

## DIGITAL STORYTELLING: TEMPORAL NARRATIVES IN EXPAT BLOGS<sup>1</sup>

Otilia Pacea  
Ovidius University, Constanta

*Time present and time past  
Are both perhaps present in time future  
And time future contained in time passed  
(T.S. Eliot, "Burnt Norton")*

### 1. Prolegomena to Digital Storytelling

The Wild West or otherwise the W(ild) W(ild) W(est) contains the universally popular and globally recognizable image of the lonesome individual, the cowboy, who emerged from the state of nature, the wilderness, to build a civil society, the frontier community. In the digital mirror held by popular culture, we rediscover practices of yore in the cyberspace: the digital cowboy, the blogger, who surfaced from the digital giant, the World Wide Web, to build a community of discourse, the blogosphere. The research question addressed in this study is to determine the extent of American popular culture infusion at the level of discourse, more specifically, how the weblog – this naked form of public writing that cut loose from the mainstream, viz. from any form of censorship or editorial control – creates meaning and generates patterns of social behaviour through the agency of language use.

On the digital frontier, the blog has been established as the fastest growing genre of internet communication (Herring et al. 2004). More than a decade ago, Jørn Barger was the first acknowledged blog practitioner who coined the term *weblog*, a portmanteau word blending "web" and "log", to describe the process of logging his online discoveries as he was making them (Pacea 2008: 257). One year later there were only a few sites similar in nature to the original one. At a time when those who blog did not know that

---

<sup>1</sup> This work is part of the research project "Translations of American Myths, Icons and Brands in Post-Communist Romanian Culture (TRANSMIT)", supported by CNCSIS-UEFISCSU, Project number PNII – IDEI – 802/2009.

what they were doing was indeed blogging, Jesse James Garrett and Cameron Barrett were already two popular names in the blogosphere. The former compiled a list of “other sites like his” as he found them in his travels around the web. In November 1998, he sent that list to Cameron Barrett, who published it on his own weblog *Cameworld* and soon others sent their URLs to be enlisted. At the beginning of 1999, it was easy to keep track with only twenty-three blogs included in Cameron Barrett’s list but soon more and more internet users became interested. It was just a matter of few months until a whole community emerged. Shortly thereafter, in early 1999, Peter Merholtz announced that he was going to start pronouncing *weblog* wee-blog → (we) blog → blog as well as to refer to those who employ the new form of e-writing as “bloggers” (Blood 2002 quoted in Rodzvilla 2002). The term *blog* has thus already established its lexical supremacy not only among ordinary internet users but also among linguists: Merriam-Webster dictionary publishers, who crowned “blog” *Word of the Year* in 2004, did not wait longer and one year later they included the word “blog” in their print edition. Google the word today and the search will be rewarding: 70 million hits for the key search word “weblog” vs. over 2 billion hits for “blog” in June 2010.

In general, a blog is a website that contains an archive of regularly updated articles, which are called blog posts entries or simply, posts. They are written mostly by one author (although multi-authored blogs are also available) and presented in a reverse chronological order, foregrounding the latest. The posts are generally made available to the public, unless the blog author intends otherwise, in which situation blog authors do have the option to control the public access to their blog content by simply protecting it by password. The posts are mostly raw texts but hyperlinks and other media such as picture, video or sound files can contribute as well. A special blog feature allows readers to comment on blog posts and publish their comments. Conversely, blog authors may disallow certain comments that are written by readers if they consider them inappropriate. In what regards their content, two major categories divide blogs between **personal** and **thematic** blogs. The first category includes (individually authored) personal journals, in which blog authors discuss their own life. The second category can be divided between filters and k-logs (knowledge logs), in which blog authors discuss general or specific topics such as current events, politics, arts, entertainment, sports or technology, rather than their own private life, the former featuring blogs that filter the internet based on the author’s interests and opinions and k-logs (knowledge logs) and the latter featuring blogs that contain information and observations focused around an external topic (Pacea 2010, in press).

