

A Strange Metapaper on Computing Natural Language

by [Manuel Portela](#), [Ana Marques da Silva](#)

09-02-2018

[PDF](#)

<https://doi.org/10.7273/6gws-hf93>

This essay was peer-reviewed.



Without anonymous peer review, there can be no formal recognition of literary scholarship, and ebr is no exception. That said, our journal looks for occasions to turn our confidential reports into public riPOSTes, if the reviewer is so inclined. In this essay, our colleagues from Coimbra, Manuel Portela and Ana Marques da Silva, stage reflections on the peer reviews that their own scholarly work has generated, in earlier submissions to other peer review outlets. The "metapaper" that results, is a further step in the initiative not to do away with peer review, but to bring the process into the public sphere.

Abstract

This is a paper about writing a paper about computational creativity in natural language generation. The first part contains the second-order paper, i.e., a general explanation about the first-order paper or embedded paper, which constitutes the second part. [This embedded paper](#), by the same authors, contains its own abstract, keywords, and reference list. It is titled "If then or else: Who for whom about what in which." Three actual peer reviews of that embedded paper have been integrated into the framework of the second-order paper as an attempt to illustrate the discursive and pragmatic conditions of the communicational situation of the first-order paper. This framing of one text inside another is intended to highlight the form of the paper as a specific writing constraint while using it as a self-exemplary instance of the difficulties and limitations of computing natural language. The whole metapaper is intended as a writing experiment on self-description and on linguistic creativity. Or is it just a joke?

Keywords: reflexivity; parody; writing under constraint; natural language processing.

Metainroduction

We will start by explaining (Section A) the context for our *sui generis* approach to computational creativity in natural language generation as exemplified in the embedded metapaper below. Then we will analyze our own embedded paper as (Section B) a procedural generative non-computational

form of writing which contains a philosophical reflection (Section C) about the conditions for the emergence of textual form and textual interpretability, and about current practices of natural language automation based on computational generative works. Finally, we will call attention to (Section D) our own embedded metaper as evidence of both the challenges of modelling natural language through computational generativity, and the political and social implications of the ongoing natural language automation. The distinction between embedded paper and framing paper breaks down when, in a final double coda (Section E), we discuss the discursive conditions that define the academic paper as a particular textual constraint. We suggest that readers jump ahead to the embedded paper at this point. Cf. below, ["If then or else: Who for whom about what in which."](#)

A. Context

As literary scholars, we have been reading programmed generative works for several years with the aim of understanding the poetics of literary production involving natural language generation (Portela 2013, 2017; Marques da Silva 2016, 2017). Our research has been focused on a literary and cultural reading of Natural Language Generation (NLG) rather than a strictly linguistic and computational perspective (which is the main focus of the research papers presented in this workshop). This paper was originally presented at the "INLG 2017 Workshop on Computational Creativity in Natural Language Generation", September 4, 2017, School of Engineering at the University of Santiago de Compostela). Although we recognize and greatly benefit from the contributions of engineering approaches (Gervás 2017; Manjavacas et al. 2017), we want to bring to this discussion some fundamental theoretical questions about language and automation. We are grateful to the organizers of the workshop for this opportunity for submitting our ideas to cross-disciplinary examination and critique. We admit beforehand that our paper may be even more absurd than it sounds. We suspect that it is not computable, even in its parodic elements.

The second aspect for sketching the context from which we are approaching the workshop topic is the fact that we have been focused on corpora of generative works that offer critical insights about ongoing processes of automating natural language production in various human practices, from literary creativity to everyday interactions with digital devices and systems. Such works are interesting not primarily for producing meaningful and original texts (which they do) but for reflecting on their conditions of production. Thus the literary works chosen for analysis are studied as examples of NLG works that can be illuminating about generative poetics, but also as probes into the nature of automation of natural language, which, in its turn, can be seen as just a particular domain in the current accelerated process of softwarization of human culture, in particular of communication media (Manovich 2013).

The question that underlies our embedded metaper is this: what are the conditions for textual interpretability? In other words: how does a textual form emerge? In yet other words: what is the relation between known features of natural language (such as generativity) and the emergence of textual form as an interpretable verbal action? We have no answers for these questions, but we have attempted to make a textual experiment whose result is the paper itself (instead of any formalized textual generative system). Our paper is thus a self-exemplary instance of the conditions required for the emergence of interpretability in written uses of natural language. This is the third element required for explaining the *sui generis* context of our paper.

B. A procedural generative non-computational form

The procedural method used for writing "If then or else: Who for whom about what in which" allowed us to identify three interactional layers required for the production of fully interpretable textual forms, which we have named as "textual text", "meta-textual text" and "networked text". In order to become interpretable, textual forms have to somehow articulate those three dimensions: an assemblage of first-level textual signs (a string of well-formed discourse) depends on explicit or implicit signs that frame their interpretation at a higher level (as a particular genre, for instance), and also on explicit or implicit references to other texts. Texts mediate themselves through both these meta- and network-levels of reference.

