DHQ: Digital Humanities Quarterly
2023
Volume 17 Number 2

Computational art Explorations of Linguistic Possibility Spaces: comparative translingual close readings of Daniel C. Howe's *Automatype* and *Radical of the Vertical Heart 亻*

John Cayley <john_cayley_at_brown_dot_edu>, Brown University https://orcid.org/0000-0002-5219-1537

Abstract

A code-critical close reading of two related works by Daniel C. Howe. The artist's *Automatype* is an installation that visualizes and sonifies minimal-distance paths between English words and thus explores a possibility space that is relatively familiar to western readers, not only readers of English but also readers of any language which uses Latin letters to compose the orthographic word-level elements of its writing system [Howe 2012-16a]. In Radical of the Vertical Heart 亻 (RotVH) Howe engages with commensurate explorations in certain possibility spaces of the Chinese writing system and of the language's lexicon. Translinguistically these spaces and, as it were, orthographic architectures, are structured in radically different ways. A comparative close reading of the two works will bring us into productive discursive relationship not only with distinct and code-critically significant programming strategies, but also with under-appreciated comparative linguistic concepts having implications for the theory of writing systems, of text, and of language as such. Throughout, questions concerning the aestheticization of this kind of computational exploration and visualization may also be addressed. His website is programmatology.com.

This essay discusses two distinct but related works by Daniel C. Howe. *Automatype*, 2012-16, is, in the words of the artist's description, "a networked installation (composed of analog TV monitors and RaspberryPIs running custom software) that explores the creation of aesthetic, linguistic meaning via anticipation, juxtaposition, and association. The algorithm at the heart of the work continually computes the minimum number of substitutions required to transform each valid English word into the next, deriving a near infinite number of combinations of words and phrases, letter-by-letter, substitution-by-substitution."[Howe 2012-16a] The algorithm running on each of the (usually nine) displays in an *Automatype* installation is the same, and close reading here effectively focuses on this constituent of the work. *Radical of the Vertical Heart 亻*, 2019-21, the most recent work in what Howe considers a series,

> searches the Chinese lexicon by repeatedly making minimal changes (stroke-by-stroke) to the sub-components of characters, in order to arrive at a new word. Rather than evaluating letters, as in alphabetic machines, this logographic [sic] reader analyses the radicals, components and strokes of characters. When the machine lands on sensitive words, such as those disallowed by China's Great Firewall (or those now illegal in Hong Kong), a break occurs and the machine jumps from traditional [Chinese script forms] to simplified.

[Howe 2019-21a] As a critical feature of *RotVH,* this is aligned with other work by Howe, often engaged with important problems and structures of our predominant digital culture – surveillance, security, power asymmetries, freedom of expression, etc. This essay touches on these matters, particularly insofar as they are addressed by algorithm, data curation, and data structuring, and also comparatively, insofar as they are drawn out by and critiqued in this work's remarkable translinguistic and transcultural contexts. But for more extensive discussion of the digital politics of Howe's work, and its aestheticization, the reader is referred to this other work itself.

Both these works have been presented to their readers and audiences as examples of electronic literature, often sited in necessarily sculptural gallery installation.[1] They are simultaneously examples of computational art and, as such, they are afforded the potential for re-presentation in other forms, particularly distinct graphic and audiovisual manifestations on computer monitors. The code or software of these works is at the core of what they are, artifactually.[2] What the code produces as "display" may be sited or made manifest in any number of 2D or 3D forms and spaces. This discussion, however, concerns itself, chiefly, with the code itself and what is presented by the code as, basically, typographic form on a 2D display, with some accompanying audio: letters and words in the case of *Automatype;* strokes, characters, and two-character compounds in the case of *RotVH*. Audio is used to signal particular relationships between generated forms and items in lexicons that have been derived from the languages of the respective works. 2

The typographic forms generated by these works are of course referred, by those who engage with them, to language. At any one moment, they can often – not always – be read as such, as language; or rather as words *in*, at least, one of two natural languages, English on the one hand, Chinese on the other. Except that the situation immediately becomes more complicated. *Automatype* does have an English lexicon at its disposal but, as it "explores" a mathematically abstracted space of Roman-letteral "spelling," quite apart from "spellings" that are non-lexical in any language, it will encounter and display word-like forms that may be shared by natural languages other than English (any of those whose typography also uses Roman letters), or forms which are not orthographically English although they may be orthographically correct and readable in *another* language. 3