The first blog editors were web enthusiasts, in general web designers and IT programmers. The original weblogs or protoweblogs used to locate

information efficiently at a time when search engines could not. They were link-driven sites, a unique proportioned mixture of links, commentary, personal thought and essays. Many blogs today still have the same original style. However, the influx of blogs over time has changed the original definition of *weblog* from “a list of links with commentary and personal asides” to “a website that is updated frequently, with new material posted at the top of the page” (Blood 2000).

In the blogosphere, if not ubiquitous yet certainly pervasive, is the **expat blog** that is created and maintained by self-exiled blog authors who have experienced relocation in an English-speaking environment, which can be extended to refer to the digital environment as well (Pacea 2010, in press). In this study, the expat blog defines midway between personal journal and thematic blogs - a hybrid blog subgenre. Formally, it contains the reverse-chronological internet publication of expatriate notes drawing on the experience before, meanwhile and after temporary or permanent relocation.

Of all researchers who have over-capitalised on the opportunities of the newly emerged internet culture, the linguist cannot but take delight in exploring its latest electronic genres that surfaced from the ecology of public writing. Such spontaneous language production beyond any standardizing process, “written language in its most naked form” (Crystal 245), is most appealing.

Amid a high-tech revolution, a new linguistic variety is born of parents of modern technology. Technical constraints work in both directions with impact on the new electronic medium at large and on the linguistic variety in particular. Variation according to user – *a set of linguistic forms which have a regular connection with a particular social situation* – and variation according to use – *a set of situations* – classify linguistic variety as the most appropriate concept to define the language of the internet, also due to the nature of the connection which is not obligatory but functional: “enough of a connection to give speakers an intuitive feeling of the existence of a link between the language and the situation” (Crystal :101).

Blogging has been widely recognized as an important internet genre that brings the most substantial contribution to the internet linguistic variety with an internet audience of 188.9 million reported in August 2008 (Technorati’s annual report *State of the Blogosphere 2008*). More interested in non-empirical and non-linguistic approaches that pioneered electronic communication research such as the social impact of blogging (Blood 2002; Gilmore 2003; Lasica 2002, 2003; Park 2003; Delwiche 2004), the content analysis of blogs (Herring et al 2004, 2005, 2008; Papacharissi 2004), the rhetorical analysis of blogs (Miller and Shepherd 2004), few scholars turned to the analysis of the linguistic properties of blogs that remain yet under-researched. Among notable exceptions, Herring and Paolillo (2006) used selected linguistic features (the relative frequency of pronouns, determiners,

and certain function words; cf. Argamon et al 2003) to analyze gender differences in blogs, and Puschmann (in press/2010) examined variation in expressions of futurity in blogs.

## **2. Research Question**

In a longitudinal study of linguistic variation we advance the paradigm of the temporal divide as the major classifier for the expat blog subgenre whose presence in the blogosphere is consistent. Using google search technology to rate the self-publishing phenomenon, more than half a million blog feeds on key word “expat blog” popped up in June 2010. From the underlying dimensions of functional linguistic variation already identified in the literature (Biber 1988, 1995, 1998) that produced the primary blog subgenres - the personal blog and the thematic blog, we selected two linguistic features, present tense verbs as positive loading and past tense verbs as negative loading to build the paradigm of the temporal divide: immediacy vs. remoteness. We also included two content categories we adapted from a text analysis software programme – Diction 5.0 – that is under-used in the linguistic study of the language of blogs (Huffaker, Calvert 2005): concreteness and complexity. Linguistic variation between English native and English non-native speakers’ selection of such linguistic features that we classified as immediacy vs. remoteness was analysed in two separate corpora: Corpus of Expat Blogs in English (CEBE) and Corpus of Expat Blogs in English as a Foreign Language (CEBERO). Based on linguistic co-occurrence, two electronically-dependent patterns of social behaviour emerged that mirror the paradigms of the temporal divide: the social actor and the bon vivant. Quantitatively, we hypothesized that expat blogs roughly correspond to Diction text types of newspaper editorials and novels/short stories. Qualitatively, we interpreted linguistic data in relation to the external functions.

