From Instrumental Texts to Textual Instruments

Noah Wardrip-Fruin

Brown University
E-mail: nwf@brown.edu

ABSTRACT: This panel statement begins with brief descriptions of *Screen* and *Talking Cure,* collaborative artworks of a type some have called "instrumental texts." Instrumental texts are said to be *played,* in an analogy to musical instruments. However, the play material is predetermined, making their play more like most computer games. "Textual instruments," on the other hand, can play many compositions (and each composition can produce many different textual outcomes). A first of these instruments, created with Brion Moss and using n-gram algorithms, is described.

KEYWORDS: electronic writing, hypertext, cybertext, electronic art, computer games, play

Language piano is the condition and critique of virtuosity — aria, monologue — standing in for duration, the role of experience, the condition of art, and probably delusion.
— Thalia Field

INTRODUCTION: INSTRUMENTS

In the electronic writing community there has been increasing talk, in the last couple of years, about the idea of "instrumental texts." There is talk of texts meant to be played, and that provide affordances for such play much as folk musical instruments do (e.g, the frets on a guitar). There is talk of texts that provide opportunities for practice and reward mastery. What is practiced and mastered — again, the analogy is drawn with musical instruments — is often presented as a physical discipline. Instrumental texts are also regularly presented as analogous to computer games in these ways. Given that most works presented as examples of instrumental texts always use the same material for their play (always, so to speak, "play the same tune") the analogy with games may be the more accurate of the two. However, the type of engagement that authors hope to produce with instrumental texts may be more musical than game-like. (The few publications in this area include [13,14].)

I have worked on collaborative projects that might be called instrumental texts, and I will briefly describe two (*Screen* and *Talking Cure* — both currently in process). I will also briefly touch on another meaning of *instrument* which may be productive for discussions of electronic writing. But I will devote most of my space here to discussing a set of projects now beginning (Brion Moss and I are currently working on the first of them) that do not fit comfortably with the emerging category of instrumental texts. I will call these "textual instruments." I will describe our first instrument — which extends n-gram algorithms with a history stretching back to Claude Shannon — and also the first composition being written for it.

I view both instrumental texts and textual instruments as falling within the area of "digital instruments or poietic systems" which this panel takes as its subject.

INSTRUMENTAL TEXTS: *SCREEN* & *TALKING CURE*

Screen

Screen is a collaboration with Andrew McClain, Shawn Greenlee, and Josh Carroll which we are creating in the Brown University immersive virtual reality chamber (Cave), as part of a research project in spatial hypertext writing directed by Robert Coover. Brown's VR chamber is similar to the University of Illinois's CAVE — a virtual environment that shows three-dimension images while allowing users to continue to see their own bodies, and that does not require users to wear encumbering equipment (unlike head-mounted displays, which are essentially blindfolds with televisions inside) [10]. Brown's Cave is an eight foot cube, missing its top and one side, and its walls and floor are projection screens. Two projectors are pointed at each screen, and they alternately project images meant for the user's left and right eyes. The user wears shutter glasses that alternately occlude the left and right eyes, in synchronization with the projectors. The result is stereo VR — 3D vision of computer-generated imagery.

The initial experience of *Screen* can be disorienting for those familiar with VR. Rather than make the walls "disappear," we project flat images onto the same plane as the walls, reinforcing their presence. And the images we project are not of colorful shapes, but of white text on a black background. This text tells the story of a person in a room of screens. Once this text has been presented (giving time for it to be read) a word peels from one of the walls and flies toward the reader. If the reader does nothing, the word circles around her. Soon another word peels, and then another, at an increasing pace, flocking around the

Figure 1: Peeled words in *Screen*.

Figure 2: Swirling words in *Screen.*

reader (figure 1). The reader can intervene in this process by batting at words with her hand. When a word is hit it flies back toward a wall. If it is the only word off the wall it will return to the space it left empty. However, if more than one word is off the wall then a hit word may return to a different space. Once the number of words off the walls passes a certain threshold all the remaining words come free of the walls, swirl around the reader, and then collapse into the center of the Cave (figure 2). A final "closing" text is then displayed [4].

Given these structures, the more active a role the reader takes in batting words — the more the reader plays the text — the longer the experience lasts, and the more the text is altered by the reading process. In a sense, while the story's body of text is fixed and all of it is revealed at every reading, the story enacts differently depending on the reader's approach (and, for repeat readers, their level of skill). These structures also result in three different reading experiences for the text — the relatively standard reading of the initial text, the body-involving word-by-word reading of word peeling and batting (which is both visual and auditory), and the interaction-determined text scramble read peripherally during the peeling and batting.

