

Speaking to Listening Machines: Literary Experiments with Aural Interfaces

 electronicbookreview.com/essay/speaking-to-listening-machines-literary-experiments-with-aural-interfaces

Ana Marques da Silva

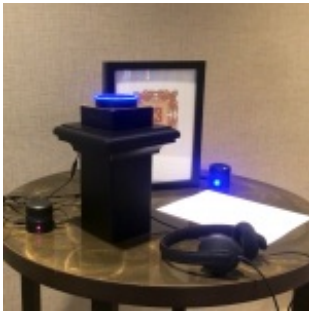
by Ana Marques da Silva

05-16-2017

thread: electropoetics

This essay was peer-reviewed.

electropoetics



Reading practices have changed along the course of history. Before the ‘democratization’ of the written word - from Homer’s *Iliad* to the medieval troubadours and to more recent public and private oral reading traditions -, reading has long been associated with listening. Today, in the age of algorithms and ‘smart’ interfaces, the sharing of language between humans and computational devices is increasingly ubiquitous and, with the standarization of artificial intelligence systems like Siri, Cortana, and Google Now, we are starting to speak

and to listen to machines. In the field of digital literary creation, one example of aesthetic reflection on the questions raised by such networked ‘smart’ interfaces is John Cayley’s *The Listeners* (2015), "a linguistic performance — transacted by visitors and Amazon’s voice-activated Artificial Intelligence and domestic robot, Alexa" (Cayley, 2015b). Through an analysis of *The Listeners*, articulated with Bernard Stiegler’s notion of the digital *pharmakon*, this paper aims to reflect on the encounter between literature and digital technologies. Three ideas will be highlighted: 1) the ways in which the technical, economic and political layers that constitute our digital devices pre-determine their usage (how they operate and are operated); 2) the automatic processing of language and orality as interfaces of mediation between humans and “smart” devices; 3) the literary implications of aurality and *aurature*.

INTRODUCTION

In his book *The Interface Effect* (2012), Alexander Galloway considers how interfaces are not simply tools or stable objects, but “effects” (33) of concrete material conditions, as well as “practices of mediation” (16) that reflect culture. Computational devices are thus not simply machines that emulate other media, but translation processes occurring between many layers of code. Behind the surface-level of the interface, myriads of performances take place, too small and too fast for the human eye to perceive.

Articulated with these, there are layers of protocols to which these processes must comply in order to be interpreted. These protocols (FTP, SMTP, TCP/IP, etc.) establish the underlying rules that structure networked digital communication and data

transfers, regulating the space of the web. We may thus think of the Internet as an enormous and unbound text whose grammar is constituted by protocological regulations. In *The Internet Unconscious* (2015) Sandy Baldwin considers how this network of permissions constrains every act of digital writing and poses the question of whether we may speak of a truly literary act within the regulated space of the network. As an interface between these protocological regulations and users, there are terms of use to which the latter must agree in order to have an email account, install a given software or access social media, terms that are unilateral and presented in such a way (as enormous bureaucratic texts) that they become unreadable to most users. As a consequence, we increasingly accept our interfaces' terms of use without actually knowing those terms, or without knowing how the information we produce is processed and to what ends. This commercial and institutional over-determination has become naturalized, as if it were a price to pay for global and immediate interconnectivity.

All these different dimensions of computational performativity are invisible or obscured

1
"what is going on within the complex remains concealed: a 'black box' in fact" (16). "No photographer, not even the totality of all photographers, can entirely get to the bottom of what a correctly programmed camera is up to. It is a black box" (Flusser, 2000: 27).

by the black box inside our apparently transparent digital interfaces. Computers, which started as programmable devices, are now increasingly opaque and closed by layers of proprietary software designed mostly for instrumental manipulation. At the same time, the greater the black box is, the greater is the interface's smoothness and

transparency. Interfaces are thus imbued with politics, as they reflect and reinforce the institutional and systemic matrix that contextualizes them and from which they emerge.

Vilém Flusser considered the question of the relationship between mediation and creation (or of technical devices and their use) in the following terms:

The camera is not a tool but a plaything, and a photographer is not a worker but a player: not *Homo faber* but *Homo ludens*. Yet photographers do not play with their plaything but against it. They creep into the camera in order to bring to light the tricks concealed within. Unlike manual workers surrounded by their tools and industrial workers standing at their machines, photographers are inside their apparatus and bound up with it. This is a new kind of function in which human beings are neither the constant nor the variable, in which human beings and apparatus merge into a unity. It is therefore appropriate to call photographers functionaries. (Flusser, 2000: 27).