Those conditions for interpretability have been reflexively modeled in our paper as follows:

- Level 1 ("textual text"): «for the first version, each sentence was alternately written by one of the authors, so that one (and only one) sentence by A1 was followed by one (and only one) sentence by A2 (May 30); for the second version, authors could add one sentence in-between any two sentences of the first version, but each new sentence could only be introduced after a sentence not been written by the same author (May 31) — the sum of versions 1 and 2 originated the textual level that we describe as "the textual text"» ;
- Level 2 ("meta-textual text"): «for the third version, both authors commented on version 2, trying to highlight the network of concepts and associations implicit in sentences, arguments and tropes of versions 1 and 2 — this level we have called "the meta-textual text"» ;
- Level 3 ("networked text"): «they further added, as footnotes, theoretical references and examples of works and text generators that illustrated certain ideas and problems (June 1-2) — a level we referred to as "the networked text"».

This three-level division is merely a heuristics for making visible processes that are intertwined and interactional. Levels 2 and 3 do not have to be textually explicit in order to perform their function of textual mediators of level 1. What our experiment wants to highlight is that conditions for textual interpretability are not a mere internal function of the linguistic system or of the programming system. They originate in wider discursive and social processes of mediation. Delegation of symbolic production and symbolic exchange in autonomous intelligent systems is one those mediating processes.

C. Textual form and textual interpretability

From our perspective, specific techniques of production (permutational and combinatorial; statistical; stochastic; machine-learning approaches using neural network algorithms; etc.) are less important than the underlying principles of instrumentality that use the automation of language as part of the cybernetic logic of social control. We also question the conceptual division between the functional generation of natural language and the so-called creative generation of natural language, since they are equally embedded within specific discursive and social constraints, one of which is the ongoing process of automation of symbolic production (including the acts of writing, reading, speaking, listening, and translating). Instead of reifying creative computation as a special case, we analyze works that bring their own conditions of production and reception into critical focus. These works interrogate the production of the literary within current cybernetic and networked textual spaces, providing a critique of engineering approaches that work on the basis of simplified and mechanistic notions of the "literary".

What have we learned about textual production through our procedural collaboration? Each sentence establishes a particular lexical and semantic field, within particular syntactic and prosodic structures, which then become triggers for further writing through various mechanisms of semantic, phonetic, rhythmic and syntactic association (metonymic, metaphoric, paranomastic, parallelistic, etc.). Such associations are motivated by an open interpretation of the previous sentences or groups of sentences, by a self-conscious engagement with an emerging textual form, and by a network of textual references that enable each of us to generate new meanings. Semantic coherence and syntactic cohesion develop in incremental steps through recursion and revision. The act of writing extends our cognitive awareness about what might be said next as the intentionality is distributed across an accretion field of juxtaposed sentences. This process proceeds in successive loops that spiral into further ideas and sentences. It is through this embedded self-awareness that natural language parses its constituent elements for further combinations. Writing enhances this procedural dimension because the externalization of syntactic and semantic structures opens up new reading and writing possibilities. A constrained rule-based process of collaboration becomes an experiment with intentionality as the textual emergence of meaningful language, that is, language produced and interpreted by linguistically constituted subjects.

D. The embedded paper

A number of writing constraints of the mode of production of the academic paper are laid out through a procedural rule-based human generative process. Once the argumentative and discursive form of the paper begins to take shape, specific strategies for grounding concepts and theories are brought into play – quotations, references, commentary, annotations. A textual network is made explicit, and the paper's dense and abstract language is given further context. The seams that connect the various narrative levels are foregrounded by specific choices of page layout and type style (normal, bold, italics) that serve for marking interruptions and shifts in perspective. The paper struggles to retain marks of its mode of written and social production: on one hand, the sentences produced by each writer are not specifically attributed and their detailed and successive revisions are not tracked; on the other hand, the paper takes great pains to explain and self-document its constrained collaborative writing process. Its twisted, convoluted and oblique argument is kept ambiguous and open. Perhaps its aim is to show the productivity of its procedural program as a form of constrained non-algorithmic writing. Is it suggesting that this form of natural language generation cannot be automated? That this level of complexity is beyond computational creativity?

Its thematic cohesion may be said to come from a double thread in its argumentative rhetoric. One line of argumentation deals with the nature of language in relation to the self. We could sum it up in the idea that the authors explore the question of how human subjectivity is mediated through language. Another thread in the argument is its underlying concern with the political and social implications of the ongoing natural language automation. Thus the text attempts to frame the specifics of artificially generated natural language – whether as written or as spoken discourse – within general processes of algorithmic culture, which are metaphorically (and perhaps also hyperbolically) described as a mode of social engineering and control. This problematics is highlighted by the paper's slightly enigmatic title, which calls attention to the conditions of computational processing of natural language. The title can even be interpreted as a pastiche of a self-conscious snippet of pseudocode, one in which the "if-then-or-else" nested sentence structure of executable language becomes suddenly aware of the wider conditions of execution that cannot be

contained in its code – those of social action and political determination. We suggest that readers jump ahead to the last section of the first-order paper at this point. Cf. below, ["E. Coda 2."](#)

If then or else: Who for whom about what in which

Manuel Portela (University of Coimbra)

Ana Marques da Silva (University of Coimbra)

Abstract

This article discusses generativity in natural language production by adopting two different strategies: on the one hand, it reflects on its own human and collaborative process of writing as a textual instantiation of the feature of the faculty of language called "generativity"; on the other hand, it uses a series of literary generative works of different kinds to interrogate the cultural, political and aesthetic significance of the computation of language as a social practice. Computational creativity in natural language generation is thus contextualized in ongoing processes of datafication and automation of symbolic production in networked algorithmic culture.

Keywords: language and generativity; algorithmic culture; computational creativity; self-description.