## **3. Corpus Design**

Previous empirical research was conducted for blogs that were created and maintained in English, although there was no interest in questioning the use of English between native and non-native; in fact reference to language proficiency was never made. In recent years, the construct of ‘advanced foreign language proficiency’ has begun to attract increased attention (Byrnes and Maxim 2004; Leaver and Shekhtman 2002). Foreign language proficiency in the electronic medium represents a new frontier [sic] in foreign language education circles (Belz, Reinhart 2004). To my knowledge, no electronic corpora are available for EFL internet linguistic variety.

The present study is based on the analysis of two snowball samples of 120 blog entries that were randomly selected from 40 single-authored blogs, following links from a source blog, three posts per blog, amounting to a pool of 56,774 words. The blog entries were controlled to meet the expatriate condition in the sense that we opted for those entries that provided an account of the expatriate experience in the case of the blog authors who might have relocated in their native country. Blogs were handpicked upon the condition that were created by American expatriates who blog in English in the case of one corpus and that they were written in English by Romanian expatriates who use EFL to blog about their expatriate lives in the case of the other corpus. The data was available in reversed chronological from the earliest post collected in 2004 to the latest collected in May 2009. There are several directory sites, which provide paired listings of the various nationalities of the expatriate bloggers and the countries they belong to, among which the most productive is [expat.blog](http://expat.blog) ([expat-blog.com](http://expat-blog.com)) - a rich, helpful and user-friendly website created by the expatriates for the expatriates.

Two corpora were created: Corpus of Expatriate Blogs in English for American English Natives (CEBE) and Corpus of Expatriate Blogs in English as a Foreign Language for Romanian Natives (CEBERO). The digital frontier of the American expat blog network extends from Germany, Turkey, Holland, Argentina, Saudi Arabia, France, UK, Panama, Romania and Thailand; whereas the Romanian expat blog network includes mostly the United States from Washington DC, Louisiana, Illinois, Atlanta to Pittsburgh, Alabama, Colorado but also UK, Ireland, France, Czech Republic, Poland, Finland, Spain, Germany and Lebanon. Our assumption is that such a geographically and culturally complex network expatriate blogs is more productive than an exhaustive data collection that may offer a panoramic view of the blogging phenomenon but less homogeneous results at the level of discourse.

Affiliation and language choice (English) were the main criteria for our list of expatriate blogs. Previous gender-focused blog research is in favour of female bloggers posting more as well as more frequently than male bloggers (Huffaker, Calvert 2005). Indeed, of the twenty most active bloggers in each of the two corpora, there were fourteen female and six male blog authors. Because of the abundance of female blog authors, the sample was separated by gender and a stratified analysis was then conducted. Three blog entries were selected for each of the twenty blogs per corpus, which finally amounted to 38,285 words per CEBE and 18,489 words per CEBERO. Our final sample thus comprises 120 blog entries drawn from twenty weblogs per corpus which were gender-controlled, 14 female blog authors and 6 male blog authors.

#### 4. Methodology

Consistent with other empirical analysis of web genres (Bates & Lu 1997, Ha & James 1998, McMillan 1999, Bauer 2000, Huffaker, Calvert 2005), we employed methods of content analysis to identify and quantify linguistic properties of the blogs in our corpus.

Diction 5.0 is a text analysis software programme, extensively utilized for its reliability, validity and popularity among communication researchers and subsequently for lexical analysis, viz. for the study of vocabulary and word choice. This lexically based software programme searches texts for five semantic features (the master variables) and thirty-three sub-features (separate dictionaries such as familiarity, human interest, concreteness, tenacity, numerical terms, ambivalence etc). Each semantic feature is scored on the basis of the 33 dictionaries according to a series of formulas incorporated into the software.