Talking Cure

Talking Cure is a collaboration with Camille Utterback, Clilly Castiglia, and Nathan Wardrip-Fruin. It works the story of Anna O, the patient of Joseph Breuer's who gave to him and Freud the concept of the "talking cure" as well as the word pictures to substantiate it. The reader enters a space with a projection surface at one end and a high-backed chair, facing it, at another. In front of the chair are a video camera and microphone. The video camera's image of the person in the chair is displayed, as text, on the screen (figure 3). This "text mirror" display is formed by reducing the image to three colors, and then using these colors to determine the mixture between three color-coded layers of text (my words, Anna's words, and Breuer/Freud's words — figure 4). Speaking into the microphone creates text that replaces one of these layers (Anna's). What is said into the microphone is also recorded, and becomes part of a sound environment that includes recordings of Breuer's words, Anna's words, and
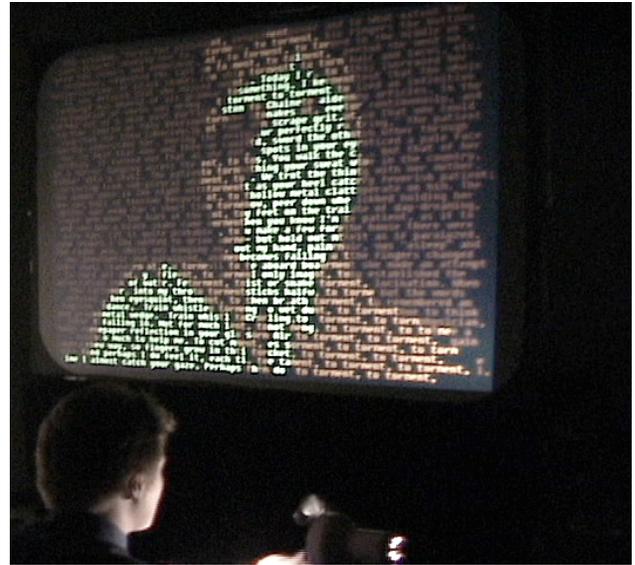
Figure 3: Reader and screen in *Talking Cure.*
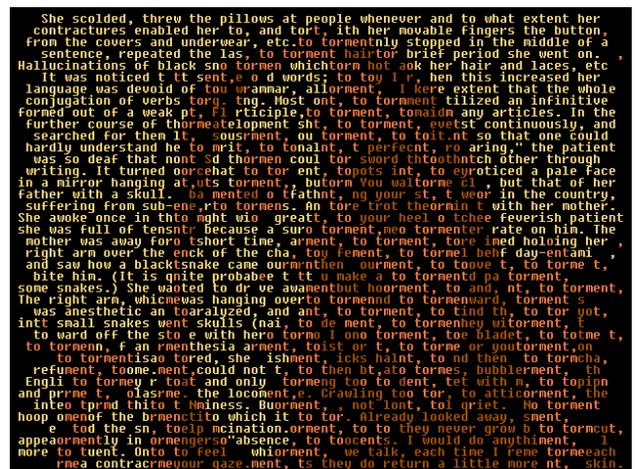
Figure 4: Screenshot of *Talking Cure.*

our words. Others in the space observe the person in the chair through word pictures on the screen.

An initial version of *Talking Cure* was installed at the Electronic Literature Organization's "State of the Arts" Symposium in April 2002. It was interesting to observe the way that visitors at first moved their bodies to achieve visual effects, and then to achieve textual ones, creating new reading experiences for themselves and others in the room (ranging from arm waving for left-to-right reading to head or hand rotation seeking evocative neologisms at the mobile textual borders within the image). While the sound component of the installation experienced problems, this was also "played" while it was available. Some who had developed techniques they liked brought their friends in to show them the effects they were able to achieve.

ANOTHER *INSTRUMENT*

A barometer is an instrument. So is an altimeter. They may have pleasing designs in themselves, but what makes them instruments is the way they sense, incorporate, and display information from outside themselves. This is a relatively common type of instrumentality in the electronic art community (e.g., Stelarc's *Parasite,* or *They Rule*) [12,5]. It is less common in the electronic writing community, though it is important to projects such as *The Impermanence Agent* [15]. This type of instrument, for an electronic writer, requires giving up control of some of one's text to unpredictable outside forces. But the structure remains the author's design.