1. Introduction

Incipit. This article was written by two human language generators (its authors) according to the following procedural constraints: for the first version, each sentence was alternately written by one of the authors, so that one (and only one) sentence by A1 was followed by one (and only one) sentence by A2 (May 30); for the second version, authors could add one sentence in-between any two sentences of the first version, but each new sentence could only be introduced after a sentence which had not been written by the same author (May 31) — the sum of versions 1 and 2 originated the textual level that we describe as "the textual text"; for the third version, both authors commented on version 2, trying to highlight the network of concepts and associations implicit in sentences, arguments and tropes of versions 1 and 2 — this level we have called "the meta-textual text" —, and they further added, as footnotes, theoretical references and examples of works and text generators that illustrated certain ideas and problems (June 1-2) — a level we referred to as "the networked text". Versions 1, 2 and 3 were written as running text without paragraph breaks. Finally, in the fourth version, both authors rewrote text, meta-text and networked text, defining paragraphs and sections, separating commentary and notes while integrating them into the main text, and expanding sentences from versions 1, 2 and 3 in order to fit the conventions of the academic paper and the formatting guidelines of the NAACLHLT [North American Chapter of the Association for Computational Linguistics: Human Language Technologies] template (June 5-6). In this fourth moment of composition the textual, the meta-textual and the net-textual became the (almost) "clean text" of the final draft.

Rather than offer a seamless integration of procedures and layers, we have kept several markers of those shifts and layers as far as was possible within the NAACLHLT template. This will allow readers of this paper to track some of the changes and processes that resulted in these particular textual strings, which we intend to offer as an example (and, perhaps, also as a model) of a how a natural language text is creatively generated through iteration and recursion involving two human

subjects. As can be seen by looking at its syntactic and semantic structure, textual generativity subsumes the meta-textual and the net-textual as the general condition of textual production. In programmed generativity, the question becomes: how does a computer-generated text talk about itself and how does it link itself to other texts? In other words, how can programmed generativity emulate the linguistic processes of reference and self-reference so that the particular syntactic cohesion and semantic coherence of a discursive field emerges?

The aim of this highly reflexive exercise is to highlight how the generative productivity of language is necessarily constrained by discursive and interpretative patterns, from the point of view of human production and reception, and how the computational implementation of natural language generativity should also be analyzed as a particular kind of speech act. When considered as a speech act, that is, a particular form of social action by means of language, the conditions of production and reception of computer-generated natural language cannot be accounted for without the consideration of the particular pragmatics of natural language as output of executable language and of the social actions it is meant to perform. Both process and product, computer-generated natural language instantiates the algorithmic automation of symbolic and cultural production as a stage in the development of writing media as software (Manovich 2013).

2. Who for whom about what in which

"What does it matter who is speaking", someone said.

[**Comment:** The text begins by questioning the relation between language and self. If the human speaker of language does not matter, does it matter when the generator becomes the speaker? And in what sense can the generator speak? This sentence, which was originally written by Samuel Beckett (85), has been repeatedly used for theorizing about the problems of authorship, that is, of attributing origin to a particular utterance. And yet, even when used to claim the irrelevance of a personal self as the subject of language, it is attributed to an author. It doesn't matter who is speaking but it does matter who is speaking. See Note 1.]

[**Note 1:** Philip Nickel (2013) has coined the notions of "speech actants" and "proxy speech" to account for artificial speech that fulfils the conditions of speech acts, including illocutionary and perlocutionary force: "Similarly, NLG systems do not need to have general situational awareness, adaptive intelligence and unlimited linguistic generativity in order to perform speech acts on behalf of some other agent." (500)]

Between harmony and dissonance, all voices are choirs.

[**Comment:** The second sentence expands the idea of selfless language to suggest that each voice already is a multiplicity of voices.]

Each writing creates an alien voice.

[**Comment:** The third sentence introduces writing as a mechanism for estranging the voice of the speaker. But is writing a multiplier of voices or just a technique for revealing the multiplicity of voices already contained in language?]

Constantly deferring itself. They know not what they speak.

[**Comment:** Is that a feature that the speaker shares with the generator? Not knowing what s/he speaks?]

They babble their way out of confusion. Is there language without a voice? Or a voice without a language?

[**Comment:** Now a pair of chiasmatic sentences hints at the possibility of autonomizing voice from language, but also at their nature as mutually constitutive: language developing from externalized vocalization and, at the same time, enabling the articulation of a speaking voice.]

What happens when language speaks itself?

[**Comment:** This is perhaps the core of the problem: in what sense can a language speak itself? A language must speak its material and social conditions of production. An alternative question would be: who is the subject of the textual generator?]

What is it made of? Where does its code come from? Is language a biological organism? Like a virus? An interface between the brain and the mind? Does it need a host, to speak? Am I hostage to the voice of language?

[**Comment:** Images are now associated on the basis of the bio-linguistic hypothesis for the faculty of language mixed with a theory of language as tool for the constitution of its subjects. I have a biological capacity for language but my voice is already pre-constituted in the language I have to learn to speak.]

If so, how do I get free? Is "I" a special kind of virus in the code of language? When I enter language "I" am already there. See Note 2.

