If Pandora had a Blog: Towards a Methodology for Investigating Computer-Mediated Discourse

Otilia Pacea, Ovidius University, Constanta, Romania

ABSTRACT

In the context of internet genre migration and proliferation, conventional taxonomies are no longer valid. To classify blogs between thematic and personal blogs is to blissfully ignore the legions of successful content prosumers, from political blogs to travel blogs, from food blogs to MAD (mom and dad) blogs, from fashion blogs to milblogs. With the recent explosion of social media, the digital landscape shifted and today there are more voices online than ever before. For blogs, however, the original purpose for communication has always been twofold: to inform and to emote. Computer-mediated communication may be overpopulated with a myriad of mixed forms and blogs might be dead or simply, difficult to reach with so much overlapping. Yet high-impact blogs still remain and are widely read. This paper explores the language of high-impact blogs, testing a new methodology for genre analysis to solve genre hybridity in the case of computer-mediated discourse.

Keywords: Blogs, Classification, Computer-mediated Discourse, Diction, Empirical Measure, Genre, Hybridity

THE BLOG IS DEAD: LONG LIVE THE BLOG!

In 2013, Jason Kottke, an American blogger and former web designer, who maintains one of the longest continuously running blogs on the web, proclaimed the death of the blog. He argues that the original informational task that blogs have been fulfilling for the past two decades is now being taken over by various other media forms that have recently emerged, many of which still featuring blog-like content and form, but consistently evolving and mutating into more tightly integrated services such as Facebook and Twitter.

Recent statistics speak rather of a rebirth of the blog with the advent of the social media: over 6 million people are reported blogging on blogging sites as compared to 12 million people who are still blogging but via social networks (ignitespot.com). According to the same statistical report, 77 percent of internet users read blogs today. Even more optimistic reports suggest a number of over 4 million blog posts that are written every day (internetlivestats.com). Awareness and usage of blogs is not yet yesterday’s news: major corporations but also lower-ranked companies still populate the blogosphere to engage more customers. Statistics confirm that 61 percent of American consumers have made a purchase based on a blog post (ignitespot.com). Consequently, more than a third of Fortune 500 companies still maintain active blogs (Forbes).

DOI: 10.4018/IJSSS.2015070102
Also many spread the word of a modest decline in the younger blog population. The overt preference of the so-called millennials (aged 18-33) for the quickly-catching social networking sites is in fact counterbalanced by the increased blog popularity among older generations (pewinternet.org). What remains, though, without doubt, are the blog-like features such as links, personal updates and musings that percolate the digital discourse today. The form might be dead, but the content is to stay.

Of all such prominent characteristics, it was first and foremost the confessional character of blogs that caught the eye of seminal blog research (Herring, 2004, 2005, 2006; Blood, 2002; Crystal, 2006 etc). The personal features were overemphasized and highlighted in dictionary definitions as well. When Merriam-Webster dictionary publishers crowned “blog” word of the year in 2004, they did not overlook its personal character in defining “blog” as a Web site that contains an online personal journal with reflections, comments, and often hyperlinks provided by the writer. Prior to the Merriam-Webster award in 2004, in March 2003, the Oxford English Dictionary was the first to add “weblog” (both noun and verb) as well as “blog” to its corpus. Personal hints are aimed in OED definition of the blog that refer to a frequently updated website consisting of personal observations, excerpts from other sources, etc. typically run by a single person, and usually with hyperlinks to other sites; an online journal or diary. The personal features in blogs were rightfully overstated since they link to the fact that in their blogging practices, individuals control language in a way that was quite new, not found elsewhere, as suggested by David Crystal in the most engaging and provocative large-scale study on the impact of the internet on language use: “[…] a variety of writing intended for public consumption which appears exactly as the author wrote it, which is not constrained by other genre conventions, and which privileges linguistic idiosyncrasy. I call it, on analogy with free verse, free prose” (Crystal, 2006, p. 246).

Blogs include a perplexing wide variety of subgenres, from the already-established personal blog, journal or lifelog (Schaap, 2004), the filter, the knowledge logs or k-logs (Herring et al., 2004, 2006) to the corporate blogs (Puschmann, 2010; Cho and Huh, 2010) but also recently-emerged forms such as milblogs (Wall, 2010), travel blogs (Banyai and Glover, 2013), food blogs (Salvio, 2012), fashion blogs (Rocamora, 2011), mommy blogs (Lopez, 2009), expatriate blogs (Pacea, 2010) etc.

On content criteria, blogs may be thus classified into personal blogs, corporate blogs, travel blogs, food blogs, fashion blogs, book blogs, expatriate blogs, MAD blogs (mum and dad blogs) etc. On auctorial criteria, blogs can be created and maintained by politicians (the political blog), by professional journalists (the journalist blog), or by researchers or experts (the academic blog, the educational blog or the personal development blog). Most blog definitions are in general trait-based in the fashion of the so-called recipe theory in that their boundaries are built on a convention of like-statements in conjunction with not-statements (cf. Freedman’s recipe theory).

To address so much content overlapping that has significant effects on genre genealogy, a unifying framework for analysis is proposed in this paper. Language data are collected from the most-linked Technorati-ranked blogs. Computer-aided text analysis methodology (DICTION, Hart, 1997) is employed to calculate the level of hybridity in high-impact blogs which may vary from highly-informative and detached styles to personal and involved styles. Individual language scores which are generated by DICTION’s standard dictionaries are concatenated to outline the overall tone and theme of the blog posts that can be classified accordingly. The findings are correlated with existing blog classifications to validate a universal computer-assisted corpus-based methodology for genre analysis in computer-mediated discourse.