

**PLEASE NOTE:** my dissertation was published in October 2010 with Göttingen University Press under a Creative Commons license. Go to <http://tiny.cc/corporatebloggingthesis> for direct PDF download and Amazon order information.

# **The corporate blog as an emerging genre of computer-mediated communication: features, constraints, discourse situation**

Doctoral thesis outline

Cornelius Puschmann

Department of English Language and Linguistics

University of Düsseldorf

[cornelius.puschmann@uni-duesseldorf.de](mailto:cornelius.puschmann@uni-duesseldorf.de)

The central focus of this dissertation project rests on a register study<sup>1</sup> of an emergent type of professional discourse: the *corporate blog*. A phenomenon that has only appeared in the course of the last two years, corporate or enterprise blogging is the use of blogs by business professionals to further organizational goals<sup>2</sup>. Corporate blogs serve a variety of purposes – they are used both for extra-organizational (marketing, public relations, customer relationship management, recruiting) and intra-organizational communication (internal management, expert blogging) as well as for other issues pertinent to large organizations, e.g. knowledge management.

My research approaches the object of study from two different directions. In the first part of my thesis, I will seek to describe corporate blogs empirically via a text-typographic corpus study<sup>3</sup>. My assumption is that systematically collecting and heuristically assessing textual data will contribute to a larger picture of this evolving genre. Core-linguistic aspects of corporate blogs, especially lexical, morphosyntactic and discursive-interactive characteristics will be at the center of this part of the investigation. My goal is to conduct both a qualitative evaluation of individual texts and specific linguistic constructions in them and to approach the corpus data holistically, using scalar statistical measures such as lexical density<sup>4</sup>, f-score<sup>5</sup> and propositional idea density<sup>6</sup>, as well as degree of dimensional variation<sup>7</sup>. I will attempt to evaluate the usefulness of inductive (data-

---

1 I use both the terms *register* and *register study* with the meaning assigned to them by Biber and Finegan (1994), who lament that “the analysis of genre, register and stylistics is seen as relating to written varieties and as based on literary and rhetorical interests rather than linguistic ones”. This view in the discipline, paired with a noticeable neglect of written material that is still partly common in many areas, has led to a relatively low volume of publications on the topic of (written) registers, at least outside of the domain of historical linguistics. The field thus appears fragmented between (computer-mediated) discourse analysis, traditional stylistics, ethnographic and rhetorical approaches and pragmatics (where the focus – similar to sociolinguistics – seems to be largely on synchronous communication, e.g. chat and instant messaging).

2 cf. Weil (2006).

3 With this approach I follow Biber's (1992) suggestion of a *macroscopic analysis* of a written genre.

4 cf. Stubbs (1996).

5 cf. Heylighen and Dewaele (2002).

6 As implemented by Brown et al (2007).

7 Following Biber (1992), but with a restricted set of features.

driven) methods in this context<sup>8</sup> and use data from reference corpora to contrastively analyze corporate blogs in relation to other text types, both along the domain and modal axes<sup>9</sup>.

Proceeding from this empirical and largely quantitative view on the text type, I will then switch to an assessment of the corporate blog as a domain-specific genre or register, that is, to the contextual metastructure framing the process in which individual textual instantiations are produced. While the first part uses only the linguistic data itself as evidence (and is thus purely empirical) the second integrates meta-discursive observations from the speech community<sup>10</sup>, as well as the highly significant technological situatedness of blogs into the description. By adding the pragmatic and sociotechnical dimensions to the picture, I will seek to explain the emergence of corporate blogs plausibly as more than just an inevitable technological development, but instead as the result of a shift from one-to-N-oriented, unidirectional broadcasting to a communicative paradigm more similar to immediate interpersonal communication. I will illustrate that corporate blogs are seen by their proponents as a means of overcoming a looming communicative crisis threatening institutions because of their static and highly conventionalized genre practices, which often seem opaque, construed and manipulative to outsiders.

A central point of reference in both parts of the investigation will be the strategic, expressive and discursive creativity of the individual blogger, the stylistic variability that results from it, and how this degree of individual freedom of expression can be aligned with organizational goals. The question of whether stylistic heterogeneity must always be the passing characteristic of an emerging genre or whether it can be its core constituent will be reviewed. Furthermore, questions such as the relevance of authorial authenticity and reliability will be discussed in the light of deceptive practices such as *ghostblogging* and *flogging*, which are considered moral transgressions by many non-corporate bloggers.