The programme presents normative data for the five variables and thirty-three sub-variables, which invites the user to locate a given passage in the universe of discourse. These normative data have been generated by running more than 22,000 contemporary language documents through Diction, which amounts to a 10,000 word corpus. They range from public speeches to poetry, from newspaper editorials to music lyrics, from business reports and scientific documents to TV scripts and informal telephone conversations. The default in Diction is to combine all these data into a single normative profile so the user can see what a passage looks like in general. Most of the data for the normative values are generated from a broad sampling of (print) texts produced primarily in the United States. Diction's master variables were created with the rationale that "if only five questions could be asked of a given passage, these five would provide the most robust understanding" (Hart 2001: 45). The choice of each variable was deliberate and inspired by the theoretical work of a number of social thinkers. From the range of sub-variables, we selected three sub-features: present concern, concreteness and past concern.

A second category of variables used by Diction are calculated variables which are based on word ratio, from which we adapted complexity, a measure of word size based on Flesch's (1951) notion that convoluted phrasings make a text's ideas abstract and their implications unclear (*Rhetorica.Net*. Web).

The data produced in our Diction analysis was supported by the findings of a second computer-aided text analysis – Linguistic Inquiry and Word Count (LIWC) – that determines the rate at which the authors/speakers use positive or negative emotion words, self-references, big words, or words that refer to sex, eating, or religion within emails, speeches, poems, or transcribed daily speech (Pennebaker, Booth, Francis 2007).

## 5. Discussion

The approach builds on the centrality of linguistic co-occurrence in analysis of text variation (Ervin-Tripp 1972, Hymes 1974, Biber 1988). Co-occurrence reflects shared function: features co-occur frequently in texts and point to the communicative functions they serve, which are associated with the situational contexts, in our case the electronic medium. As argued above, from the set of linguistic features that were identified in the literature as either positively or negatively influencing the constituency of the most prominent textual dimensions, we selected two major co-occurring features, viz. present tense verbs and past tense verbs, which together with concreteness and complexity, build our temporal paradigm of the temporal divide:

1. Immediacy: Present + Concreteness
  2. Remoteness: Past + Complexity
- (Table 1)

The text analysis programme allowed us to compare the expat blog subgenre to thirty-six distinct genres, among which we selected two categories: journalism with its sub-category newspaper editorial and literature with its sub-category novels/short stories. The former represents an assortment of newspaper editorials from diverse geographical locations in the United States with topics ranging from sexual harassment, infant mortality, casino gambling, drug enforcement and other public matters. All editorials were written in the 1990s and obtained from online versions of these papers. The latter compiled fiction from a diverse collection of seminal writers from Nathaniel Hawthorne, Stephen Crane, J.D. Salinger, James Fenimore Cooper, Lewis Carroll, Kate Chopin and Francis Bacon; the excerpts were taken from such online sites as the Internet Classics Archive and Book Lover's Resource Page.

Our first research goal was to quantify the temporal divide in both corpora and thus establish a linguistic identity for the expat blog subgenre on the basis of linguistic co-occurrences patterned on the dichotomy immediacy vs. remoteness.

Present tense verb and concreteness raw frequencies overlap in newspaper editorials and expat blogs. Past tense verbs occur with higher frequency in expat blogs than in newspaper editorials: 55% of the expat blog entries in the Romanian expat blog corpus and 45% in the US expat blog corpus scored above the highest point in the normal range for past concern. Complexity scores were lower for expat blogs with 15% of the expat blog

entries in the Romanian expat blog corpus and 70% of the expat blog entries in the US expat blog corpus below the lowest point in the normal range.

Present tense verbs occur with higher frequency than in short stories/novels (30% of the expat blog entries in the Romanian expat blog corpus and 70% of the expat blog entries in US expat blog corpus scored higher than the highest point in the normal range). Concreteness scores were quite similar (5% of the expat blog entries in the Romanian expat blog corpus below the lowest point in the normal range and 10% of the expat blog entries in both corpora above the highest point in the normal range). Past tense verb and complexity raw frequencies overlap in short stories/novels and blogs.