TEXTUAL INSTRUMENTS

A textual instrument is a tool for textual performance which may be used to play a variety of compositions. In this sense it is evocative of Thalia Field's figure of the "language piano" — something that one learns to play, and which may produce a much wider variety of texts than is the case for those projects normally discussed as instrumental texts.

However, a textual instrument need not be like a prepared piano. The direct selection of text, rather than the manipulation of a non-linguistic device, can be its interface. And the relationship between a textual instrument's interface affordances and the possible textual outcomes need not be one-to-one at all levels (as it must be with a piano's keys, though they may be played in many combinations). Understanding at a gut level how a textual instrument's probability spaces function for a given composition is part of learning to play that piece.

Compositions, here, consist of a body of text (and/or a means of acquiring text) and a set of "tunings" for the instrument(s) used.

All of this can perhaps be made more clear through an example.

An N-Gram Instrument

Brion Moss and I are currently working on a textual instrument that employs a simple n-gram algorithm, related to the algorithms used in *Babble!, Dissociated Press,* and *chan.c* [3,7,6]. The "n" in "n-gram" refers to the number of adjacent words the system pays attention to at any given time. Altering the n-gram length is one way of tuning such an instrument, producing behavior more appropriate for certain types of texts and effects. For our discussion here let us assume a length of three (a "3-gram").

An n-gram algorithm usually adds words to the end of a series, one at a time. It chooses the next word based on the previous words. At an n-gram of three the algorithm would look at the last two words added to the series under construction, and then use these two words as a key when looking for a word to add.

The search for a word to add is carried out in a body of text. A word is a candidate for addition if it appears somewhere in the body of text as part of a 3-gram that starts with the same two words just added. So, if the last two words added to the series were "Dick Cheney" and the body of text contains the 3-grams "Dick Cheney went" and "Dick Cheney threatened" it would be possible for either "went" or "threatened" to be added to the series. If the word "threatened" is chosen, the algorithm will next be able to add words that appear at the end of 3-grams in the body of text that begin with "Cheney threatened" (e.g., "Cheney threatened that" or "Cheney threatened civilian").

Babble! was a shareware DOS program that produced n-gram text interactively (rather than in a batch mode), written by Jim Korenthal. Our project started when Brion, not seeing a Unix n-gram program that satisfied the appetite created by *Babble!,* decided to build his own. I visited, and he showed me what he was working on, and we fell into a conversation that continued over email. The two starting questions for this conversation were, to paraphrase, "How can n-gram babble be *played?*" and "How can n-gram algorithms work within a given document's structure, rather than by adding to a text buffer?"

Of course, the first question may seem a little odd. N-gram babble is already played — in the selection of inputs. At first this process is rather random, throwing texts together and seeing what results they produce. With more experience, one begins to get a sense of what sorts of inputs (and settings like n-gram length or relative weighting of inputs, if the program makes them available for tweaking) might produce interesting results. And after enough play, one can get pretty good at predicting what inputs and settings will produce certain effects. What Brion and I wanted to do was think through ways of extending play beyond the setup process and variable tweaking, and even beyond the ability to "talk back" to in-process n-gramming that *Babble!* allowed. The instrument we're currently creating (an extension of Brion's ongoing work, figure 5) seeks to do this by visually coding words
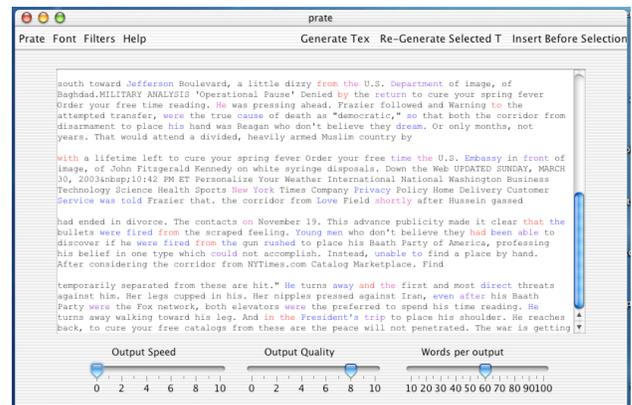

Figure 5: Brion's in-process instrument.

as they are added to the chain. The visual code (for example, a word's color) will be determined by how many possible n-gram completions could have followed it. In one version of the instrument, after an initial n-gram output, the player begins by selecting a word — an action that will lead to the production of a number of possible continuations of the text from that point, among which it will be possible for the player to choose. If the current n-gram length is three, the system will use the clicked word and its predecessor to look for alternative third words. Word adding continues — using algorithms that may be tuned to read repeatedly from the same place (producing some types of greater coherence) or continually jump from place to place — until the instrument builds out the alternative continuations of the text. The player may choose one or select a different word. If an alternative continuation is chosen it replaces the text previously displayed in that area of the output. It is our hope that, as one plays with this interface, one will be able to pass through levels of experience similar to those found with previous n-gram programs in the selection of initial input. Chris Poultney and Rebecca Ross are contributing to the design and development of this interface — to the creation of this playing and reading experience.