Because it is devised as a systematic study of both textual instantiations (text typology) and of the socio-cognitive frame in which they produced (register/genre), this work must take theoretical models and approaches from multiple areas of linguistics and from entirely other fields into account.

My conception of genre as functionally underdetermined closely follows Swales, who describes a genre as “a class of communicative events [...] with a shared communicative purpose” (1990). In the same vein Bazerman speaks of typified texts and remarks that “by using [them] we are able to

---

8 The context here is the notable contrast in how different disciplines approach the concept of genre. Many researchers involved in computational linguistics and computer science have a utilitarian concept of genre, regarding it as a system for the labeling of text in clearly delineated categories which are primarily formally defined. In contrast to this, rhetorics and other theory-minded fields seem to have largely abandoned form and instead see the functional dimension of a genre as the constitutive element.

9 That is, a comparison of corporate blogs both with other types of corporate discourse (memo, mission statement, press release) and with other types of blogs (political, personal).

10 In the tradition of an ethnographic study of a domain-specific variety, such as those conducted by Smart (2006) and Artemeva & Freedman (2001).

advance our own interests” (1994) - something certainly very closely aligned with the goal-oriented corporate mindset. The conductive role of genres in social interactions is equally stressed by Miller (1984) and later applied specifically to blogs (Miller and Shepard 2004). Work in the ethnographic tradition of genre research especially in the domain of professional communication is exemplified by Artemeva (2005) and Smart (2006) and based on prior theoretical exploration, especially by Bhatia (1993) and Berkenkotter and Huckin (1995). Important contribution from the field of research into genres of computer-mediated communication and specifically blogging come from Herring et al (2004, 2005) who draw on previous work in CMC genre analysis (Yates and Orlikowski 1992, Orlikowski and Yates 1994), as well as from Efimova and de Moor (2005), Anjewierden and Efimova (2006), Efimova and Grudin (2007) and Schmidt (2007). The approaches used in these studies propagate a holistic view of blog research, not limiting themselves to linguistic features, but also investigating aspects such as network structure and blogger motivation. Works with a strong applied and computational focus also investigate blogs, analyzing large amounts of language data in the context of automated genre classification (Santini 2004, 2005), authorship attribution (Stamatatos et al. 2000, 2001; Koppel et al. 2003), age (Schler et al. 2006) gender (Koppel et al. 2002) and personality (Nowson et al. 2005). In many of these studies feature sets such as that developed by Biber are devised<sup>11</sup> and then used to test a data set for statistically salient patterns such as lexeme frequencies, part-of-speech frequencies and n-gram frequencies. The theoretical considerations of rhetorics and traditional genre studies seem furthest away from these approaches, which share a largely utilitarian outlook on genre.

To summarize, my goal will be to integrate these diverse approaches into a complete register study to the extent possible, because all of them have something to offer from the point of view of the researcher. I am guided by the realization that a text is more than a sequence of words and that a genre is in turn more than the sum of its collected instantiations, while at the same time recognizing that the systematic study of variation among texts using statistical measures has opened a completely new avenue into linguistic investigation.

## References

- Anjewierden, A., and L. Efimova. “Understanding weblog communities through digital traces: a framework, a tool and an example.” *On the Move to Meaningful Internet Systems 2006: OTM 2006 Workshops*. 4277/2006 (2006): 279-289.
- Artemeva, N. “A Time to Speak, a Time to Act: A Rhetorical Genre Analysis of a Novice Engineer's Calculated Risk Taking.” *Journal of Business and Technical Communication*. 19.4 (2005): 389.

---

<sup>11</sup> In contrast to Biber's original method many newer approaches focus more strongly on syntactic features, for example in the form of POS n-grams (cf. Santini). This is due largely to an increased reliability of automated taggers and thus less need to focus largely on lexical types, as was the case in the past due to computational constraints.