[**Note 2:** Talan Memmott's "Self-Portrait(s) [as Other(s)]" (2003) is an intermedia work in which twelve self-portrait paintings and twelve biographical notes are cut-up and recombined. Described as "a recombinant portrait and biography generator", this work draws attention to the narrative conventions through which biographies are constructed, but also to the presence of others in the constitution of one's sense of self. Thus it provides an image of the fluidity of experiences and representations from which a sense of self emerges. Its pre-constitution in the conventions through which life is narrated becomes apparent in the multiplication of possibilities created by generative visual and verbal recombination, but also in its highly patterned discursive and visual structure. One could see this juxtaposition of text and image as the ensemble of discrete subject-positions that I can occupy when I self-refer to myself as "self" or as "I". The fact that it remains in constant flux, changing at each iteration, is itself an image of that process of linguistic self-production within the meaning structures of language.]

[**Comment:** Again, the text is very much aware that language provides the self with a category for him/her to participate in and appropriate its system of differences. Insofar as "I" is the category that allows for self-reference and for structuring all references in a deictic system, "I" have to enter "I" as a pre-defined variable in its semantic and syntactic system.]

Is language everywhere, and "I" a product of its code?

[**Comment:** The contrast between self and otherness thus seems to be a product of syntax, rooted in the structural and relative positions of subjects in any given context.]

I inhabit the empty self of language. Gathering its pieces, I move and play in the field of language. Strawberry fields forever. Full of sound and visions. Each word has its own viewpoint.

[**Comment:** In these five sentences, the text has linked the idea of the split-self (self as linguistic category and self as historical being) to the idea of words as discrete units of perceiving. The transition from one concept to the other is metaphorically produced by the transit created by the word "fields": *language fields, strawberry fields, sounds, visions, words, viewpoint*. What remains unclear is what is it this emptiness of language? Is it its ability for resignification through combination?]

Their lights crossing, moving everywhere. They open up perception, but they also confine us to their categories. We are grammatological creatures. Meaning as an accident of syntax, a secondary effect of permutation.

[**Comment:** Here the text suggests that meaning is a result of creativity: we cannot avoid creating meaning. Meaning isn't there, as an aspect of a thing, it is created by every subject. Hence creativity is a secondary effect of permutation, a secondary effect of our linguistic condition, since it is the structure of language that gives us a perspective on the world, as subjects. At the same time, the last sentence also points to theories of language based on the hypothesis of the emergence of the faculty of language as a consequence of genetic mutations.]

Corrupting and expanding the code. Or maybe just playing out its instructions. Where are the limits of language? Are they in the speaking body through which it speaks? And what are the limits of that body? Once embodied in writing its viral nature spreads beyond its living host.

[**Comment:** These sentences raise the question of natural language generation as the result of structural material constraints, such as a grammar or a body. At the same time, they point to an understanding of writing as the body of language, as the medium and the performance that enable the expression of the system of language. Expressing, just as computational code is expressed as it is executed, in what it generates, or writes. Language's performative existence is a creative one in the sense that it generates itself as it exists, and also in the sense that it generates things (words, concepts, mental images) as it is expressed, as it writes itself on the world and as it writes the things it names onto the world. This form of creativity is generative: it creates with no goal outside the creative act, indifferent to the value of what it creates. See Note 3.

[**Note 3:** In the words of Oliver Bown: "From the broad perspective of poesis [...] all the patterns, structures and behaviors that exist in the world can be taken as evidence of creativity. This jars with the traditional psychological view of creativity, and implies a distinction between two varieties, generative and adaptive. Generative creativity takes an indifferent approach to the problem of value, it is value-free creativity. In generative creativity, things are not created for a purpose. Things can come into existence without being created for their value" (2012: 363).]

Inhabiting everything we see. To read is to be infected by the written virus of the code of language. Hopelessly finding meanings everywhere. Finding one's voice in alien snippets of code. Looking for and testing the possibilities of the code. Saying what has not been said before, letting language invent itself.

[**Comment:** Here the text returns to the question of the relationship between subjectivity and the production of meaning, highlighting how the latter may be understood as a result of a generative and creative process.]

Letting the code express itself. Like a blind man lost in the desert, laying stones and little sticks to build a map. A map without a territory, referring only to itself, full of sound and fury.

[**Comment:** A series of sentences about the creativity inherent in the proliferation of language leads to Macbeth's speech about the brevity and meaninglessness of human existence, and thus about the meaninglessness of language as description of experience.]

Making something from the empty self of language. Sensing the passing of time in the rhythm of language. Existing in the places invented through language. Searching for language, for more language, searching with language for more language. Creating new places for language to grow, serving nothing but language itself. ~~Every body is a speaker, building itself through its voice and the voices around itself.~~

3. If then or else

And yet, if language is a tool for being, what happens when its self-replicative processes are abstracted from sentience?

[**Comment:** This self-referential proliferation of the empty meaninglessness of language seems significantly different from Macbeth's existential expression of the madness and pointlessness of ambition, revenge, remorse, guilt, fear, desire. Perhaps that is what is meant by "abstracted from sentience": once disembodied from intentions and situational contexts, the text is sequestered by the mechanism of its machinic production. See Note 4.

[**Note 4:** An extreme example of this combinatorial logic can be seen in the "Library of Babel" (2015-2017) by Jonathan Basile, a computational interpretation of Jorge Luis Borges' "Library of Babel", which "demonstrates the paradoxical effect of automating endless factorial permutations of the alphabet. On the one hand, the relentless logic of the algorithm results in the constrained expression of purely abstract differences that instantiate themselves as a textile of letters, punctuation marks and blank spaces. On the other hand, the impossibility of exhausting semiosis through the sheer force of calculus becomes evident as meaning can only happen probabilistically, discontinuously and interactively at scales other than the highly granular and machinic character by character permutation. Even if seen as a conceptual enactment of the continuum of expression upon which signifiers cut out their own form as differential meaningful strings, Basile's experiment shows the profound alien nature of the semiotic excess of computationally constrained writing in its literalized and randomized production of alphabetic infinity." (Portela 2017)]

In such an abstract environment, how does feedback work? Can a language generator feel its own use of language, or is it just a simulacrum of subjectivity?