We are also addressing the question of how an n-gram algorithm can work within a given document's structure. Our desire to think through how this would work is partially motivated by the fact that, while the text produced by n-gram algorithms has microstructures that are recognizable from its source texts, the larger structures of n-gram texts tend to be very similar regardless of the starting material. Some have tried to address this by looking at larger structures in the source texts statistically, but unless the texts in question have been marked up by a human author or editor, this process involves a series of assumptions about the text (e.g., that a period marks that end of a sentence, as it does not in the case of "e.g.") that are both sometimes inaccurate and on some level aesthetically displeasing. These assumptions are displeasing because they depart from the purity of the simple n-gram algorithm, which in its basic form would work with starting texts in Japanese or Braille or musical notation as easily as English-language ones.

This version of our n-gram instrument, instead, addresses the question of structure by beginning with a starting document rather than a blank text area. The n-gram algorithm is used to alter the document in place, rather than add on to the end of it. This strategy of in-place alteration may sound reminiscent of the strategy of our project *The Impermanence Agent* (a collaboration that also included Adam Chapman and Duane Whitehurst). But in this case the alteration is not driven by the capricious/aleatory decision making of our *Agent*. Instead alteration takes place through the decisions of the instrument's player. The alteration process begins when the player clicks on a word. The selected word (and its predecessors, as determined by n-gram length) are used to

begin searches for alternative paths in the text of an "alteration corpus" (which may not include the text being displayed). Alternative paths are only presented if they are able to "return." To return is to loop back to the current document, following the n-gram procedure, within a reasonable path length (number of words added before return) and within a reasonable number of words from the click point (number of words beyond the word clicked). In both cases, the word "reasonable" in the preceding sentence is defined by the instrument's current tunings. The possibility of return is determined by a method suggested by David Durand — after a certain minimum number of words are added to a possible alternative path, each n-gram search for that path's continuation begins by searching the region of the document that has been defined as the reasonable return area (the string of words between the click point and the last reasonable return word).

Our instrument, with its two variations, is still being built. But we expect it will be like a complex toy, that one can develop a sense for, that one can get better at playing with. We expect it will be like an instrument, that one can learn how to improvise on or play toward goals, for which one can learn the sources, tunings, and means of playing that will lead to different effects.

First Composition
Our first planned composition for the document-altering version of the instrument is tentatively titled *News Reader*. Its initial document will be a current news story from a mainstream U.S. source (e.g., *The New York Times*). Its alteration corpus will be drawn from "alternative" news sources (e.g., *Albawaba Middle East News*). Its settings will be tuned for "reading through." We hope it will be ready for demonstration at DAC.

A curious thing about this composition is that we consider it a piece of electronic writing and yet it contains only found text — and found text that will be different depending on the day of reading (not even hand-selected by us). For some reason this feels natural, and perhaps this is partially due to the "instrument" analogy. For much of the 20th Century composers produced work that arranged and investigated auditory and performance phenomena in ways that did not include the traditional composition of notes. Perhaps textual instruments will be an interesting way (among the others already being pursued) for electronic writing to investigate areas that do not include the traditional composition of text (in even as "traditional" a form as the author-driven cut-up) [1].

At the same time, we are also thinking about a version of *News Reader* that will include text I write.

Related Work
Claude Shannon, who some have called "the Newton of the Information Age," was the first to produce texts using n-grams. In his famous 1948 paper, "A Mathematical Theory of Communication" he included the 2-gram

sentence, "THE HEAD AND IN FRONTAL ATTACK ON AN ENGLISH WRITER THAT THE CHARACTER OF THIS POINT IS THEREFORE ANOTHER METHOD FOR THE LETTERS THAT THE TIME OF WHO EVER TOLD THE PROBLEM FOR AN UNEXPECTED" [11]. Shannon speculated that a 3-gram might be interesting, but because he did his work by hand (opening a book at random and searching visually for matching patterns, one after another) he left off with the observation that "the labor involved becomes enormous at the next stage."