- Artemeva, N., and A. Freedman. "'Just the Boys Playing on Computers': An Activity Theory Analysis of Differences in the Cultures of Two Engineering Firms." *Journal of Business and Technical Communication*. 15.2 (2001): 164.
- Bazerman, C. "Systems of genres and the enactment of social intentions." *Genre and the new rhetoric*. (1994): 79-101.
- Berkenkotter, Carol, and Thomas N. Huckin. *Genre Knowledge in Disciplinary Communication: Cognition, Culture, Power*. L. Erlbaum Associates, 1995.
- Bhatia, V. K. (Vijay Kumar). *Analysing genre: Language Use in Professional Settings*. Longman, 1993.
- Biber, Douglas. *Variation across Speech and Writing*. Cambridge University Press, 1992.
- Biber, Douglas, and Edward Finegan. *Sociolinguistic Perspectives on Register*. Oxford University Press, USA, 1993.
- Efimova, L., and A. de Moor. "Beyond Personal Webpublishing: An Exploratory Study of Conversational Blogging Practices." *System Sciences, 2005. HICSS'05. Proceedings of the 38th Annual Hawaii International Conference on*. (2005): 107a-107a.
- Efimova, L., and J. Grudin. "Crossing Boundaries: A Case Study of Employee Blogging." *System Sciences, 2007. HICSS 2007. 40th Annual Hawaii International Conference on*. (2007): 86.
- Herring, S.C. et al. "Bridging the gap: a genre analysis of Weblogs." *System Sciences, 2004. Proceedings of the 37th Annual Hawaii International Conference on*. (2004): 11 pp.
- Herring, Susan C. et al. "Conversations in the Blogosphere: An Analysis 'From the Bottom Up'." *Proceedings of the 38th Hawaii International Conference on System Sciences (HICSS'05)*. 4 (2005): 107b.
- Heylighen, F., and J. M. Dewaele. "Variation in the Contextuality of Language: An Empirical Measure." *Foundations of Science*. 7.3 (2002): 293-340.
- Koppel, M., and J. Schler. "Exploiting Stylistic Idiosyncrasies for Authorship Attribution." *Proceedings of IJCAI'03 Workshop on Computational Approaches to Style Analysis and Synthesis*. (2003): 69-72.
- Koppel, M., S. Argamon, and A. R. Shimoni. "Automatically Categorizing Written Texts by Author Gender." *Literary and Linguistic Computing*. 17.4 (2002): 401.

- Brown, C. et al. "Measuring Propositional Idea Density through Part-of-Speech Tagging." (2007). Annual Meeting of the Linguistic Society of America, Anaheim, Ca.
- Miller, C. R. "Genre as social action." *Quarterly Journal of Speech*. 70.2 (1984): 151-167.
- Miller, C. R., and D. Shepherd. "Blogging as Social Action: A Genre Analysis of the Weblog." *Into the Blogosphere: Rhetoric, Community, and Culture of Weblogs*. University of Minnesota. <http://blog.lib.umn.edu/blogosphere>. (2004).
- Nowson, S., J. Oberlander, and A. Gill. "Weblogs, genres and individual differences." *Proceedings of the 27th Annual Conference of the Cognitive Science Society*. (2005): 1666–1671.
- Orlikowski, W. J., and J. Yates. "Genre Repertoire: The Structuring of Communicative Practices in Organizations." *Administrative Science Quarterly*. 39.4 (1994).
- Santini, M. "A Shallow Approach to Syntactic Feature Extraction for Genre Classification." *Proceedings of the 7th Annual Colloquium for the UK Special Interest Group for Computational Linguistics*. (2004).
- . "Genres In Formation? An Exploratory Study of Web Pages using Cluster Analysis." *Proc. CLUK*. 5 (2005).
- Schler, J. et al. "Effects of Age and Gender on Blogging." *AAAI 2006 Spring Symposium on Computational Approaches to Analysing Weblogs (AAAI-CAAW)*. (2006): 06-03.
- Schmidt, J. "Blogging Practices: An Analytical Framework." *Journal of Computer-Mediated Communication*. 12.4 (2007).
- Smart, Graham. *Writing the Economy: Activity, Genre, and Technology in the World of Banking*. Equinox Publishing (UK), 2006.
- Stamatatos, E., N. Fakotakis, and G. Kokkinakis. "Computer-Based Authorship Attribution Without Lexical Measures." *Computers and the Humanities*. 35.2 (2001): 193-214.
- Stamatatos, E., G. Kokkinakis, and N. Fakotakis. "Automatic text categorization in terms of genre and author." *Computational Linguistics*. 26.4 (2000): 471-495.
- Stubbs, Michael. *Text and Corpus Analysis: Computer-Assisted Studies of Language and Culture*. Blackwell Publishers, 1996.
- Weil, Debbie. *The Corporate Blogging Book: Absolutely Everything You Need to Know to Get It Right*. Portfolio, 2006.

Yates, J., and W. J. Orlikowski. "Genres of Organizational Communication: A Structural Approach to Studying Communication and Media." *The Academy of Management Review*. 17.2 (1992): 299-326.