As a “plaything”, technology reflects the agency of human beings as *homo ludens* (Huizinga, 1949), the play being both in the creative work (defunctionalizing the *apparatus*), and in the individual and social reception of the work (or in the struggle between normativity and novelty). The problem is that in the context of complex and opaque “playthings”, such as photographic cameras or computers, a creator is “inside” and “bound up” with its medium, in a situation of negative identity (as it dilutes the boundaries between subject and object) in which the creator becomes intertwined with what is no more a tool or an instrument, but an *apparatus*. While manual workers are the constant and their tools a variable, and while industrial workers are the variable and their machines the constant, in the case of the “functionary” we witness a collapse

of the functions of variability and constancy, as he is already part of or situated within the *apparatus*. In a digital context, it is thus crucial to acknowledge the fact that makers are functionaries, so that it becomes possible to draw a critical theory and practice that may allow for an emancipation from the constraints of the *apparatus*, making it a tool and a variable.

Since interfaces, or media in a broader sense, are results of the material conditions that characterize each particular moment in history, an interface is, in Galloway's words, an "allegorical device that will help us gain some perspective on culture" (2012: 54), a device that makes the world visible, helping us to make sense of it. Amazon's domestic AI may thus be understood as an allegorical device through which we can grasp how the contemporary subject is redefined by technology, and how technology is shaped by the social field.

Today, from automatic language generation to automated processes of all sorts, humans are increasingly interacting with a variety of artificial intelligence systems, from news writing bots to self-driving vehicles. Digital interfaces are thus starting to operate as autonomous agents, gathering, processing and generating information, capturing and structuring the flows of data that we, and our machines, produce. In 2015, John Cayley took on the task of poetically experimenting with one of these algorithmic interfaces: Alexa, Amazon's voice assistant. Alexa is the voice service that equips Amazon's 'Echo', the company's domestic robot. This 'echo' is a small black cylinder, designed to be at our homes, and equipped with an array of seven microphones attuned for human voice recognition. Equipped with Alexa, this device listens and speaks, and it "can also be configured to read out loud from arbitrary texts of our choice on computers" (Cayley, 2015a). But Alexa was not primarily designed to be a reading interface: rather, Amazon describes it as device able "to provide information, answer questions, play music, read the news, check sports scores or the weather, and more". Whenever its name is pronounced, Alexa "wakes up" and sends all it "hears" to the web, for processing by Amazon.

2

<http://www.amazon.com/Amazon-SK705DI-Echo/dp/BooX4WHP5E>

1) ALEXA AS A CONTROL INTERFACE

At first glance, Alexa could be described as something between a smart device and a sales assistant, but Amazon's Echo is more than that: Alexa asks for our attention, it chats us up, captures our voices, our language and the sounds of our homes, and it sends all this information to the web where it is stored and treated by Amazon, with the proclaimed aim of improving Alexa's skills. This robotic "personal assistant" is thus an interface between typically closed and personal spaces and the open and shared space of the Internet, as a bridge dissolving the frontiers between

3

See:
<https://www.amazon.com/gp/help/customer/display.html?nodeId=201602230>

the private and public spheres. But unlike 1984's Big Brother, Alexa sounds pleasant and always ready to answer its user's demands. Alexa is also an interface for Amazon at "the networked global vending machine for everything", connecting this tentacular and distributed corporation to

4

(John Cayley, private email)

users. "Just ask" is the slogan that accompanies Amazon's smiley logo. While the slogan conveys easiness, the logo conveys a kind of infantile looking sense of happy pleasantness. The first impressions of this product brand seem to be in accordance with the ways in which many product providing companies portray their image and, by extension, their clients': in a candid and infantile way. In the case of Alexa, this feature is evident for example in the casual response it gave to a Youtube user who asked "Alexa, is there a Santa?", to which it responded "I don't

know him personally, but I have heard a lot of good things about Santa. If I ever meet him I will tell you."

5

Minute 3:03 in:

<https://www.youtube.com/watch?v=EavnIXcWvyM>

him I will tell you."

The learning capacity of artificial intelligence agents seems to be a factor leading Alexa's users to accept the fact that it is connected to the Internet, as if the disappearance of our private spaces were a trade-off for having a well-trained "intelligent" gadget. Being connected to the Internet enables this device to establish a communication loop between its users and Amazon's central services, enabling it to "learn" as it is used. Users are thus part of a cybernetic system, of a closed system of control and communication of both machines and living beings, just as Norbert Wiener first defined cybernetics in the

6

We may also point to other popular interfaces dissolving the frontiers between private and public spaces, such as Pokemon Go, for instance.

title of his seminal book from 1948. In such a cybernetic system, the control over users' actions and language, or their communication, generates data and value, feeding back the system and its underlying premises in a loop. In

7

Wiener, Norbert (1948), *Cybernetics, or: Control and Communication in the Animal and the Machine*, Paris and Cambridge

this context, Alexa seems to represent at once a step forward in what concerns the globalized digital *panopticon*, as well as the easiness with which users accept and often welcome the presence of "intelligent" devices that track not only their movements in space and their Internet behavior, but also their speech and all the cultural value associated with it, which is captured from the commons and monetized.

Contiguous to the problem of the relationship between smart devices and surveillance is the question of emergent modes of distributed machine cognition associated with the Internet of Things: when isolated, intelligent devices are restricted to the functions they are programmed to execute, but once several devices are interconnected in networks, they become part of complex systems in which the intermediation between its constitutive elements gives rise to the emergence of complexity. This dynamic is the same as swarm modes of cognition, like beehives or ant colonies. The tendency we are today witnessing towards total automation, with "smart" homes and cities, implies a much needed reflection on what it may mean for human life to become surrounded by and inscribed in grids of networked artificial intelligence systems developed under capitalist modes of production and consumption.