Linguistic variation on the immediacy vs. remoteness establishes the sub-genre of the expat blog between newspaper editorials and fiction, closer to newspaper editorials in the use of present tense and concreteness that constitute the immediacy dimension of the temporal paradigm, which can be associated with LIWC scores in favour of formal texts (Table 2). As regards the remoteness dimension of the temporal divide, the expat blog is closer to short stories/novels in the use of past tense and complexity.

Also computed in comparison to fall into one of the two temporal paradigms, immediacy raw frequencies tend to overlap in the Romanian and American expat blog corpora. Remoteness raw frequencies are similar insofar as the lowest point in the normal range is concerned. What can be interpreted as a significant point of difference is the value for the highest point in the normal range, with the remoteness raw frequencies in the Romanian expat blog corpus higher than the highest point in the normal range.

Drawing on Biber (1988, 1995, 1998) we also advanced comparative research between corpora of English native and English non-native but advanced proficiency users on the assumption that when past tense occur with high frequency in a text, present tense verbs tend to be notably less used in that text, viz. the more frequently past tense verbs, the less frequently present tense.

A nonparametric rank statistic was used as a measure of the strength of the associations between two variables to compute correlation between two data series. For the data obtained computing raw frequencies for present and past tense verbs in CEBERO, Spearman's  $r = -0.115$ : the critical value of  $r$  for a number of 20 scores is .447 ( $\alpha = .05$ ). Although there appears to be a weak negative association between present and past tense verbs use, the value of the rank order correlation is not statistically significant. Present and past tense verb frequencies in CEBE resulted in a statistically different value, i.e. Spearman's  $r = -0.6105$ : the critical value of  $r$  for a number of scores of 20 is .447 ( $\alpha = .05$ ). The correlation is 61% as stronger as it possibly could be.



There is a significant negative relationship between present and past tense verb use in American expat blogs, which confirms the hypothesis that if present tense verbs occur more frequently, past tense verbs tend to occur less frequently. Hence, while a strong negative correlation exists between present and past tense verbs use in American expat blogs, no association exists in Romanian expat blogs.

## 6. Conclusion

The difference in linguistic variation across corpora that has resulted can be linked to external criteria, situationally-influenced, altogether contributing to the advancement of patterns of social behaviour that emerged online and are consequently projected offline. Under the heading category, the digital cowboy we recycled from the western mystique and already established in a previous analysis of an extended corpus of expat blogs (Pacea 2010, in press), we contend that two sub-categories become prominent: the social actor and the bon vivant.

“A little time will melt her frozen thoughts” says the Duke in *Two Gentlemen of Verona* (III.ii.9). It is how the social actor perceives time in the expat blog subgenre that is created and maintained by Romanian natives but English fluent. The social actors are fully aware they play a part. They perform actively onto the social arena (Pacea 2008: 264). They seek to adapt yet they are trapped in their non-native past (take for instance the higher scores obtained by the Romanian native expat bloggers for past concern). Fragmentation defines the temporal narrative in the expat blog subgenre. The social actors cannot escape their geographical and first-language past, hence Romanian-oriented past whereby the temporal paradigms are well-defined. Also in addition to technological constraints and digital time, their present can refer virtually to any time along the time. Past is present and present is past or anything in between. Also, most expat bloggers changed location for personal and professional development: either to continue studies abroad or for better job positions. Possible scenario: they continued their studies, found a significant other, remained. They have never closed the door to their geographical past. Fragmentation is self-conscious, make-believe, artificial.

If for the social actors, English is the language of choice, for the bon vivant represented by the American expat bloggers who comfortably blog about their expatriate experience, the linguistic choices they make and that are relevant for our temporal paradigm are reminiscent of the First Voice’s opening monologue in Dylan Thomas’s *Under Milk Veil*: “Time passes. Listen. Time passes”. Fragmentation is inviting as well as a conscious, genuine construct of the temporal narrative which appropriates content in the expat blog subgenre by English natives to conventional travel books or logs

that classify midway between involved production immediacy (involved production) and remoteness (elaborated production).