In the case of *RotVH*, the situation is similarly complicated but radically different. Arbitrary or regularly abstracted rearrangements or substitutions of any particular character's strokes or sub-elements will not – or only in extremely rare cases – produce an orthographically readable character in Chinese or a language other than Chinese, but any of the orthographically readable characters which *are* produced may be read, entirely differently, in any number of Chinese dialects or, indeed, in any number of other dialects of certain natural languages such as Japanese or Korean. Since Daniel Howe currently lives, works, and exhibits in Hong Kong, readings of *RotVH,* for example, have often been anticipated in Mandarin (Putonghua) and Cantonese, which are considered mutually unintelligible dialects of Chinese. 4

Raising such complications at the outset of this close reading does at least two things. It highlights the way in which similarly motivated algorithmic processes may generate language-driven computational artworks with entirely different readings, particularly since these "readings," quite apart from being metaphoric – that is, critical or aesthetic – are also readings of distinct "written" linguistic materials which co-constitute the works. And it also reminds us of deep problems concerning the relationship between, on the one hand, what we call "text" or "writing" or "typography," or now also – in the context of computation – "*strings* of characters," and, on the other, language as such. 5

For text is not language unless and until is either actually read, or unless and until it is considered to be, potentially, readable. This is a statement from my own philosophy and, indeed, ontology of language. [3] In the context of my theory I am happy to refer this statement to a species-unique human faculty of language;[4] in the present context I ask only that those reading this essay agree that the words of *Automatype* and the characters of *RotVH* are subject to the possibility, at least, of human reading and interpretation, and that this does have a bearing on their appreciation as art, in particular any art of language that they propose. Their code, however, executes and generates typographic forms without regard to any human reading that may or may not be taking place. From a linguistic perspective, it is the various lexicons and associated data structures embedded in these projects' software –pre-determined and adopted by the artist prior to any execution of the code – which establish relationships with human reading. The processes which generate their displayed linguistic forms do not. 6

They do not, that is, unless, as readers and scholars we believe and assert, amongst other things, that reading and interpretation are reducible to regular, formal, combinatorial processes. One of the important tasks of critical code studies is to articulate this relationship between the code-composed programs ("programs" both literal and figurative) insofar as they are generative of linguistic form, and the practices of reading that we bring to them in order to appreciate what they express as language or, indeed, language art. To be clear, I am not in the camp of those who take language to be computable (reducible to computation) in any sense of this hypothesis that is abstracted or divorced from evolution 7

or, indeed, history. In close reading these two works, I highlight the disjunctures between coded processes and human reading, and will even argue that these disjunctures, articulated, are themselves amongst the works' most significant and affective readings. The disjunctures are brought into relief by the similarity of algorithmic process across these works, in contrast to their radically different integrations with the languages they engage. Both works are, I believe, easy enough to understand in terms of what their code is doing, while their readings are also clearly indicative of quite complex differences of language and culture; and also, in the case of *RotVH*, sociopolitics.

8 Concluding these introductory and anticipatory remarks, it bears mentioning that this relation between code-generated text and human reading has the same structure, fundamentally, as that which obtains for the increasing predominant practices of Natural Language Processing that are driven by Deep Learning. Except that, when Deep Learning is operative, human readers' ability to articulate and understand the coded processes are – is it right to say "literally"? – redacted. Deep Learning language models' "encoding" of "representations" is often presented as a "black box," received as such even by willing experts. And while, in principle, exhaustive formal analysis may be held to be possible, what is actually concealed from us, as human readers and consumers (sic) of the generated pseudo-language, is a number of unsubstantiated assumptions concerning things that we do not (and perhaps cannot) know, scientifically or otherwise. [5] I would summarize this by saying that, although we speak and read, we still do not know what language is, nor how or why some of it may be aesthetically significant for us. As to this how and why, when we consider *Automatype* and *RotVH*, at least we have a chance.

9 The artist documentation pages for these works offer links to "Project Home Pages," Online Versions, and various public manifestations. The Project Home Page for *Automatype* is minimal, and features an evaluative description extracted from a review by Brian Kim Stefans [Stefans 2011]. The Home for *RotVH* however takes us to a page which links the two works explicitly, categorizing them both as "Atomic Language Machines" (ALMs), defined as "discrete computational entities with the potential to change the direction, intent, or magnitude of a literary vector. In general, ALMs can be defined as members of the simplest class of mechanisms able to realize linguistic advantage." [Howe 2021] "Linguistic advantage" is one of those terms deployed in the context of computational linguistics to provide or explain motivation in Natural Language Processing, something used to test against a generated linguistic form in order to decide if it has, in some sense, advantageously "succeeded" or achieved a "goal" which is usually interpretable by (and "advantageous" to) humans. For both *Automatype* and *RotVH* it is advantageous to spell out an item from the lexicons of English and Chinese, respectively.