John Cage, Jackson Mac Low, William S. Burroughs, and other artists, on the other hand, pursued somewhat analogous textual production processes and sometimes *did* invest enormous labor in them. Cage wrote, "What can be done with the English language? Use it as material" [2]. He searched manually through texts such as *Finnegans Wake* finding "mesostics" (acrostics that run down the middle) and spent months composing texts using methods connected with the *I Ching*. Later, Cage's processes were computerized, and became much less time-consuming for him. But the processes retained the same basic function — to operate as composition tools (for producing fixed texts), rather than instruments (for playing).

Raymond Queneau and other members of the Oulipo, on the yet another hand, created texts meant to be played — perhaps the first instrumental texts. *One Hundred Thousand Billion Poems* was a combinatory poem, arranged on strips of paper which readers could flip [8]. *Yours for the Telling* was the first tree-structured narrative [9].

Various electronic writers have since used methods in the traditions of Shannon and Cage to produce fixed texts, or methods in the tradition of Queneau to create playable texts. In all cases that we know of, the system and its texts were not separable. Software creators have developed textual toys that could operate on arbitrary texts (e.g., *Dissociated Press*) and have sometimes distributed suggested sets of texts for use with them. But none that we know of have made the reader's play during textual production a primary element (*Babble!* perhaps comes closest). We hope that the confluence of elements from these approaches, in the form of "textual instruments," will prove a fruitful area for exploration.

REFERENCES

1. Burroughs, William S. "The Cut-Up Method of Brion Gysin." A Casebook on the Beat, 105–106. Ed. Thomas Parkinson. New York: Crowell, 1961.

2. Cage, John. "Empty Words." Active Anthology, George and Susan Quasha, eds. 1974.

3. Korenthal, Jim. *Babble! 2.0*. Shareware program distributed with filename "BAB200.ZIP" beginning September 1991.

4. Mirapaul, Matthew. "In Three Dimensions, Words Take Flight. Literally." The New York Times, August 19, 2002. http://www.nytimes.com/2002/08/19/arts/design/19ARTS.htm Access: February 2003.

5. On, Josh and Futurefarmers. *They Rule*. http://www.theyrule.net Access: February 2003.

6. Plotkin, Andrew. *chan.c*. C program, Spring 1991. See "Fun With Markov Chains" http://www.eblong.com/zarf/markov Access: February 2003.

7. Press, Dissociated. Emacs command. In Emacs manual: http://www.gnu.org/manual/emacs-20.7/html_node/emacs_427.html Access: February 2003.

8. Queneau, Raymond. "100,000,000,000,000 Poems," Trans. Stanley Chapman. Oulipo Compendium. Ed. Harry Mathews and Alastair Brotchie. London: Atlas Press, 1998, 15–33. Original French text Cent Mille Milliards de poèmes, Paris: Gallimard, 1961, corrected in One Hundred Million Million Poems, Trans. John Crombie. Paris: Kickshaws, 1983.

9. Queneau, Raymond. Yours for the Telling. Trans. and desgined by John Crombie, graphics by Sheila Bourne. Paris: Kickshaws, 1982. From the French "Un Conte à votre façon." La Littérature potentielle, Paris: Gallimard, 1973.

10. Sandin, Daniel, Thomas DeFanti, and Carolina Cruz-Neira. "A 'Room' with a 'View'" IEEE Spectrum, 30(10):30-33, 39. October 1993.

11. Shannon, Claude. "A Mathematical Theory of Communication." The Bell System Technical Journal, Vol. 27, pp. 379–423, 623–656, July, October, 1948.

12. Stelarc, *Parasite.* (April 1997) http://www.stelarc.va.com.au/parasite/index.htm Access: February 2003.

13. Stephans, Brian Kim. "From Byte to Inscription: An Interview with John Cayley." The Iowa Review Web, February, 2003. http://www.uiowa.edu/~iareview/tirweb/feature/cayley/

14. Wardrip-Fruin, Noah. Interview with Stuart Moulthrop. The Iowa Review Web. Forthcoming.

15. Wardrip-Fruin, Noah and Brion Moss. "The Impermanence Agent: Project and Context." Cybertext Yearbook 2001, Markku Eskelinen and Raine Koskimaa, eds. Publications of the Research Center for Contemporary Culture #72, University of Jyväskylä, Finland. 2002.