Our preliminary findings shed light on the nature of the electronic linguistic variety, more specifically on one of its most productive genre – the blog. The linguistic analysis of the expat blogs consolidates research in the new linguistic variety that is yet under development. Our goal was to set up a basis for investigating language that is used in the electronic environment and that can evolve into more complex tasks in the future. Our preoccupation with native and non-native English use in the language of blogs has also pioneered empirical research in corpus linguistics on the internet. “Real-world” or in our case virtual-world language use that has become more accessible in corpora, which are now open to all linguists, offer unprecedented research opportunities to check intuitions about language in general, similarities and differences in the use of certain words or phrases, and to test the prominence of certain expressions, new words included.

**Table 1. Diction’s Category Components for the Expat Blog Subgenre**

<b>Subtractive/ Additive Category Component</b>	<b>Effect on Certainty</b>	<b>Definition</b>	<b>A Selection of Examples</b>
<b>Present Concern</b>	[+]	A selective list of present-tense verbs extrapolated from C.K. Ogden’s list of general and picturable terms, all of which occur with great frequency in standard American English. The dictionary is not topic-specific but points instead to general physical activity, social operations and task-performance.	cough, taste, sing, take, canvass, touch, govern, meet, make, cook, print, paint
<b>Concreteness</b>	[+]	A large dictionary possessing no thematic unity other than tangibility and materiality. Included	peasants, African-Americans, Catholics, carpenter, manufacturer, policewoman,

		are sociological units, occupational groups and political alignments. Also incorporated are physical structures, forms of diversion, terms of accountancy and modes of transportation. In addition, the dictionary includes body parts, articles of clothing, household animals and foodstuffs and general elements of nature.	communists, congressman, Europeans, courthouse, temple, store, television, football, CD-ROM, mortgage, wages, finances, airplane, ship, bicycle, stomach, eyes, lips, slacks, pants, shirt, cat, insects, horse, wine, grain, sugar, oil, silk, sand
<b>Past Concern</b>	[+]	The past-tense forms of the verbs contained in the Present Concern dictionary.	
<b>Complexity</b>	[-]	A simple measure of the average number of characters-per-word in a given input file. Borrows Rudolph Flesch's (1951) notion that convoluted phrasings make a text's ideas abstract and its implications unclear.	

**Table 2. CEBE/CEBERO (A LIWC Analysis)**

<b>LIWC dimension</b>	<b>Your data</b>	<b>Personal texts</b>	<b>Formal texts</b>
Self-references (I, me, my)	5.00	11.4	4.2
Social words	6.97	9.5	8.0
Positive emotions	2.51	2.7	2.6
Negative emotions	1.04	2.6	1.6

Overall cognitive words	5.51	7.8	5.4
Articles (a, an, the)	7.40	5.0	7.2
Big words (> 6 letters)	18.35	13.1	19.6

LIWC dimension	Your data	Personal texts	Formal texts
Self-references (I, me, my)	5.35	11.4	4.2
Social words	7.44	9.5	8.0
Positive emotions	2.53	2.7	2.6
Negative emotions	1.06	2.6	1.6
Overall cognitive words	5.79	7.8	5.4
Articles (a, an, the)	7.59	5.0	7.2
Big words (> 6 letters)	17.72	13.1	19.6