10 The "atomism" of ALMs is referred by Howe to a concept of "simplest class of mechanism" and is linked with other such classes of mechanism which are cited as deployed, for example, in *The Readers Project* [Cayley and Howe 2009].[6] Here, I read "atomic" as indicative of, as it were, an *elemental* similarity between the class of mechanism driving *Automatype* and that which animates *RotVH*. The ALMs page quotes the following code snippet abstracted from the actual code driving *RotVH*:

```
/*
 * A new minimum edit distance for logographic (Chinese) characters.
 * 1) Decompose each character of each word into subcharacter parts
 * 2) Compute the Levenstein distance for the pair of decomposed strings
 * 3) Add 'cost' component, the # of characters modified - 1
 */
int logographicMinEditDistance (wordA, wordB){

  int cost = Math.max (0, this.levenshtein.get(wordA, wordB) - 1);

  string dcompA = subCharacterDecomposition(wordA);
  string dcompB = subCharacterDecomposition(wordB);

  int total = this.levenshtein.get(dcompA, decompB);

  return total + cost;
}
```

Figure 1. The function logographicMinEditDistance: source code and comment.

11 "Levenshtein distance is a string metric for measuring the difference between two sequences. Informally, the Levenshtein distance between two words is the minimum number of single-character edits (insertions, deletions or substitutions) required to change one word into the other." [Wikipedia contributors 2021] This metric is the measure of "linguistic advantage" that underlies both *Automatype* and *RotVH*. When we consider *Automatype* – working with English words spelt out in Roman letters – this distance is a relatively straightforward concept to grasp. Typographic words are, indeed, sequences or "strings" of characters. For each word in a lexicon (or its derivatives) we can calculate a Levenshtein distance to any other word. Having done so, we can start with any word we like and, having found its "nearest" Levenshteinian lexical neighbor, we may animate a typographic display which performs the minimum number of edits on our original string of characters in order to transform it into its neighbor. This is precisely what *Automatype* does, and by disallowing any "turning back" (to a previously displayed lexical item) it effectively represents an ALM that would travel least-distance paths from item to item until it had visited and exhausted all the items of its English lexicon. [7]

12 Another connotation of "atomic," perhaps only latent in Howe's ALMs webpage, comes into play at this point. The strings and sequences over which Levenshtein distances are calculated must be composed of integral, indivisible – at least for the purposes and processing of the algorithm – tokens, the "atoms" of the symbolic system of which they are expressions. The atoms of *Automatype*'s sequences are letters taken from a familiar, quite clearly delineated set, one that is widely shared over a number of linguistically integrated domains. In linguistic reality, the Roman letters used for English orthography are problematically related to what they purport to represent, the constituents of spoken English "sound-images" (as Saussure, for one, would have it). But leaving these problems of reference, representation, and transcription to one side, "English" letters can be read straightforwardly by humans as from a finite set of integral elements. And although minimally semantic, they are included as lexical items in most dictionaries. Letters are, generally speaking, typographic atoms of reading and are amenable to contemporary computation as such.

13 But what are the correspondent "atoms" of Chinese typography, of the Chinese writing system? Lexicons in the Chinese cultural sphere typically have characters as the "head words" of their articles although modern dictionaries may also use "compounds" of characters, usually consisting of two characters. In modern Chinese these compounds correspond, linguistically, to the words of English dictionaries. Nonetheless, the Chinese character is, culturally speaking, the "atom" of reading for the languages which use this logographic or, more properly, morphographic system of inscription. *RotVH* works with two-character compounds. If characters were *RotVH*'s atoms then Levenshtein distance would be calculated on the basis of *character* insertions, deletions, and substitutions. Since single characters are always morphemes and may often correspond with English words in modern Chinese (while typically doing so in pre-modern Chinese) they are clearly of a different linguistic order as compared to Roman letters. Morpheme insertion, deletion, and substitution would read – if it did motivate the algorithms of *RotVH* – as incommensurate with the corresponding operations of *Automatype*,

although, in this speculative condition, *RotVH* would still exhaustively traverse a lexicon consisting of two-character compounds and would still sometimes display non-lexical items. (And it would still also be able to signal and respond to any of its encounters with politically charged two-character terms). But it would never be non-semantic. Its operations would always yield readable (and perhaps occasionally poetic) meanings.