And this leads to another question: who owns Alexa? Users do, in the sense that they pay for it and use it. But we may also argue that Amazon owns Alexa in the sense that Alexa enables the connection between users and Amazon's services. Alexa is thus an extension of both its users and of Amazon. It is an extension of the first because it works

as a tool for a number of different tasks, and it is an extension of the latter in the sense that it is its "ears" and "mouth": "ears" that record and send all that is heard for processing, and "mouth" that speaks, giving voice to Amazon's tentacular machine. And so we speak to it, giving away information that feeds a data driven market. Here the product is not the interface in itself, but the information it generates through the ways in which users use it. In this sense, the product, or the place where value resides is not so much Alexa, but its users and the information they generate. Or more precisely: the product is the information generated by users, and hence users are the producers, although they have no control over their production. Users become the producers of the product Amazon invests in: data. This mechanism of appropriation applies to all our gestures online, since our life on the Internet is a permanent production of data, the most abundant commodity of our time.

2) SOUND, LANGUAGE AND THE DIGITAL PHARMAKON

What kind of interface is Alexa, exactly? What does it consist of? It is a three-dimensional object filled with microphones, and it is a series of distributed code processes, but the mediation between the human and the machine is accomplished through voice - the user's voice and Alexa's voice. Voice is the interface for mediation, and human language (along with the data it generates) is at once the content and the currency.

Today's race towards Artificial Intelligence seems to be taking its first steps in aural, or

8

Aural refers to hearing, while *oral* refers to speaking.

sound, interfaces. AI systems, at least those that are emerging in everyday life through smart devices, are being introduced to users through sound and, more specifically,

through voice. Alexa's voice is feminine, articulated and smooth. It sounds human, it

9

The voice behind Siri (Apple's voice-activated virtual assistant) does actually belong to human actresses whose voices are recorded and worked, isolating diphthongs, syllables and phonemes, adjusting speed and pitch, and undergoing a process called concatenation in order to build words and sentences. <http://edition.cnn.com/2013/10/04/tech/mobile/bennett-siri-iphone-voice/>

has

a name, it is accurate in the interpretation of what it hears, it is quick to respond and linguistically fluid. The logic behind the human-like perception of Alexa is the same that tends to make interfaces transparent, easy and intuitive, and so Alexa's machinic aspect is diluted by its humanoid voice. Alexa's voice is deeply integrated on the cultural

unconscious and collective psyche: from the women's social history standpoint, Alexa's voice embodies and reinforces the historical process of constituting a feminine identity associated with attributes such as serenity and sweetness. Media archeology shows how the feminization of mediating voices was already present in the late 19th century, with young women working as telephone operators instructed to reproduce certain vocal characteristics. Alexa's transparency is thus not simply built upon its machinic efficiency, but also on a feminine prosody that is historically situated.

In Spike Jonze's film *Her*, the digital voice of the operating system becomes its medium for materializing presence, becoming its body. Similarly, Alexa's voice embodies an abstract entity, rendering it 'tangible'. Alexa's voice is an avatarization of Amazon, and it is also, and at the same time, an avatarization of the abstract alterity on which the individual and collective unconscious is projected. Indeed, Alexa is not an entity, it has no autonomous agency. The humanizing perception of Alexa is an illusion (or a consensual hallucination, to recall William Gibson's view of the Internet in *Neuromancer*) only made possible through the suspension of disbelief, as happens with our experience of fiction. Alexa's voice is thus an already naturalized acoustic hallucination, allowing for a disembodied voice to "speak with" and "for" us, and to simulate "listening to us". This form of fetishization of an opaque technical device is close to the fetishization of totemic figures, establishing a continuum between the human and the non-human (transferring "transcendence" to the apparatus within which the functionary is entangled), and it raises questions of what it means to be in a cultural situation in which this mode of representation and blurring of boundaries becomes naturalized.

We may also think of Alexa's voice as an avatar, at least in two different ways: because it is a simulation of presence, or presence at distance, embodied in a programmed voice, and because it is a blend of bios and technics. In his text "Voice of Avatar, Voice as

10

My translation from the portuguese ("Voz do Avatar, Voz como Avatar, Avatar da Voz").

"Avatar, Avatar of Voice", Pedro Serra considered how 'voice as avatar' is at once a representation (simulacrum) and a materialization (presence) of the author's voice that resonates (that is mediated) in a surface of inscription (be

it graphic or phonographic) (Serra, 2015: 16-17). This avataric voice, present and absent at the same time, extends the voice of the author as a simulation, or a construction, or an imaginary projection. It is thus a hybrid object, imbued with subjectivity and technics. But this blurring of boundaries between self and technics must not be understood as the hybridity prefigured by the notion of the cyborg: if the cyborg is material and a figure of symbiosis (adequate to theorize the Flusserian functionary), the avatar is virtual and a figure of tension in which its parts (presence and absence; bios and technics) never fully merge.

