

Book Reviews

discourse.cpp

O.S. le Si (edited by Aurélie Herbelot)

University of Cambridge

Berlin: Peer Press, 2011, 55 pp; paperbound, ISBN 978-3-00-33516-7, €9.00, or free on-line at peerpress.de/discoursecpp.pdf

Reviewed by

Lori Emerson

University of Colorado at Boulder

discourse.cpp is a short collection of computer-generated poetry edited by computational linguist Aurélie Herbelot. The poetry was produced by a program, named O.S. le Si, that was derived from Herbelot's research on context-based ontology extraction from text (Herbelot and Copestake 2006; Herbelot 2009). In this case, Herbelot provided 200,000 pages from Wikipedia for the program to parse and output lists of items whose context is similar to words such as *gender*, *love*, *family*, and *illness*. For example, Herbelot explains that content in the opening piece titled "The Creation" was "selected out of a list of 10,000 entries. Each entry was produced by automatically looking for taxonomic relationships in Wikipedia"; and, for the piece titled "Gender," she chose the "twenty-five best contexts for *man* and *woman* in original order. No further changes" (page 47). The collection is, then, as we are told on the back cover, "about things that people say about things. It was written by a computer."

Poets—or, for the sake of those still attached to the notion of an author who intentionally delivers well-crafted, expressive writing, "so-called poets"—have been experimenting with producing writing with the aid of digital computer algorithms since Max Bense and Theo Lutz first experimented with computer-generated writing in 1959 (Funkhouser 2007). The best-known English-language example is the 1984 collection of poems *The Policeman's Beard is Half-Constructed* by the artificial intelligence program Racter (a collection which, it was later discovered, was heavily edited by Racter creators William Chamberlain and Thomas Etter). *discourse.cpp* is yet another experiment in testing the capabilities of the computer and computer programmer to create not so much carefully and intentionally crafted, rhythmically or musically pleasing verse as *broadly revealing* poetry—poetry that is not meant to be close-read (most often to discover underlying authorial intent) but rather read as a collection of a kind of linguistic evidence. In this case, the collection provides evidence of trends in on-line human language usage which in turn, not surprisingly, provides evidence of certain prevailing cultural norms; for example, we can see quite clearly English-speaking, western culture's continued attachment to heteronormative gender roles in "Gender" (page 18):

| Woman | Man |
|--------------|--------------|
| man love — | — win title |
| — marry man | — love woman |
| — give birth | — claim be |

One of the questions this collection raises is this, however: If the craft of poetry writing becomes more about programming and editing, and if the reader or critic does

not have access to the code or to all of the editorial decisions that created this collection, how, then, do we judge whether the computer-generated writing is successful or not? There's no doubt that the piece "Gender" is provocative and revealing, but is it a "good" poem? Perhaps the collection teaches us that, with the ever-increasing intertwining of human and digital computer that results in the displacement of the human as sole reader-writer now that the computer is also a reader-writer alongside (and often in collaboration with) the human, these sorts of judgments of "good" and "bad" are no longer appropriate. Once the human is no longer the sole creator and no longer in control of the final artistic or literary artifact, then we are left to judge only the process in place of judging the product; with regard to *discourse.cpp*, it is a process we do not have direct access to.

Furthermore, as Herbelot rightly points out in the Editor's Foreword, to a large extent this experimentation with the computer as reader-writer also comes out of early 20th century, avant-garde writing that similarly sought to undermine, if not displace, the individual intending author. Dadaist Tristan Tzara, for instance, infamously wrote "TO MAKE A DADAIST POEM" in 1920 (see Tzara 1924) in which he advocates writing poetry by cutting out words from a newspaper article, randomly choosing these words from a bag, and then appropriating these randomly chosen words to create a poem by "an infinitely original author of charming sensibility." Tzara was, of course, being typically Dadaist in his tongue-in-cheek attitude; but he was also, I believe, serious in his belief that the combination of appropriation and chance-based methods of producing text could produce original writing that simultaneously undermined the egotism of the author. Insofar as *discourse.cpp* comes out of a lineage of experimental writing invested in chance-generated writing and, later, in exploiting computer technology as the latest means by which to produce such writing, it also comes out of a certain tradition of disingenuousness that comes along with this lineage. No matter how much Tzara and later authors of computer-generated writing sought to remove the human-as-author, there was and still is no getting around the fact that humans are deeply involved in the creation process—whether as cutters-and-pasters, programmers, or editors. The collection, then, is a much more complex amalgam than even Herbelot seems willing to acknowledge, as *discourse.cpp* is evidence of the evenly distributed reading and writing that took place between Herbelot and the program itself.

References

- Funkhouser, Christopher. 2007. *Prehistoric Digital Poetry*. University of Alabama Press, Tuscaloosa.
- Herbelot, Aurélie. 2009. Finding word substitutions using a distributional similarity baseline and immediate context overlap. *Proceedings of the Student Session of the 12th Conference of the European Chapter of the Association for Computational Linguistics*, 9 pages, Athens.
- Herbelot, Aurélie and Ann Copestake. 2006. Acquiring ontological relationships from Wikipedia using RMRS. *Proceedings of the Workshop on Web Content Mining with Human Language Technologies*, 10 pages, Athens, Georgia.
- Racter. 1984. *The Policeman's Beard is Half-Constructed: Computer Prose and Poetry*. Warner Software / Warner Books, New York.
- Tzara, Tristan. 1924. TO MAKE A DADAIST POEM. Reprinted in *Seven Dada Manifestos and Lampisteries*, translated by Barbara Wright. Oneworld Classics, Richmond, UK, 2011.

Lori Emerson is an Assistant Professor in the Department of English at the University of Colorado at Boulder. She writes on digital literature, experimental American and Canadian poetry from the 20th and 21st century, history of computing, and media theory. In addition to directing the Archeological Media Lab, she is currently working on several book projects, including a monograph, *Reading Writing Interfaces: From the Bookbound to the Digital*, and *The Johns Hopkins Guide to Digital Media and Textuality*, co-edited with Marie-Laure Ryan and Benjamin Robertson (Johns Hopkins University Press, forthcoming Fall 2013). Emerson's e-mail address is Lori.Emerson@colorado.edu.

