuncts' to the master-text and shed light on the other face of representation, that of the viewers. Comments are texts that discuss other texts; they are *meta-texts* .... (p. 180, emphasis in the original)

Viewers may also comment on the comments to a video, thus making meta-comments, and they may also record their own video responding to the master-text video through remix or remake practices.

The analysis of Chapter 4 focuses on viewer responses to a 56-second viral home video, 'Charlie Bit My Finger – Again!' showing the reaction of a British child (Harry) when his infant brother (Charlie) bites his finger. The video, which was shot and posted by the children's father, is, according to Sindoni, 'the most viewed non-commercial video of all time, and one of the most commented and rated on YouTube' (p. 182). This video was selected as 'a reliable example of how a very basic, barely verbal, if any, homemade and user-generated video clip can produce such a massive *hypertrophic side-textuality*' (p. 185; emphasis in the original) – that is, how an everyday

home video can produce such an enormous body of commentary and other responsive texts. The analysis is based solely on verbal comments 'because they belong to the same context of situation, being mapped across the same constellations of textual materials, directly related to the master-text' (p. 188), which is the original video. Keyword analysis showed reduced surface form (e.g. through contraction and *that*-deletion), 'a generalized or uncertain presentation of information and a general fragmented production of the micro-texts that constitute the comments' (p. 195). Lexical bundle analysis revealed the prominence of lexis that embodied: (1) referential content referring to the video, especially the protagonist's name (*Charlie*) or key features of the event (e.g. *finger*, *hurt*, *mouth*); (2) meta-level response on the popularity of the video, especially commenting on its being the most *viewed* video; and (3) evaluative stances – predominantly but not exclusively positive – towards the video. Given that 'popularity is the reason why so many users want to comment' (p. 201) on this video, it is perhaps not surprising that viewers orient their comments to this fact. Repetition and reduplication (sequences of words such as *yum*, *babies*, *sweet*, and *Charlie*) were found to be common in four- and five-word bundles. The analysis further revealed that 'the most common bundles in both speech and writing are, quite strikingly, *less* frequent than expected' (p. 201; emphasis in the original), thus again affirming that the discourses of online contexts do not reproduce the patterns of offline contexts.

Further analysis grouped the comments into the categories of: *quotation* (from the master-text), *evaluation* (of the master-text), *questioning popularity* (of the master-text), *self-promotion* (of one's own or other video), *deviation* (non-relevant observations), and *removal* (of author text from the site). Sindoni notes that the comments 'can be very broadly placed in two functional macro-categories in terms of communicative purposes: *the first is oriented towards relevance, the second towards deviation*' (p. 209; emphasis in the original). This kind of commentative discourse is quite different from that seen in the videochats and community blogging of the previous chapters as the 'YouTube comments display lexical patterns that are very much linked to ... the foregrounded video' (p. 210) rather than to the other participants: 'Findings suggest that interaction among participants is very loose, whereas interaction between the single user and the video is high' (p. 217). Sindoni invokes a '*multimodal relevance maxim* ... to explain the low degree of relevance among comments and the high degree of relevance between each comment and the video to which they refer' (*ibid.*; emphasis in the original).

Throughout this interesting book, the author contextualizes her research within existing theory while offering her own thoughtful perspectives and discussion regarding interaction in an online environment. In Chapter 2,

Sindoni, following other researchers, situates videochat 'as an example of the evolution of the human body *in toto* in relationship to man's natural environment, technology and communication in general' (p. 89; emphasis in the original). In the chapter's final paragraph, she observes:

Participants do not ultimately interact with each other in videochats, but with the medium. The context of situation is both shared and *not* shared at the same time. Each participant is an active producer of illusions for the other, in orienting his/her gaze, in exploiting verbal and non-verbal strategies to create a fake (or virtual) environment for the other, which is only partially real. Ultimately, the backbone of reality is constituted by interaction with a machine. Verbal and non-verbal strategies add to a sense of reality while dissolving it at the same time. (p. 89; emphasis in the original)

The concluding point of Chapter 3 likewise seeks to contextualize blogging within a shared human and machine context:

The evanescence of individual writing of individual bloggers, grappling with their lives and building up personal narratives, mixes with the permanence of their own social practices within the digital world. The credibility of their stories, be they personal or political or both, is instantiated in the credibility of the medium that mainly lies in writing, revealing its secrets only to those who already know them. (p. 169)

The final section of Chapter 4 considers how comments on a massively popular home video portray the self within a global digital environment:

The self is not a monolithic and stable entity, as it is constantly fluctuating across different representations. The latter are grounded on the semiotic weight which the represented and representing subject (often but not always coinciding) decides to attribute to each *mise-en-scène*. In other words, a video may narrate the self via either a narration/confession regarding one's own personal life (e.g., by a vlog) or via segments or fragments which are considered representative of intersubjective spheres. The latter is less obviously self-narrative .... Comments written by tubers are another face of representation. They could be interpreted as a desire to leave a textual trace after other people's textual vestige. (p. 211)

As these concluding points to her core chapters illustrate, Sindoni's book is a stimulating read that both advances knowledge and makes insightful theoretical conclusions and intertextual connections with other related work. It is generally well-written and well-edited, though with a few mentions of authors whose work is not explicitly referenced and with some occasional infelicity of expression that interferes with clarity. While it provides a wealth of valuable information and substantive findings that tend to support the 'between-ness' of the online context as a 'third space'

(Pennington, 2013) for creating discourse that differs from both speech and writing, and while it is based on analysis of a large number and diversity of online interactions across modalities and resources, much of the actual analysis remains implicit. For the most part, what is offered instead of close analysis of data within the text is findings of a summative nature, together with the author's theoretical inferences from those. The reader is therefore sometimes left with a sense that there are gaps in the logical chain linking data, findings, and the author's inferences. Yet the breadth of analysis is impressive, and the author's conclusions from her data and connections to other work are insightful, offering much food for thought. It can be concluded that Sindoni's book makes important contributions to this area of research while at the same time making it strikingly clear how much more there is to tell and to discover about the rich and complex world of online communication using multimodal forms of analysis.

### About the Author

Martha C. Pennington (Ph.D. Linguistics, University of Pennsylvania) is affiliated with University of London as a Professorial Research Associate (School for Oriental and African Studies) and a Research Fellow (Birkbeck College). She is the Founding Editor of *Writing & Pedagogy* and a Series Editor for Equinox and Brill. She recently co-edited (with Alice Chik and Tracey Costley) *Creativity and Discovery in the University Writing Class: A Teacher's Guide* (Equinox) and has several books forthcoming, including *Introduction to Bilingualism and Multilingualism: People and Languages in Contact* (Wiley Blackwell) and (with Theresa Malphrus Welford) *Welcome to My World: A Writing Course* (Equinox).

### Reference

Pennington, M. C. (2013). Trends in writing and technology. *Writing & Pedagogy*, 5 (2), 155–171.