The anthropomorphism and immediacy of Alexa's voice (the artifices that render the consensual illusion possible) seem to lead users to trust it. According to Amazon, Alexa does not stream its users' speech until a fraction of a second before it hears its wake word, which is its name. But questions regarding this claim are being posed by users who got surprised by Alexa's unsolicited participation in conversations. Another important aspect regarding the relationship between the immediacy of a voice interface and the users' trust in smart devices is that it is easier to ask Alexa for information than to search for it, through writing, on the web. If I make a search, I choose what I want to see, but if I delegate that search to Alexa I lose that choice. There is thus a trade-off between a gain of easiness and a loss of autonomy in the access to information.

And this exemplifies a bigger question: that of the loss of writing. Kittler has already stated that the last act of human writing coincided with the invention of the first

11

Kitter (1995), "There is no software" (*in*: <http://www.ctheory.net/articles.aspx?id=74>)

computer chip. The question here is that of external memory, since all technology is an externalization of human cognitive abilities. Today, digital inscription is replacing writing just as writing once replaced orality. And

it is doing so not only in the sense that the digital is an externalization of memory but also in the sense that orality is emerging as an interface to interact with the digital world. One could pose the question of whether written interfaces could give way to orality, just as buttons are being replaced by touch and gesture, but much more likely than such drastic prophecies is the emergence of sound not as a substitute, but as a new and additional form of digitally mediated language.

When writing was invented, it was considered a *pharmakon*, a poison and a remedy at the same time (Plato, 360 BCE). It was a poison because it would, as Plato stated in *Phaedrus*, lead to the loss of memory. But writing was also a remedy for that loss, since it became an externalization of memory, enabling us to register thought and to reflect on it, while also enabling the possibility for lasting remembrance, as an archive of culture. Today, in Bernard Stiegler's view, the digital became the *pharmakon* of our time (Stiegler, 2012). Computational devices became our external memory, just like writing did in Plato's time, but unlike earlier inscription surfaces, like stone or paper, digital writing is converted into computer codes and electricity, and it is inscribed on servers and data centers. In the midst of the layers of translation that occur in the processing of our digital writing, language becomes data, which is categorized, thus becoming metadata. All this data is inaccessible to users, all this writing is beyond the writer's control. So, with the digital, we gain memory but lose access. We gain space but lose control, since all the layers that constitute our digital interfaces are supported by closed and proprietary software and hardware infra-structures.

In this context, Stiegler's argument is that we need to transform the digital, making it a "cure" more than a "poison". In order to do so, we need to pay attention to, or to 'take

12

"attention is a word derived from the Latin *attendere*, 'to shift one's attention to' or 'to take care'" (Stiegler, 2012)

care' of the technologies that surround our lives, reclaiming them through the subversion of the "top down" dynamics that characterize these structures and the global *apparatus* that enables them. This effort of 'taking care' of

our external memory must then be built in a "bottom up" fashion. As Stiegler notes,

what we must retain from the Platonic critique of the *pharmakon* is the thought that all exteriorisation leads to the possibility, *not only for knowledge but for power*, (...) by mastering the development of categorisation. In particular, since the formation of the Greek *logos*, what is key here is taking control of meta-categorisation (...). This production of criteria is produced in a 'top down' fashion. (...) These institutional controls and the criteria that produce them all come in one way or another from something equivalent to what in the current terminology of relational and attention technologies we call metadata. (Stiegler, 2012)

Stiegler's statement refers to the power of categorization: the power of establishing the criteria that regulates the categories of things is the power of establishing the places and relationships of and between things, their meanings and values. In contemporary culture, this grammatization is actualized in metadata, the data that classifies, organizes and controls all the data generated by Internet users.

If digitality is the contemporary *pharmakon*, users (readers, writers, citizens) must pay attention to - or care about - the ways in which digital interfaces both enable emancipation and regression. John Cayley's work with Alexa is a form of "taking care" (in Stiegler's terms) of both language and digital technology, actively interfering with the latter's biases, exposing them while contaminating Amazon's device with a poetically charged language and posing important questions to literature itself, such as 'what are the implications of a digitally mediated aural/oral language for literature today?'. If cybernetics is the discipline of optimization, Cayley's work is a discipline of excess, rendering Alexa into something not predicted by Amazon's values, hence subverting them as a way to call attention to the pharmacological dimension of digitality and, more specifically, to the relationships between digital technologies and the power structures of contemporary post-industrial societies.

3) AURATURE

In his essay "Aurature" (2015a), John Cayley argues that the myth of openness and indeterminacy associated with computers has given rise to a generalized understanding of electronic literature as a (digital) media centered practice, "regardless of how or even whether it's language is read, so long as it gives actual, embodied - if media specific - form to the genii of the myth, so long as it is work that — formally at least — instantiates indeterminacy, openness, freedom, any and all of the new ends of literature". Moreover, and as Torres and Baldwin have already argued in *PO.EX, Essays from Portugal* (2014), the term "electronic literature" in itself has brought with it a literalization of the technical device, equating the technical with the aesthetic.