## WORKS CITED

- Argamon, Shlomo, Koppel, M., Fine, J., Shimon, A.R. "Gender, Genre, and Writing Style in Formal Written Texts". *Text*, 23: 3 (2003).
- Bates, Marcia J., Shaojun Lu. "An Exploratory Profile of Personal Home Pages: Content, Design, Metaphors". *Online and CD Review*, 21 (1997): 331-340.
- Bauer, Martin W. "Classical Content Analysis: A Review". *Qualitative Researching with Text, Image, and Sound: A Practical Handbook*. Ed. M. W. Bauer & G. Gaskell. London: Sage Publications, 2000.
- Belz Julie, Reinhart, Jonathon. "Aspects of Advanced Foreign Language Proficiency: Internet-mediated German Language Play". *International Journal of Applied Linguistics*, 14: 3 (2004).
- Biber, Douglas. *Variation across Speech and Writing*. Cambridge: Cambridge University Press, 1988.
- Biber, Douglas. *Dimensions of Register Variation. A Cross-Linguistic Comparison*. Cambridge: Cambridge University Press, 1995.
- Biber, Douglas, S. Conrad, and R. Reppen. *Corpus Linguistics: Investigating Language Structure and Use*. Cambridge: Cambridge University Press, 1998.
- Blood Rebecca. "Weblogs: A History and Perspective". *Rebecca's Pocket*. 07 September 2000.  
< [http://www.rebeccablood.net/essays/weblog\\_history.html](http://www.rebeccablood.net/essays/weblog_history.html) >.

- Byrnes, Heidi and Hiram H. Maxim. *Advanced Foreign Language Learning: a Challenge to College Programs*. Boston: Heinle, 2004.
- Cystal, David. *Language and the Internet*. 2001. Cambridge: Cambridge University Press, 2006.
- Delwiche, Aaron "Agenda-Setting, Opinion Leadership, and the World of Weblogs". Paper presented at the Annual Conference of the International Communication Association, New Orleans, LA, 2004.
- Ervin-Tripp, Susan M. "Alternation and Co-Occurrence". In J. J. Gumperz & D. Hymes (Eds), *Directions in Sociolinguistics: The Ethnography of Communication*. New York: Holt, Rinehart and Winston, 1972: 218-250.
- Flesch, Rudolf. *The Art of Clear Thinking*. New York: Harper, 1951.
- Gillmor, Dan. "Moving toward Participatory Journalism". *Nieman Reports*, 57(3), 2003: 79-80. Available from <http://www.nieman.harvard.edu/reports/03-3NRfall/V57N3.pdf>
- Ha, L., E.L.James. "Interactivity Reexamined: A Baseline Analysis of Early Business Web Sites". *Journal of Broadcasting and Electronic Media*, 42 (1998): 457-474.
- Hart, Roderick P. "Redeveloping DICTION: Theoretical Considerations. *Theory, Method, and Practice in Computer Content Analysis*. Ed. Mark West. New York: Ablex, 2001.
- Herring, Susan C. et al. 2004 Herring, S., Scheidt, L. A., Bonus, B., & Wright, E. (2004). "Bridging the Gap: A Genre Analysis of Weblogs". *Proceedings of the Thirty-Seventh Hawai'i International Conference on System Sciences*. Los Alamitos, CA: IEEE Computer Society Press. Retrieved September 2, 2008 from <http://www.blogninja.com/DDGDD04.doc>.
- Herring, Susan C., Scheidt, L.A., Bonus, S., & Wright, E. "Weblogs as a Bridging Genre". *Information, Technology & People*, 18 (2) 2005:142-171.
- Herring, Susan C. "Language and the Internet". In W. Donsbach (Ed.), *International Encyclopedia of Communication*. Blackwell Publishers, 2008:2640-2645. Preprint: <<http://ella.slis.indiana.edu/~herring/lg.inet.pdf>>.
- Herring, Susan C., and J. C. Paolillo. "Gender and Genre Variation in Weblogs". *Journal of Sociolinguistics*, 10(4), 2006: 439-459. Preprint: <http://ella.slis.indiana.edu/~herring/jslx.pdf>.
- Huffaker, David A., S.L. Calvert. "Gender, Identity, and Language Use in Teenage Blogs". *Journal of Computer-Mediated Communication*, 10(2) 2005: <<http://jcmc.indiana.edu/vol10/issue2/huffaker.html>>
- Hymes, Dell H. *Foundations in Sociolinguistics: An Ethnographic Approach*. Philadelphia: University of Pennsylvania Press, 1974.