[**Comment:** These two questions point to the fact that language is not transparent and neither is code: both are inevitably embedded with human intentionality.]

Maybe it is like a bat, blindly navigating the vastness of the code's combinations and comparing different morphologies in space. Echolocations of the world, words are deflected by objects into new directions. Reflecting, mixing, deforming and carrying the sounds of those objects toward new directions. The unheard of frequencies of speech sounds parsed by means of the discreteness of letters. See Note 5.

[**Note 5:** *Automatype* (2012), for instance, is a literary experiment by Daniel C. Howe that "uses algorithms to find the bridges between English words, Six-Degrees-of-Kevin-Bacon-style — not bridges of garbled nonsense but composed of normative English." (Howe 2012). Another example of similar processes is *ppg256* (2012), a series of poetry generators by Nick Montfort: "I determined

that common initial bigrams and common final bigrams of four-letter words could be joined uniformly at random to produce 450 distinct four letter words, 273 of which (more than 60%) were dictionary words." (Montfort 2012)]

[**Comment:** This set of images point to the notion of machine creativity as a generative process, based on the decomposition of words and sentences into their core and/or minimal elements, and on the derivation that results from the re-composition of those minimal elements into new linguistic units, according to the specific set of rules that determines a given process, such as poetry generation or computer-assisted translation. See Note 6.

[**Note 6:** AI models of creativity fall into two broad groups, because creativity itself is of two types. On the one hand, there is what we may call 'combinational' creativity. Here, the novel idea consists of an unusual combination of, or association between, familiar ideas. Poetic imagery, metaphor, and analogy fall into this class. On the other hand there is exploratory-transformational creativity, grounded in a richly structured conceptual space. A conceptual space is an accepted style of thinking in a particular domain — for instance, in mathematics or biology, in various kinds of literature, or in the visual or performing arts. (Boden, 2009)]

The rules structuring how novelty may be composed. Writing already is a computation of natural language, a machine for exploring the probabilities in its code. An automated writing machine has many different kinds of listeners.

[**Comment:** The last sentence highlights the distributed condition of computation, stressing that an automatic language generator writes and speaks not only to and with humans but also to and with other machines, or programs upon which it depends. These nets or meshes of interconnected algorithms are part of the infrastructure of digital language.]

Including those who listen for controlling, processing and measuring generated language. Scanning the context, weighing and comparing the generated language with all the natural language it reads as it writes. The algorithm is a social form with situated intentions, not a naturally occurring event, and not a linguistic fact. Enclosed in layers of opaque objects and relations, can this writing machine be understood and mastered? Objects will speak with us and they will speak for us. As we become their fuel. Clouds of networked writing processed in real time are scripting back the generation of natural language. In a constant and recursive movement, I emulate the language that emulates language. Will speaking objects write us out of language? A matrix feeding on the language we produce. We teach the machine to speak for us. As we speak with it and as it speaks through us. An evolving machine. The externalization of linguistic production is a new social fact. The web as a living archive for writing and speech. A prosthetic reflection of the cultural field. A biological self is no longer required for the computation of language. Abstracted from speaking bodies, language is processed and generated as a hybrid material made of different semiotic regimes. Relentless iteration of combinations towards pure discursive forms: filling in the blanks for poems, stories, screen-scripts, news articles. Following and reinforcing established models. It can run on endless loops from circuit to circuit. In a recursive process of translation, it becomes a conversation between machines. We sit back and enjoy the show as all symbolic production is automated and delegated. At once spectators and characters. We listen in on their data crunching, moved and alienated by their noise. But do we understand their speech? They garble their way through unicode letter by letter. See Note 7. Code: <https://github.com/jhave/Big-Data-Poetry>

[**Comment:** This section reflects on the material (technical, economic, political, cultural) situations of digital writing, positing it in a set of social conditions. More than a medium, and more than an organ, language is here understood as an externalized technology, or a prosthesis.]

[**Note 7:** In his project *Big Data Poetry* (2014-2017), David Jhave Johnston uses machine learning techniques to generate strings of language. *BDP* uses a combination of techniques of visualization, analysis, classification and substitution of objects, applying these to a corpus of language made of hundreds of thousands of lines of poetry. The result is a disarticulate and incoherent mass of language, on which the poet works by means of improvised reading, stitching together the generated language in order to transform it into a meaningful poetic experiment. See Note 8. Data: <http://www.macs.hw.ac.uk/iLabArchive...>

[**Note 8:** Efficiency of statistical natural language generators depends on the granularity of semantic annotation on the training data (such as word-level or phrase-level annotation). "Stochastic Language Generation in Dialogue Using Factored Language Models" (Mairesse and Young, 2014) illustrates the complexities of designing a dialogue system whose predicted variables can be conditioned by different utterance contexts. Since any training has to occur within a limited corpus — in this instance the corpus of the Cambridge Tourist Information System —, language generation is a constrained computational expression of a discourse field. In other words, it is a mathematical disciplining tool which scripts the behavior of the human interlocutor to match the range of probabilities of its pre-defined utterances or its generated paraphrases.]