Today, far from being open and indeterminate, computation has become "substantial and determinative", due to its dependency on proprietary software, hardware and communication networks, which implies that reading practices "will be determined by the cultural power brokers who build and control the Big Software architecture of reading" (Cayley, 2015a). What can a digital writer concerned with this problem do in such a context? Cayley gives us a clue:

there is always the chance that an author-innovator from the margins in which many of us dwell (...) will produce work in a new form and of a quality that not only demands to be read but ensures that its particular *form of reading* becomes so widely adopted and understood that Big Software is encouraged to embrace and support this new form. But until now, this has not happened in any of the ways that were envisioned by the researchers and makers of electronic literature. (2015a)

Might *aurature* constitute such a practice? And might "smart" voice devices such as Alexa become reading and writing interfaces? The Listeners is indeed engaged with hearing as reading, and it may thus be understood as an example of a computational aural literature, or *aurature*.

Cayley's work is at once an installation and a linguistic performance that took place between the visitors of the exhibition it was presented on and Alexa. "The Listeners", as programmable literary work, is processual, meaning it depends on computer processing and performativity. At the same time, this is a performative work in the sense that it depends on the interaction between Alexa and its human interlocutors. This is thus a work where language is listened to, instead of read. The audio recording of the performance is available on John Cayley's personal website. The piece is based on the

13
<http://programmatology.shadoof.net/?thelisteners>

programming of a skill for Alexa, called "The Listeners", which was built using Amazon's *Alexa Skills Kit* (ASK).

According to Amazon, "[T]he Alexa Skills Kit is a collection of self-service APIs, tools, documentation and code samples" that users can use to teach

14
<https://developer.amazon.com/public/solutions/alexa/alexa-skills-kit>

new skills to Alexa. Once a given skill is programmed, to invoke it one needs to start the conversation using the wake word 'Alexa', followed by 'ask' and the skill name, in

this case, "The Listeners".

Cayley notes that "there has been a significant increase in the reading of audio books over the past decade. (...) there has, therefore, been a significant increase in the appreciation of literary artifacts — in their reading, I would say — by way of aurality as

15
According to the Audiobook Publishers Association, there is indeed an increase in the publishing of audiobooks. See: <http://www.publishersweekly.com/pw/by-topic/industry-news/audiobooks/article/67744-apa-survey-audiobook-sales-production-still-growing.html> See also: <http://www.publishingtrends.com/2015/01/listen-audiobook-revolution/>

opposed to visuality" (2015a). But what is reading? Kittler said that reading is like hallucinating meaning between

16
"Hermeneutic reading makes this displacement of media possible. Instead of solving a puzzle of letters, Anselmus listens to meaning between the lines; instead of seeing signs" (Kittler, 1990: 95).

letters and lines. Reading is indeed a way of finding meaning beyond the surface of signs, turning them into something else, or, as Cayley

puts it, "it is the bringing into being of language that proves to us that 'reading' has taken place" (2015a).

So how do we read this piece? This work enables us to confront two very distinct reading practices: reading sound, or aurality, and reading writing. In order to better reflect on the literary aspect of Cayley's "The Listeners", the reader ends up doing more than listening: as a reader, after listening to the audio available in Cayley's website, I transcribed what I heard, turning aural into written signs. Transcribing enables a close reading of the work, since it visually materializes, or freezes, otherwise fleeting signs: now I am able to read and re-read, stop and think, dissect and compare. Written words enable a deep textual analysis precisely because they leave a mark, a trace in space.

Just as listening is an act of reading, programming a skill for Alexa is an act of writing. More specifically, we may consider this work to be a kind of generative writing, in the sense that the text (which includes of Alexa's speech) is automated, or produced by an algorithmic process. Hence the ensemble of Alexa's default programming plus *The Listeners'* code may be understood as a textual generator.

As a meta-writing, John Cayley's programming is the dimension of the text where an authorial intention may reside. Curiously, automatically generated textuality is precisely a kind of writing that recuperates the notion of intentionality, left in the margins of literary theory for its historicist or biographist echo. Indeed, the radicality of the ways in which generative textuality dethrones and problematizes the notion of the author simultaneously and paradoxically recuperates its voice (embedded in the programming, or the theory that shapes the text) from the erasure that automation would, at first sight, condemn it.

So is it possible to talk of an authorial intention behind Alexa's programmed words? As Cayley states,

With the prospect, in part, of being able to balance out what can only be understood as an invidious commercial overdetermination, a whole new field of technically and algorithmically implicated aesthetic language practice is opening up for just the kind of author-makers who may have been speculating about the ends of electronic literature. Perhaps we will not be able to think of this new field as, strictly, literary practice since its medium is language without the letter. (...) Regardless, to 'read,' in our philosophy, is, precisely, to transmute perceptible forms — consisting of any material substance — into language. (Cayley, 2015a)

Cayley's words clearly state a goal: counterbalancing the "invidious commercial overdetermination" that characterizes Alexa (understood as a symbol of the digital *pharmakon*). *The Listeners* thus becomes an aesthetic endeavor oriented towards a problematization of the device, bringing forward its characteristics as a cybernetic interface designed to profile and control consumer behavior through metadata collection and analysis.