14 The artist recognized these circumstances and went deeper into the analysis of Chinese characters, coming up with a remarkable and effective compromise for his aesthetic purposes along, implicitly, with novel proposals for conceptualizing and calculating Levenshtein distances across the domain of Chinese characters. This is Howe's comment on his "logographic distance" calculation:

15

```
/*
 * logographic distance
 * - number of different full characters less 1 (via basic Levenshtein)
 * - plus Levenshtein distance between two decompositions for each char
 * from [ ⿰ ⿱ ⿲ ⿳ ⿴ ⿵ ⿶ ⿷ ⿸ ⿹ ⿺ ⿻ ]
 * - this gives an integer distance >= 0 (with 0 for identical strings)
 * - the floating point component (0 <= f < 1) is added by comparing the
number
 of strokes in differing sub-parts normalized against a max-stroke count
 (not used in production)
 */
```

16 The distance calculation is a sum of staged Levenshtein distances. A first distance [1] is calculated between strings of integral characters (the length of these strings being always equal to two in the lexical domain that is addressed by *RotVH*). Then [2] each of the characters in the strings being compared is assigned to one of twelve patterns of (de)composition which are typical of Chinese characters. There is no generally accepted, rigorous analysis of this feature of characters, but sub-elements of characters are traditionally recognized and read into character composition and patterns of disposition for these elements fall into one of the twelve such that determinations are made and may even be assigned in some character dictionaries. Interestingly the patterns quoted above are represented by Unicode glyphs and thus – although these glyphs' reference is much further divorced from linguistic significance, in the sense of any sound-image denotation, than that of alphabetic letters – their implicit deployment in Howe's calculation resonates with the letters-as-tokens approach that is assumed in standard Levenshtein calculations over alphabetic orthographic typography. Finally, [3] calligraphic (or sinotypographic) strokes are also recognized as finest-grained elements of characters. Ordered strokes are what compose the higher-level sub-elements of the (de)composition patterns. Howe documents the possibility of using stroke-token sequences for each of the decomposed sub-elements and adding these to calculate a correspondingly finer-grained Levenshtein distance between, for example characters sharing the same (de)composition pattern. He proposes to do this in an implicitly weighted manner by adding these distances as a "floating point" (fractional) component. In practice, however, in the actual "production" version of *RotVH*, Howe decided to ignore this component.[8]

17 *RotVH* is art not science. It may well, however, engage more practically, empirically, and experimentally with scientific analyses or formulations of editing distance over the Chinese lexicon than other efforts which have been put forward in the context of science explicitly. The choice of next operation – here of insertion, deletion, or substitution – is on display to artistically motivated readers and viewers and will thus bring in other aesthetically implicated considerations apart from Levenshtein distance, the most obvious being based on keeping a "history" of all words or character compounds so far displayed, and disallowing repetitions. This principle applies to both *Automatype* and *RotVH* and is what enables these ALMs to take their shortest possible – conceptual-art – walk through an entire lexicon. For *RotVH*, Howe needs to establish additional pragmatic-aesthetic criteria, beyond the shared concept.

> [O]ther criteria also affect selection (from the larger pool of candidates) including whether they differ on the same "side" (left or right character in the word) as the last few changes, so as to avoid the

same character remaining constant for long periods, and whether any are trigger words. In the current version I ended up ignoring the floating point part of the distance metric… in order to get a set of words essentially tied in distance, so that I could then choose between them according to these other metrics.

[9] The “larger pool of candidates” refers to all those compounds for potential display which are “tied” in terms of Levenshtein distance after stages [1] and [2] above.