We, as unstable terms of comparison for algorithmically generated language. Unlike us, they only know the language they use as well-formed character strings. They blindly follow the rules that declare their semantic representation. Even when they machine-learn their way into further production and reproduction. Their cognitive processes as a mesh of mathematical threads, too flat and too fast for us to understand.

4. Yet but however

Like us, they cannot own the language they speak. The code that speaks through us speaks through them. Constantly circulating through the social engine. Defining our subject positions as natural language generators. Our speaking bodies as complex and subtle machines, feeding the cybernetic machine. Their processes are dependent on databanks where language is enclosed. See Note 9.

[**Comment:** These sentences point to some of the common aspects between artificial and natural language generators, or between computers and human speakers, highlighting how both humans and machines are situated in a linguistic system that depends on privately owned infrastructures.]

[**Note 9:** In *How It Is in Common Tongues* (2012), John Cayley and Daniel C. Howe programmed a series of n-gram searches using Google's search engine, taking the whole of the Internet as a database for making searches of combinations of strings of words that replicate Samuel Beckett's *How It Is*. This work renders explicit the appropriation and monetizing of the commons of language by Google, while also applying strategies of subversion that defy the unilateral terms of use that regulate the relationships between Google and its users. (Cayley 2012).]

I can only enter into contractual relations that further determine the language contract. I can only move and speak in predetermined paths, where and as allowed. See Note 10.

[**Note 10:** Sandy Baldwin (2015) describes the Internet not as the democratic rhizome promoted by the rhetoric of Silicon Valley in the 1990s, but as an infrastructure that reflects and intensifies contemporary neo-liberal macro-structures. Interweaving the history of the network with the analysis of gestures such as sending an email, accessing a website or signing in, Baldwin demonstrates how "we constantly enter into consensual relations with the opacity of a technical infrastructure" (58).]

Constrained by the computation of the grammar of language. And constrained by the infrastructures of computation. Language becomes a dataset of statistically relevant occurrences that can be mined for further language production and for granular analysis of individual desires and patterns of thought. A guessing machine, designed to optimize the world as a resource. Of that of which I can speak and of that of which I cannot speak, the program will not remain silent. See Note 11.

[**Comment:** Here the text further reflects on the digitization of language as a social process that renders it into a raw material and a source of value, and which could be characterized as cybernetic in the sense that it enacts a network of systems that monitor, evaluate, categorize, guide and sustain digital communication. By parodying Wittgenstein —"and whereof one *cannot speak*, thereof one must be silent" (23) —, the last sentence suggests that digitization extends the power of the symbolic to all domains of experience.]

[**Note 11:** John Cayley's *The Listeners* (2015) is a literary experiment in which the author programs a "skill" for Amazon's domestic AI (Alexa). This work adds a layer of programming to the default programming of this device, highlighting the ways in which the original programming is embedded with the values that give form to such corporations. More specifically, this work calls attention to the problems of surveillance and control raised by domestic intelligent devices, and it highlights how the internet may be understood as an unbound mass of language generated in real time by human speakers: each of our online movements generates a trace that augments the web, which may be described as an evolving linguistic database. At the same time, this work problematizes authorship and the conditions of possibility for literary production, by actively subverting the unilateral terms and protocols that structure and sustain digital language.]

A tool and a material at the same time, natural language processing becomes the glue or the ground of the cybernetic organization of the world. The world as computation and representation. The simulacrum as truth. One algorithm at a time. The true human-computer interface, the interface of interfaces. Mediating and digitizing all life. Juxtaposition of encodings.

[**Comment:** This section refers to the continuum between the digitization of language and the digitization of the world, or of our perception of reality, increasingly mediated by and encoded in binary systems.]

Reinforcing power relations, this post-human language becomes value. The commodification of language began with the selling of stories and poems and songs and with the selling of writing, but real-time analysis and real-time generation of language takes it to a different scale. Externalizing language into structures we do not control or understand. Do we have enough perspective to understand this moment in history? When all objects become infected with the virus of computer-generated natural language? Talking cars, talking elevators, talking gas stations, intelligent domestic devices. Seamless integration of utterance-producing appliances and devices. Shiny new

toys, magical and powerful toys regulating our moves. I say to my car, "talk to me". The consensual illusion of having a car "talking" to me.

[**Comment:** Here the text further dwells on the question of the opaqueness of digital interfaces and it highlights how the suspension of disbelief, as in our experience of fiction, blurs our perception of such intelligent technologies, which thus become fetishized, just as totemic figures.]

Why do we want to produce language with language-producing machines? Increasingly situated in a grid made of synthetic language, can we still speak outside the interface? Outside its strictly functional and managed rhetoric? Am I a soldier, a piece of the machine?

[**Comment:** This set of questions suggests that the opacity of intelligent technologies turns users into functionaries, in Flusserian terms, since users become the variable while the device becomes the constant. See Note 12.]

[**Note 12:** Every program functions as a function of a metaprogram and the programmers of a program are functionaries of this metaprogram. (Flusser, 2006: 29)]

What do I compute when I say "I"? If my language is commodified, am I a hostage of this distributed and omnipresent speaking and writing machine? Whose language am I programming? Who owns the tools, how do we learn how to rewrite the program? The network as vast word processor sustaining billions of local linguistic events has changed the ecology of language uses. Reorganized to fit a top-down structure. To conform the production of meaning to mere transcoding as in computer-assisted translation or in text-to-speech and speech-to-text applications. If machine creativity is a derivation of vertically established power relations, how can we consciously use it?