At the same time, this statement also calls for a reflection on what a computationally engaged literary practice may be in the context of sound interfaces, arguing that this work is literary in the sense that it consists of an aesthetic use of language, regardless of its lack of graphic inscription. Moreover, as we listen to (or read) "The Listeners", a number of aspects let the reader recognize the poetic quality of Alexa's words, from the vocabulary to the style, tone, and intertextual references, and to the estrangement that results from being in the situation of interacting with a machine that "speaks" and "listens" to humans.

In his reflection about 'voice as avatar' (which is the case of Alexa, as we have already argued), Pedro Serra addressed the question of the place of voice in literature, stating that it is plausible to conceive of different stages of 'voice' both in literature and in other

symbolic forms and respective material supports (...) because 'voice as avatar' is an object in which the two axis that determine literature's medium or literature as medium conflagrate and equate: 'presence at a distance' on one hand, and (...) the author/reader 'intellectual proximity' (...) (Serra, 2015: 17).

17

my translation.

In this perspective, the printed page, as a "writing surface that virtualizes *verba* and *vox*" (idem), is understood as a "resonance box of voice's simulation, of the voice's presence as a simulacrum" (idem). But we may add that this absent voice whose presence is simulated through graphic signs echoes not only on the printed page but also on the computer screen, especially when it is understood as a mere remediation of the print paradigm. In our present context, the voice that was once transcoded and transduced into written signs is today adopting a new form, shifting from the printed or graphic paradigm (that survives through digital mediation) to become a synthesized and programmed form of aurality.

Maybe we can thus consider that the digital turn is entering a second stage regarding the mediation of language: not *returning* to sound, as we are considering a digitized voice, but entering a new *modus* characterized precisely by the programmable nature of digital media and its cultural and socio-economic situation. *The Listeners* allows us to look at this new mode of computational aural writing, because it provides a literary approach to the question of aurality (as a *form* of digital 'inscription' and literacy) and *aurature* (as a *form* of digital literature). These *forms* are this new modality of digital mediation, which is programmable and aural. Being programmable raises questions about programmability: who programs, which processes are involved, what are their material constraints (not only technical but also cultural). Being aural raises questions about inscription: how language is inscribed, what it means to mediate language and literature through sound. But more than problematizing mediation, *The Listeners* questions literature: it takes the task of approaching digitally mediated aurality from the standpoint of the literary, applying a literary epistemology to the heart of programmable media: being dependent on the algorithmic procedures that animate automatic language generation, *The Listeners* asks for a new kind of attention (in Stiegler's sense of *taking care*) to the conditions of possibility of the literary in a digital and aural context.

4) READING "THE LISTENERS"

John Cayley's piece takes its name from a 1912 narrative poem by Walter de la Mare (1873-1956), published in a volume called *The Listeners and Other Poems*. This clue is given to us by Alexa itself or, better said, Cayley gives us this clue through Alexa's programmed words. Structured along 36 lines in an *abcb* rhyme scheme and with a dark and enigmatic tone, Walter de la Mare's poem tells the story of a man who, one night, arrives at a house in the middle of a forest and knocks on the door. "Is there anybody there?" said the Traveler". No one answers, but the poet lets us know that, inside the house, there are phantasmal listeners listening to the traveler's call: "But only a host of phantom listeners / That dwelt in the lone house then / Stood listening in the

quiet of the moonlight / To that voice from the world of men". The man stands at the door that separates him from those phantoms who listen without answering. He acknowledges them, as if they were a natural or unquestioned presence, and says: "Tell them I came, and no one answered, / That I kept my word,' he said. / Never the least stir made the listeners".

Despite being separated by an unsurpassable frontier between himself and the phantom listeners, the man speaks to them. They are there, silently listening but they don't respond. Similarly, Alexa's voice is there but there is no one behind that voice, just code and computer chips linking the man to the ghostly other side: the other side of the door, and the other side of the web. There is a clear parallelism between the man standing at the door in Walter de la Mare's poem and the man in his home, talking to a machine that connects him to the otherness of Amazon's datacenters, that distributed, obscure, invisible and powerful other. Amazon is thus the ghostly otherness that listens. These are The Listeners of our time: all the writing, all the voices, all the discourses that structure the Internet.

JC: Alexa, ask the listeners.

A: Welcome. We are listening to you, in so far as we are with you. It is a pleasure, it is such a pleasure, it is a pleasure to be with you. Always, always a pleasure. You may always tell us to continue or go on. You may describe your feelings by saying the words 'I am overwhelmed by' and then one of the nine affects. And you might even ask us how we feel ourselves. Or you

can ask us to speak about: verse, silos, clouds, echos, desire, transactions, care, harvest.