18 This is the point at which – without abandoning its inevitably computational and code-driven conceptual aesthetics, both also addressing the domain of Natural Language Processing, and both shared with *Automatype* – *RotVH* shifts its engines of motivation in order to adopt a *critical art* aesthetic, one that is designed to inform its readers, creatively, concerning the sociopolitical valences of certain items in the Chinese lexicon. *RotVH* contains a data file of “trigger” words, those compounds which are flagged by the “Great Firewall,” by Chinese state surveillance of linguistic internet traffic, or which have been signaled as politically taboo in what is now China’s Hong Kong Special Administrative Region. Here is a snapshot from the JSON file (triggers.json) to which *RotVH* refers [Howe 2019-21b]:

```
"滴蜡": { "lang": "simp", "pair": "滴蠟", "def": "using candles for BDSM" },
"滴蠟": { "lang": "trad", "pair": "滴蜡", "def": "using candles for BDSM" },
"汪洋": { "lang": "both", "pair": "汪洋", "def": "vast ocean, PRC ex-VP Wang
Yang" },
"妇联": { "lang": "simp", "pair": "婦聯", "def": "women's league" },
"婦聯": { "lang": "trad", "pair": "妇联", "def": "women's league" },
"罢工": { "lang": "simp", "pair": "罷工", "def": "workers' strike" },
"罷工": { "lang": "trad", "pair": "罢工", "def": "workers' strike" },
"元朗": { "lang": "both", "pair": "元朗", "def": "Yuen Long district, Hong Kong"
},
"陸肆": { "lang": "trad", "pair": "陆肆", "def": "ref. to Tiananmen
Anniversary"},
"陆肆": { "lang": "simp", "pair": "陸肆", "def": "ref. to Tiananmen
Anniversary"},
"學潮": { "lang": "trad", "pair": "学潮", "def": "student movement"},
"学潮": { "lang": "simp", "pair": "學潮", "def": "student movement"},
"八九": { "lang": "both", "def":  "1989, year of Tiananmen Sq massacre"},
"河殤": { "lang": "trad", "pair": "河殇", "def": "River Elegy"},
"河殇": { "lang": "simp", "pair": "河殤", "def": "River Elegy"},
```

19 This is a snapshot which demonstrates the wide-ranging scope of Chinese lexical surveillance, flagging terms with regard to: sexual practices; “reforming” politicians; gender- and class-based affiliations; the Tian’anmen Square massacre and its “student movement” (likely including student movements in themselves); and even a controversial television series. The “lang” property of each two-character item is indicative of one of two main Chinese systems of inscription. Those words with the same triggering “def” property may occur in “trad[itional]” (more strokes, greater complexity) or “simp[lified]” spellings. Traditional characters are still widely and officially used in regions of the Chinese culture sphere – notably and with political significance Taiwan and Hong Kong – which are, to whatever extent, still “outside” the People’s Republic of China. The PRC, on the other hand, has instituted and adopted its own “reformed” and “simplified” character orthography. Some spellings – untouched by reform – belong to “both” systems. There is an underlying ideality to characters (or compounds) in either spelling. Essentially, they refer to the same “form-as-read” in the Chinese of their speaker-readers (I would call this a “gram” of the implicit archi-writing) and thus they are equally “triggering” for state surveillance.

20 Whereas the audiovisual behaviors of *Automatype* – apart from those which represent edit operations – are confined to signaling the ALM’s “arrival” at a lexical word, those of *RotVH* signal not only lexical arrival but also whether or not the newly displayed word is a “trigger,” a surveilled word. Then also, after a distinct sound and a flash of red, *RotVH* is also

triggered to switch to orthographic explorations in whatever is the other system of Chinese spelling, either “trad[itional]” or “simp[lified],” depending on which of the two it was exploring when disrupted by a “trigger.” This shift of orthographic systems might perhaps be interpreted as a futile attempt to “misspell” and thus elude surveillance, but it can never be more than a jolt to either system since the underlying “trigger” is, as we have seen, the same word in any Chinese that matters to the surveillance operations of its PRC state overseers. Regardless of the overseers’ indifference (or rather their fixation on transgression) or the computational indifference of *RotVH*’s execution, the change of orthographic systems will, nonetheless, resonate with, and may mildly traumatize, Chinese readers since each system has ideological and political alignments and associations.[10]

21 The typical installation version of *Automatype* consisted of a number of networked, otherwise independently operating instances of the ALM in, for example, a 3x3 grid, each exploring the same lexicon separately but on its own path from a different starting point, making minimal displayed edits, and arriving intermittently at actual lexical items, English words. This grid arrangement overlays a form suggestive of visual or pattern poetry and thus also an aesthetic, a *poetic* that is not programmatically related (not integrated by code) with the *conceptual* Natural Language Processing aesthetic of the ALM itself. *Automatype*’s poetic overlay-in-installation has not been a significant focus of attention for this essay although it is what allowed Brian Kim Stefans, in his remarks on *Automatype* to say that “You will spend either 10 seconds or 5 minutes staring at this thing [the grid]; you will also see either a bunch of random words, or occasionally, if not always, engaging samples of minimalist poetry” [Stefans 2011].