- Lasica, Joseph D. "Blogging as a Form of Journalism". *USC Annenberg Online Journalism Review*. 18 April 2002. 24 March 2006:  
<http://www.ojr.org/ojr/lasica/1019166956.php>.
- Lasica, Joseph D. "Blogs and Journalism Need Each Other". *Nieman Reports*, 57(3), 2003: 70-73.  
<http://www.nieman.harvard.edu/reports/03-3NRfall/V57N3.pdf>.
- Leaver, Betty L. and B. Shekhtman (eds.). *Developing Professional-Level Language Proficiency*. Cambridge University Press, 2002.
- McMillan, Sally J. "Health Communication and the Internet: Relationships between Interactive Characteristics of the Medium and Site Creators, Content, and Purpose". *Health Communication*, 11 (4) 1999: 375-390.
- Miller, Carolyn R., D. Shepherd. "Blogging as Social Action: A Genre Analysis of the Weblog".
- In L. Gurak, Antonijevic, S., Johnson, L., Ratiff, C., Reyman, J. (Eds), *Into the Blogosphere: Rhetoric, Community and Culture of Weblogs*, 2004:  
[http://blog.lib.umn.edu/blogosphere/blogging\\_as\\_social\\_action\\_a\\_genre\\_analysis\\_of\\_the\\_weblog.html](http://blog.lib.umn.edu/blogosphere/blogging_as_social_action_a_genre_analysis_of_the_weblog.html).
- Pacea, Otilia. "*Think you're Escaping and Run into Yourself: Reshaping Electronic Diaspora*". (Ex)patriation. *A Collection of Essays (The Proceedings of the (ex)Patriation Conference , Constanța, Sept. 20-22, 2007)* 2007. Ed. Adina Ciugureanu et al. Ovidius University Press, Constanța, 2008: 257-266.
- Pacea, Otilia. "*It ain't all that different than real life: Recycling the Cowboy Myth in the Postmodern Electronic Discourse of the Expat Blogs*". *Redefining Modernism and Postmodernism*. Ed. Huber Zapf, Şebnem Toplu. Cambridge Scholars Publishing, Cambridge, , 2010. In press.
- Papacharissi, Zizi. "The Blogger Revolution? Audiences as Media Producers". Paper presented at the Annual Conference of the International Communication Association, New Orleans, May 2004.
- Park, David W. "Bloggers and Warbloggers as Public Intellectuals: Charging the Authoritative Space of the Weblog. Paper presented at Internet Research 4.0, Toronto, Canada, October 2003.
- Pennebaker, James W., R.J. Booth, M.E. Francis. *Linguistic Inquiry and Word Count: LIWC 2007*. Austin, TX: LIWC, 2007 ([www.liwc.net](http://www.liwc.net)).
- Puschmann, Cornelius. *The Corporate Blog as an Emerging Genre of Computer-Mediated Communication: Features, Constraints, Discourse Situation*. Eds. Svenja Hagenhoff, Dieter Hogrefe, Elmar Mittler, Matthias Schumann, Gerald Spindler & Volker Wittke: Göttinger Schriften zur Internetforschung, Universitätsverlag Göttingen, 2010. In press.  
 Preprint:[http://cornelius.ynada.com/publications/Outline\\_Thesis\\_Corporate\\_Blogging.pdf](http://cornelius.ynada.com/publications/Outline_Thesis_Corporate_Blogging.pdf).

Rodzvilla, John (ed.). *We've Got Blog: How Weblogs are Changing Our Culture*. Cambridge MA: Perseus Publishing, 2002.

\*\*\* *Technorati State of the Blogosphere 2008*:

<http://technorati.com/blogging/feature/state-of-the-blogosphere-2008/>