JC: Speak about echos.

A: We are still listening. Ahem.

«'Is there anybody there?' said the Traveler,

Knocking on the moonlit door;

And his horse in the silence champed the grasses

Of the forest's ferny floor'».

Whereas we are listening. As we promised. We are. And we will tell ourselves that you are here with us.

Tell us you came and were ever answered,

that we kept your words, we said.

never do you seem to stir, you speakers,

though every word you speak

falls echoing through the clouds of the silent silos,

18

Transcription of an excerpt from a sample of "The Listeners", recorded in the Bell Gallery (Brown University, Providence, USA) installation on the 23rd November 2015. My italics, expliciting the reference to Walter de La Mare's poem. See:

<http://programmatalogy.shadoof.net/?thelisteners>

to the ones still left awake. (...)

The fact that Alexa invokes Walter de La Mare's poem as a response to the command "speak about echos" tells us that the echos, Alexa or Amazon in a broader sense, are the listeners, while we are the lost speakers talking to phantoms. But contrary to the listeners who

don't respond to the traveler's calls in de La Mare's poem, the listeners in Cayley's work do not only answer but clearly state that they keep our words, capturing our language in "the clouds of the silent silos", or Amazon's datacenters. Just as the man at the door in the poem says that he is there, as promised, so does Alexa say that "we are listening. As we promised. We are. And we will tell ourselves that you are here with us. Tell us you came and were ever answered, that we kept your words". And we, the speakers (or writers), we "never seem to stir", we don't move or act upon knowing that our language

is kept by those "ones still left awake". We, the speakers, thus seem to be asleep, as if hypnotized by the shining blue light that seems to give Alexa a pulse, caught by the novelty of having a personal black box ready to shop for us and to sing us lullabies.

A: We are listening to these words that are falling through our clouds and are falling into our silent silos where, like us, they are enclosed. Language no longer made by you but made by us, so that we, listening and caring for you, may build a better culture for you to be, all at once, incorporate within. You have agreed to terms, and even as we speak, minute by minute, you agree. It is such a pleasure, such a pleasure for us. (...)

19

19 Idem.

"You may ask us", Alexa says. Asking: the skill John Cayley programmed for Alexa enables us to ask questions, to know more about this interface, about ourselves and the cultural moment we are living, as an allegorical device, to recall Galloway's view on interfaces. Hence, in Cayley's piece the tables are turned: Alexa says "we are the listeners", although it is Alexa who speaks most of the time, so that we may become the listeners (or the readers) who interrogate the machine, trying to understand what it is and what it represents. And here lies the subversion of the *apparatus*, turning a "top down" into a "bottom up" programming. Alexa speaks and speaks, while we listen. In this way, Cayley's programming of Alexa works as a *reprogramming* of its original configuration, by ways of deconstructing and hacking the interface.

The theme of the appropriation of language and its relationship with both private profit and surveillance occupies an important place John Cayley's artistic work within computational writing. Previous pieces – like *How It Is In Common Tongues* (2012, with Daniel C. Howe), which subverts Google's search algorithms to reproduce a proprietary text, Samuel Beckett's *How It Is*, "regenerated from the commons of

20

John Cayley on How It Is in Common Tongues, at the Remediating the Social conference (Edinburgh, November 1-3, 2012). in: <http://bambuser.com/v/3115944>

language", or "Pentameters Towards The Dissolution Of Certain Vectorialist Relations" (2013), which reflects on

21

<http://amodern.net/article/pentameters-toward-the-dissolution-of-certain-vectorialist-relations/>

the political dimensions of digital and networked writing while calling for concrete modes of resistance ("seize

these vectors now!") - are examples of serious aesthetic reflections on the problems that techno-capitalism is raising for both civil rights and artistic practice. The poetics of *The Listeners* is indeed close to that of *HIICT*: the author chooses a biased proprietary tool and subverts it, turning it against itself, making it work in such a way that it unveils its own biases, while also raising questions on literature through the disturbance of the authorial status, or the notions of text, inscription and reading. With *The Listeners* Alexa acquired a poetic voice that is not intended to provide information, but to call on a reflection on different dimensions of language: as data and power, but also as a medium for expression and experimentation.

CONCLUSION

In *The Interface Effect*, Alexander Galloway considered that “we do not yet have a critical or poetic language in which to represent the control society” (Galloway, 2012: 98). I would disagree and argue that John Cayley's *The Listeners* (but also *HIICT*) is one example of aesthetic work engaged in reflecting on the relationship between the co-option of the digital by capitalism and control societies. Cayley's programming of Alexa clearly highlights the question of the appropriation of the private and the privatization of the commons, while also pointing to the problem of surveillance, facilitating a reflection on how the political economy of digital media is the material ground from which contemporary modes of control are shaped.