22 Not only are *Automatype* and *RotVH* closely related in that they share essentially the same coded, programmatic procedure for “walking through” a lexicon and thus essentially the same computational and conceptual NLP aesthetic; they are also similarly structured in that both have been given, by their maker, an additional, overlaid aesthetic. And although the code of *RotVH* refers and reacts to its trigger word data, this is an additional and distinct coded operation of *RotVH*, dependent on an additional human-compiled data file, only of significance to or affective of this ALM’s readers for reasons that are sociopolitical rather than merely linguistic. The immediately following concluding remark should be part of a much larger discussion, but we might begin to take away something beyond the code-critical from this comparative reading by reflecting on how the overlaid aesthetic for a project within the domain of global English tends to engage with formal arrangement and poetics, whereas the closely related, subsequent project, addressing what is now perhaps the planet’s “other” global language, engages sociopolitics and critical art practices.

23 There is at least one more important general point to make that is based on the descriptions and analysis that we have just undertaken, and it has transcultural critical resonance. From both a code critical perspective and from one attuned to careful, responsible humanistic readings of computational art, any writing on this art must remain or become more critically aware of the culture of computation and its history. This is an imperative within the sphere of what is the globally predominant regime of computation, and it has hardly been addressed, as it must be, within an overarching context that is *trans*cultural at the level of distinct “civilizations.”[11]

24 We have seen that, addressing the Chinese system of inscription, an operational analysis of its elements must be done by way of bespoke or imported data structures and bespoke or imported functions – even when this is for the purposes of animating the same Natural Language Processing concept or operation – here, a Levenshtein distance-based “shortest walk” through a lexicon. By contrast, for global English in particular, and for “western” languages having integrated alphabetic systems of inscription, the data structures and functions are already more or less built into actually existing computational infrastructure. The historical reasons for this integration of computation and the alphabet are quite well known. But this is no reason for critical complacency, particularly when we recall that text-as-orthography is in no way consistent with even a linguistic-scientific analysis of language as such.[12] We call orthographic spellings “words,” but this is both pragmatic, living-culture convention and scientific misdirection. If historical and contingent computational infrastructures reinforce our misdirected conventions, this has implications far beyond the misapplication or bespoke adaptation of these infrastructures to systems of inscription for which they are ill-adapted. Close reading of *Automatype* and *RotVH* allows us to encounter and explore these contemporary transcultural critical aporias. These two ALMs both read “words” but read them differently, because even the spellings of these words are culturally situated and involve radically different relations to linguistic structure. The ALMs nonetheless deploy algorithms for lexical traversal

which are essentially the same. The code is radically different in each case, not only due to the cultural situation of the "words," but also because the code itself – and contemporary computation as whole – is culturally situated and adapted to particular global locations.

Notes

[1] Daniel C. Howe, *Automatype,* 2012-16; *Radical of the Vertical Heart* 忄, 2019-21. Both works are quite extensively documented on the artist's website with landing pages for each as follows: *Automatype*, https://rednoise.org/daniel/automatype; and *Radical of the Vertical Heart* 忄 (henceforth *RotVH*), https://rednoise.org/daniel/radicaloftheverticalheart (both accessed February 27, 2022). Live behavior of the basic, non-installation, versions of these works are linked from these pages and the reader is invited to consult these online. There is an entry for *Automatype* in the *ELMCIP Knowledgebase* at https://elmcip.net/node/4001, and *RotVH* (is included in the forthcoming projected for 2022) *Electronic Literature Collection 4*, see https://collection.eliterature.org. I have provided separate bibliographic entries for the code repositories for these two works, publicly available on GitHub, *Automatype [Code]*; *Radical of the Vertical Heart* 忄*[Code]*.

[2] This essay is not so much concerned with actually quoted code or code "style." The code is addressed, critically, with respect to the underlying algorithms which are, for the most part, "paraphrased" from the code itself, or re-expressed in what might be more accurately referred to as pseudocode. Nonetheless the code repositories have not only been reviewed by the author; he has also "cloned" them for examination and even potential future development and collaboration. (An interesting reconfiguration of the usual literary or art critical relations.)

[3] Cf. [Cayley 2018b]; [Cayley 2018a]; [Cayley 2020]. And also, https://nllf.net.