5. Conclusion

This paper has no conclusion. It is an open-ended writing experiment about a collaborative writing process that offers itself as evidence of the complexities of both non-formalized and formalized natural language generativity. Its aim is to show the heterogeneity of any human- or machine-generated natural language utterance as a particular speech act, which involves the creation of discursive conditions for the interpretability of its utterances beyond the discrete parsing of its constitutive elements. In the present case, the textual dynamics of text, meta-text and textual network was illustrated by means of the literary form of the academic paper. Several generative works were analyzed as creative practices that use computational generativity to interrogate the ongoing automation of natural language production. *Explicit.*

Colophon

This text was begun on May 30, 2017, 9:25 am. This text was finished on June 6, 2017, 6:55 pm.

Acknowledgments

Foundation for Science and Technology (FCT). PhD fellowship reference: PD/BD/52247/2013.

References

- Baldwin, Sandy. 2015. *The Internet Unconscious: On the Subject of Electronic Literature*. London: Bloomsbury.
- Basile, Jonathan. 2015-2017. *Library of Babel*. <https://libraryofbabel.info/>

- Beckett, Samuel. 1994. *Stories and Texts for Nothing*. London: Grove Press.
- Boden, Margaret A. 2009. Computer Models of Creativity. *AI Magazine* 30.3: 23-34.
- Bown, Oliver. 2012. Generative and Adaptive Creativity: A Unified Approach to Creativity in Nature, Humans and Machines. Jon McCormack and Mark d'Inverno (eds.). *Computers and Creativity*. Berlin: Springer. 361-381.
- Cayley, John. 2015. *The Listeners*. <http://programmatology.shadoof.net/?thelisteners>
- Cayley, John, and Daniel C. Howe. 2012. *How It Is in Common Tongues*. Providence, RI: Natural Language Liberation Front. <http://thereadersproject.org/hiiict2012.html>
- Flusser, Vilém. 2006. *Towards a Philosophy of Photography*. London: Reaktion Books.
- Howe, Daniel C. 2012. *Automatype*. <https://rednoise.org/~dhowe/automatype/>
- Johnston, David Jhave. 2014-2017. *BDP: Big Data Poetry*. <http://bdp.glia.ca/>
- Mairesse, François, and Steve Young. 2014. Stochastic Language Generation in Dialogue Using Factored Language Models. *Computational Linguistics*. 40.4: 763-799.
- Manovich, Lev. 2013. *Software Takes Command*. London: Bloomsbury.
- Memmott, Talan. 2003. *Self Portrait(s) [as Other(s)]*. Hayles, N. Katherine, Nick Montfort, Scott Rettberg, and Stephanie Strickland, eds. (2006). *Electronic Literature Collection (volume 1)*. College Park, Maryland: University of Maryland. [.../memmott_self_portraits_as_others.html](http://memmott_self_portraits_as_others.html)
- Montfort, Nick. 2012. *XS, S, M, L: Creative Text Generators of Different Scales*. Trope Tank, MIT. <https://dspace.mit.edu/handle/1721.1/78887>
- Nickel, Philip J. 2013. Artificial Speech and Its Authors. *Minds & Machines*, 23: 489-502. DOI 10.1007/s11023-013-9303-9
- Portela, Manuel. 2017. Writing under Constraint of the Regime of Computation. Joseph Tabbi, ed. *The Bloomsbury Handbook of Electronic Literature*. London: Bloomsbury. 181-200.
- Wittgenstein, Ludwig. 2016. *Tractatus Logico-Philosophicus*. Translated by K. Ogden. Chiron Academic Press.

E. Coda 1

The reviews clearly identify the major flaws and inadequacies of "If then or else: Who for whom about what in which" as a research paper. Reviewers acknowledge its parodic and performative structure, but also its failure to engage with state-of-the-art research in the field. They rightly point out the pointlessness of the experiment for automated natural language generation, and its insufficient reflexivity about the writing experiment itself.

----- REVIEW 1 -----

This papers discusses generativity in natural language production on an intriguing, self-reflective meta level. The paper reads more like a work of art -- the authors call it an "experiment" -- than an academic paper (although it includes a number of theoretical references and considerations). This makes it hard to assess whether the paper fits the scope of the workshop and, perhaps more acutely, how the oral presentation would be organized. Because of the lack of a solid theoretical or practical

conclusion, I am not tending towards recommending acceptance at the workshop, which was primarily intended an academic event.

Some of the main issues which I see, include:

- The paper promises to offer recommendations as to how a computational creativity can/should be practically implemented but these recommendations are hard to find in the text, which in fact offers very few observations as to the computational/digital aspect of the matter. In this sense, the paper does not live up to the promises made in the abstract, which is a clear weakness that should be addressed.
- Most tangible, scientific claims are included in the form of quotations from existing papers (and literary authors), and the individual novelty of the paper is therefore hard to assess but probably limited. The authors could have done a better job at highlighting the novelty of their own contribution.
- The academic literature which is processed in the paper seems like a relatively random sample and it not presented in a clear structure.

<p > The comments are an interesting stylistic feature of the paper, but they are also puzzling to the reader because their status remains somewhat unclear: do they comment on the writing process while being also a part of it? Then how is their status then different from the running text?