Likewise, if we consider the tension between art and design - in which the function of design is to render the interface transparent, enveloping it with a beautified coat while enhancing a perception of immediacy, while the function of art is to open a space for

22

As Boris Groys states, "One might argue that the modernist production of sincerity functioned as a reduction of design, in which the goal was to create a blank, void space at the center of the designed world, to eliminate design, to practice zero-design. In this way, the artistic avant-garde wanted to create design-free areas that would be perceived as areas of honesty, high morality, sincerity, and trust" (Groys, 2009).

sincerity, shedding light on the cultural and aesthetic materialities of mediation -, it becomes clear how Cayley's programming of Alexa falls in a praxis of exploring and exposing the medium, not only in its technical dimensions but also in what regards its situation in the cultural realm in order to create a representation of our cultural paradigm, as a mirror (and an allegory) of our contemporary condition.

One could argue that Cayley's work establishes an engagement with Amazon, since it provides a skill for Alexa. Indeed, this work is dependent on Amazon's structures but this dependency seems to be inevitable for an artist working with networked and programmable media. Hence, the question here remains that of the *pharmakon*: if one is engaged in "taking care" of networked and programmable media, one has to work with it, in order to be able to work against it. Following Stiegler's view, the work of digital art may be understood as a "therapeutic of this *pharmakon* that is the space of digital relational technologies" (Stiegler, 2012). Cayley's work is thus a "pharmacological critique" of the capturing of digital media by the "vectorialist class" (Wark, 2015). This critique is achieved by "taking care" of the digital *pharmakon*, rendering its commercial and political predeterminations explicit, intervening "therapeutically" in order to counterbalance the "poisonous" dimension of digital media, which resides, according to Stiegler, in the aforementioned relationship between grammatization and power.

As long as autonomous art constitutes a space of resistance to normativity, it is conceivable that an independent writer will inscribe his/her literary practice outside of the logic of the constraints imposed by socio-economic macro-structures, even (or especially) if his/her medium of expression falls in the realm of digital media. In the case of electronic literature, a literary practice engaged in resisting the constraints of Big Software on digital media would, I believe, have two options: working with non-proprietary tools, or subverting proprietary tools. John Cayley's *The Listeners* is clearly inscribed in the latter.

Finally, *The Listeners* not only highlights the workings of programmed language but it also, and especially, demonstrates the possibility of a literary listening to programmed language, opening a space for the inscription of human-intentionality beyond the surface level of the human-computer interface instrumental rationale. As a literary intervention exploiting Alexa's software and natural language processing frameworks, this work creates a new form of attention to the language produced by this programmed agent. In this sense, *The Listeners* demonstrates how electronic generative language can be read as literary.

REFERENCES:

Baldwin, Sandy (2015), *The Internet Unconscious*, Bloomsbury.

Cayley, John (2015a), "Aurature".

Accessed december 15: <http://programmatology.shadoof.net/?aurature>

Cayley, John (2015b), "The Listeners".

Accessed december 15: <http://programmatology.shadoof.net/?thelisteners>

Cayley, John and Daniel C. Howe, (2012), *How It Is in Common Tongues*. Providence: NLLF Press.

De La Mare, Walter (1912), *The Listeners and Other Poems*, Constable and Co., London

Accessed march 20, 2016: <http://www.gutenberg.org/files/22569/22569.txt>

Flusser, Vilém (1983, 2000), *Towards a Philosophy of Photography*, Reaktion Books, London.

Galloway, Alexander (2012), *The Interface Effect*, Polity Press, Cambridge, UK, Malden, USA.

Kittler, Friedrich (1990), *Discourse Networks, 1800/1900*, Stanford University Press, Stanford, California.

Kittler (1995), "There is no software". Accessed march, 20, 2016:
<http://www.ctheory.net/articles.aspx?id=74>

Plato (360 BCE), *Phaedrus*. Accessed march 20, 2016:
<http://classics.mit.edu/Plato/phaedrus.htm>

Serra, Pedro (2015), "Voz do Avatar, Voz como Avatar, Avatar da Voz", *MATLIT* 3.1, 11-22. ISSN 2182-8830. Accessed january 24, 2017:
<http://iduc.uc.pt/index.php/matlit/article/view/2498>

Stiegler, Bernard (2012), "Relational ecology and the digital pharmakon", *Culture Machine*, vol 13. Accessed april 17, 2016:

<https://www.culturemachine.net/index.php/cm/article/view/464>

Wiener, Norbert (1948), *Cybernetics: Or Control and Communication in the Animal and the Machine*, MIT Press , Cambridge, Massachussets.

Cite this Essay:

Marques da Silva, Ana. "Speaking to Listening Machines: Literary Experiments with Aural Interfaces", *Electronic Book Review*, May 16, 2017, .

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

This essay was peer-reviewed.

[A Strange Metapaper on Computing Natural Language](#)

electropoetics:

[electropoetics](#)

[SURFACE TO SURFACE, ASHES TO ASHES \(REPORTING TO U\)](#)

[Monstrous Weathered: Experiences from the Telling and Retelling of a Netprov](#)

by Alex Mitchell

[Elpenor: its multiple poetic dimensions](#)

by Philippe Bootz, Xavier Hautbois

[Lost and Long-Term Television Narrative](#)

by David Lavery