[4] Members of this species, humans, bring language into being by reading (grasping, understanding) language's material traces. These traces are expressed and appreciated chiefly in aurality but also in culturally integrated graphic visuality in writing; or in commensurate modes of language practice (in any perceptible material). The faculty that allows this to happen is necessarily constitutive of language, and co-constitutive of "the human." As such it relies on the evolved and world-integrated embodiment of human animals, such that the "differences" of its integral symbolic system are never abstract. They are better understood as Derridean "differances," as sayable meanings that emerge as they are expressed – with the kind of ambiguity or polysemy, for example, that embodied animation entails, and as neither pre-determined (from a finite set) nor subject to regular formulation as they are uttered and read. Thus, the constitutive "faculty" of which I speak is in sharp contradistinction with the predominant scientistic understanding of language, too-often based on a hypothetical Chomskyean, designative, pre-engineered difference machine, built-into the "brain" that runs combinatorial transformations on the abstracted "deep structures" of inherently computable and "disembodiable" mind/language constructs.

[5] Readers may find a recent "annual review" useful in this context: Ellie Pavlick, "Semantic Structure in Deep Learning" [Pavlick 2022]. Pavlick is not, I think, amongst those researchers who are content to "black box" Deep Learning however opaque or beyond-human-scale/perception the motivation of language models may appear to be; and certain aspects of her problematization of the relationship between "model representation" and "model behavior" resonate with arguments in the present essay.

[6] A collaboration with the author. It is perhaps worth noting that, on the ALMs page, the artist allows himself a certain poetic license, with language that is a blend of technical and fanciful, somewhat mythologizing, for example, the "early theoretical research associated with the 'Natural Language Liberation Front'https://nllf.net"

[7] As such, the piece also qualifies as a work of computationally animated *conceptual* literary art, a "walk through the dictionary" not unlike Allison Parrish's @*everyword*, published by bot on Twitter from 2007 through 2014 and documented in this printable publication: [Parrish 2015]. Parrish's work generates much of its significance and affect from the socially mediated situation and circumstances of its publication, and thus also from the transactional interventions and engagements of its readers through Twitter affordances (such as retweeting, liking etc.) Nonetheless, although the lexicons used for *Automatype* and @*everyword* were different, there is clearly a profound underlying "sameness" – a shared ideality if not an identity – to the "texts" of these two works. And either could have used the other's lexicon without impugning their respective concepts or motivating principles. Moreover, it is similarly the case that the text's generation by code, in @*everyword*, has little or no bearing on its human reading, although the relationship between code and reading is radically different from that which we are exploring here, with respect to *Automatype* and *RotVH*. It seems that there is an opportunity, which we cannot take up at present, for another interesting comparative close reading with code critical implications.

[8] There are hints that the artist will address aspects of what follows in future versions, but it should be noted that *RotVH* does not make use of actual or potential "vocabularies" of character sub-elements. This is a much-contested aspect of research on the Chinese script and its abstracted analytic properties. Whereas some sets of sub-elements are well-established, particularly the so-called "radicals" or "classifiers"

which are used in many Chinese dictionaries (and are referred to in the title of *RotVH*), even the standard 214 radicals of the canonical Kangxi dictionary have multiple forms (of ordered strokes, etc.) and there are problems in exhaustively assigning characters to some particular "radical." Nonetheless a finer-grain Levenshtein distance between characters might be able to take account of sub-element sets, leading to, for example, a Levenshtein distancing such as this: stage [1] based on strings of integral characters; stage [2] assignments of the constituent characters to one of the 12 (de)composition patterns; stage [3] tokenized identification of constituent sub-elements from a comprehensive set. Then [4, was 3 in the text] Levenshtein distances between string representations of the sub-elements' ordered constituent strokes. But it is a significant graphic problem to animate, in any "readable" fashion, the operations of stroke insertion, deletion, and substitution that might be called for or desired. This brings up a related fact concerning the animated display of *RotVH's* current version. The part of its animation during which a sub-element is built up is *not* a representation of Levenshtein edit operations. This is simply an animation of the stroke-ordered process that a human calligrapher would, by traditional convention, use to write the character's sub-element. In the current version of *RotVH* only the graphic removal of a sub-element and its subsequent animated rendering *as a whole* is related to Levenshtein distance, but in the current version it often represents a Levenshtein-distance "tie" since the before and after characters may have the same (de)composition value. Otherwise, it counts as a substitution operation.

[9] Personal communication with the artist, November 30, 2021.