----- REVIEW 2 -----

The paper presents a curious experiment on language generation. The two authors of the paper wrote the text in four different rounds. In the first one each author wrote one sentence alternatively, in the second each author could include new sentences after sentences written by the other author, and in the third and fourth rounds the text was commented and annotated with extra information.

The experiment presented in the paper is novel and interesting. However, even when the paper is written in fluent English, due to its nature it is quite dense and philosophical in several points. This problem is increased because the goal of the paper is not clear, so I felt lost in several points and not sure about what the authors were trying to transmit. NLG systems usually have a goal in mind when generating a text. What was the goal of the authors when generating theirs?

In addition, although the authors state that "this article discusses generativity in natural language production" and "the aim of this exercise is to highlight how the generativity productivity of language is necessarily constrained by discursive and interpretative patterns", the paper lacks a proper discussion about these points and the relation of the obtained text and the fields of Computational Creativity and Natural Language Generation. The authors should state clearly the main insights learnt from the experiment, and how they could be useful for the automatic generation of text.

----- REVIEW 3 -----

This paper explores the process of language generation as a product of different components: the language building blocks and restrictions, the producer of the language and all the language that has already been processed by the producer, the pragmatic embedding of any utterance, the cultural influences on language and interpretation, etc. The authors have chosen an original form, by guiding their writing process in different stages and explicating these stages in the resulting text. It is their aim to show how computer generated text will, just like human utterances, be interpreted as a speech act, a social action.

The paper is rather philosophical, asking several open questions. In this sense, it definitely succeeds in providing the reader with food for thought. The 'meta-textual text' gives useful context and the 'networked text' links this paper to works on natural language generation, some of which applicational, to show recent developments. The text has a high density of ideas. As both the content and the format of the text play an important role in the message that is conveyed, it is hard to condense a clear line from the paper. It might be good to add some more 'meta-meta' text, guiding the reader a better idea of the main story. Also, it lacks a clear message to the scientific community. Where to go from here?

The paper is open-ended, but the authors could have gone further than they did now. For example, version 3 is now clearly highlighted as an addition to versions 1 and 2. However, it is not clear how version 1 was changed into version 2 by adding sentences. It would be interesting to see which parts were added during this stage. In addition, the authors do not elaborate on their experience during this collaborative writing experiment. How did the imposed restrictions influence their writing, and what does this imply for automatically generated text?

Besides from these points, I think the endeavour original enough to deserve a venue. We suggest that readers jump back to Section A of the first-order paper at this point. Cf. above, ["A. Context."](#)

E. Coda 2

The paper is ultimately unable to tell what it means. Why? How relevant is this conceptual writing experiment for computational creativity in natural language generation? We think that our initial question may have to be rephrased in a different form: when and how can we say that a textual form satisfies its minimum conditions for interpretability? In other words: can creative natural language generation simulate reference and self-reference in ways that result in the emergence of interpretable textual forms, that is, *of forms that perform their own actions* rather than acting as proxy speech actants (Nickel 2013) who act on behalf of some other agent? Proxy speech actants *of whose language uses our human actions will become perlocutionary effects*? Is a fully externalized generative system for producing natural language the ultimate extirpation of the self who is finally deprived of the interface to itself? We can only speculate.

References

Gervás, Pablo. 2017. 'Template-Free Construction of Poems with Thematic Cohesion and Enjambment'. *Proceedings of the Workshop on Computational Creativity in Natural Language Generation (CC-NLG 2017)*, 21–28.

Manjavacas, Enrique, Folgert Karsdorp, Ben Burtenshaw and Mike Kestemont. 2017. 'Synthetic Literature: Writing Science Fiction in a Co-Creative Process'. *Proceedings of the Workshop on Computational Creativity in Natural Language Generation (CC-NLG 2017)*. 29-37.

Manovich, Lev. 2013. *Software Takes Command*. New York: Bloomsbury Academic.

Marques da Silva, Ana. 2016. 'Speaking to Listening Machines: Literary Experiments with Control Interfaces'. In *Interface Politics: 1st International Conference 2016*, 681–90. Barcelona: Publicaciones GREDITS.

———. 2017. 'Zoom in, Zom out, Refocus: Is a Global Electronic Literature Possible?' *Hyperrhiz: New Media Cultures*, no. 16. <http://hyperrhiz.io/hyperrhiz16/essays/1-da-silva-world-elit-possible.html>.

Nickel, Philip J. 2013. 'Artificial Speech and Its Authors'. *Minds and Machines* 23 (4): 489–502. <https://doi.org/10.1007/s11023-013-9303-9>.

Portela, Manuel. 2013. *Scripting Reading Motions: The Codex and the Computer as Self-Reflexive Machines*. Cambridge, MA: The MIT Press.

———. 2017. 'Writing under Constraint of the Regime of Computation'. In *The Bloomsbury Handbook of Electronic Literature*, edited by Joseph Tabbi, 181–200. London: Bloomsbury Academic.

Cite this Essay:

Portela, Manuel and Ana Marques da Silva. “A Strange Metapaper on Computing Natural Language”, *Electronic Book Review*, September 2, 2018, <https://doi.org/10.7273/6gws-hf93>.

Readers wishing to respond to an essay in *ebr* may send ripostes or short glosses to the journal's Managing Editor, [Will Luers](#).

[PDF](#)

This essay was peer-reviewed.