[10] Cf. these interesting comments from the artist,

> There is also the idea of an (externally-imposed)(catastrophic?) break here… Yes, there are words in common, especially as written characters, to both but one (human or non-human reader) is either reading in C[antonese, typically integrated with traditional characters] or M[andarin, official Putonghua of the PRC and aligned with simplified characters], it would seem to me, though my students sometimes try to create politically-inflected mergers of the two. But historically-speaking some threshold of *pressure* (a favorite word here) may be required before one willingly switches (as example, the political implications of choosing to speak — or not to speak — M[andarin] as a Hong Konger are quite clear). And while generally the case, close relationships between the political and the linguistic are perhaps more explicit/tangible in Chinese history, especially regarding government policies of the last century that attempt to manage language as a resource (see Mullaney, etc.)

(Personal communication, December 4, 2021) The Mullaney here is [Mullaney 2017].

[11] The reference to "civilizations" is historical and simply intended to signal the historical distinction between Eurocentric "western" civilization and that of a sinocentric "east." The artist's comments in note 15 above has already cited one of the few studies addressing these transcultural issues with respect to language practice and computation. [Mullaney 2017] Mullaney's book explicitly presages what promises to be his essential follow-up work on the history of computation in the Chinese culture sphere.

[12] Which has implications, please recall, for Deep Learning language models and generators.

Works Cited

Cayley 2018a Cayley, John. "Grammalepsy: An Introduction", *Electronic Book Review*, 08-05, 2018, http://electronicbookreview.com/essay/grammalepsy-an-introduction. Accessed August 12, 2018.

Cayley 2018b Cayley, John. *Grammalepsy: Essays on Digital Language Art*. Electronic Literature. New York and London: Bloomsbury Academic, 2018, https://doi.org/10.5040/9781501335792.

Cayley 2020 Cayley, John. "The Language That Machines Read". In *Attention à la Marche = Mind the Gap: Penser la Littérature Électronique in a Digital Culture = Thinking Electronic Literature in a Digital Culture*, edited by Bertrand Gervais and Sophie Marcotte, pp. 105-113. Montreal: Les Presses de l'Écureuil, 2020.

Cayley and Howe 2009 Cayley, John, and Daniel C. Howe. *The Readers Project*. 2009–. http://thereadersproject.org. Accessed November 28, 2021. [Also published/documented: *Electronic Literature Collection*, vol. 3, 2016, http://collection.eliterature.org/3/work.html?work=the-readers-project. Accessed August 17, 2017, ELMCIP Knowledge Base https://elmcip.net/node/864. Accessed August 14, 2017.]

Howe 2012-16a Howe, Daniel C. *Automatype*. 2012-16. https://rednoise.org/daniel/automatype. Accessed November 27, 2021.

Howe 2012-16b Howe, Daniel C. *Automatype [Code]*, 2012-16. https://github.com/dhowe/Automatype.

Howe 2019-21a Howe, Daniel C. *Radical of the Vertical Heart* 亻. 2019-21. https://rednoise.org/daniel/radicaloftheverticalheart. Accessed November 27, 2021.

Howe 2019-21b Howe, Daniel C. *Radical of the Vertical Heart* 亻 [Code], 2019-21. https://github.com/dhowe/rotvh.

Howe 2021 Howe, Daniel. C. (2021) “Atomic Language Machines,” *rednoise.org*. {Avilable at: https://rednoise.org/daniel/pages/alms/index.html(Accessed 28 November 2021).}

Mullaney 2017 Mullaney, Thomas S. *The Chinese Typewriter: A History*. Studies of the Weatherhead East Asian Institute, Columbia University. Cambridge: The MIT Press, 2017.

Parrish 2015 Parrish, Allison. *@everyword: The Book*. New York: Instar Books, 2015.

Pavlick 2022 Pavlick, Ellie. (2022) “Semantic Structure in Deep Learning,” *Annual Review of Linguistics*, 8:23, pp. 447-471. First published as a Review in Advance on November 23, 2021. {Available at: https://doi.org/10.1146/annurev-linguistics-031120-122924}

Stefans 2011 Stefans, Brian Kim. “Third Hand Plays: ‘Automatype’ by Daniel C. Howe”, *Openspace, SFMoMA* September 8, 2011) https://openspace.sfmoma.org/2011/09/third-hand-plays-automatype-by-daniel-c-howe. Accessed December 4, 2021.

Wikipedia contributors 2021 Wikipedia contributors. (2021) “Levenshtein distance,” *Wikipedia, The Free Encyclopedia*. {Available at: https://en.wikipedia.org/wiki/Levenshtein_distance (Accessed 28 November 2021).}

This work is licensed under a Creative Commons Attribution-NoDerivatives 4.0 International License.