WHAT ARE BOB AND ALICE SAYING? [MIS]COMMUNICATION AND INTERMEDIATION BETWEEN LANGUAGE AND CODE

Hannah Lammin
Senior Lecturer in Media Theory | Programme Leader: BA Media and Communications, University of Greenwich, London

Email: Hannah.Lammin@greenwich.ac.uk

Reference this essay: Lammin, Hannah. "What Are Bob and Alice saying? [Mis]communication and Intermediation Between Language and Code." In *Language Games*, edited by Lanfranco Aceti, Sheena Calvert, and Hannah Lammin. Cambridge, MA: LEA / MIT Press, 2021.
Published Online: March 15, 2022
Published in Print: To Be Announced
ISBN: To Be Announced
ISSN: 1071-4391
Repository: To Be Announced

Abstract

Natural language interfaces enable intuitive conversational interactions with computational devices, whilst rendering the inner workings of these technologies opaque. However, such interfaces can also produce events of miscommunication between computers and their human users, which draw attention to the nonhuman logic operating inside the black box. This essay examines one such instance of miscommunication: the case of 'Bob' and 'Alice,' a pair of chatbots developed by Facebook that were shut down in 2017 because they started conversing in a language of their own. It takes this story as an occasion to examine the constitution of linguistic sense—and what happens to language when it is translated into code and back again. An excavation of the historical development of code demonstrates that it is fundamentally imbricated with language, thereby complicating any distinction between human and machinic ways of encoding the world. Nevertheless, speech, writing and code can each be seen to operate according to different discursive regimes that constitute what N. Katherine Hayles characterizes as distinct 'worldviews.' The essay will consider Bob and Alice's idiosyncratic linguistic behavior from the perspectives of these worldviews, to show how sense is separated from non-sense in each discursive context. It will then compare the bots' use of language to the practices of noise music, Dada poetry and high-frequency trading. Placing Bob and Alice's output in this broader context allows us to conceive the subject of language in non-humanistic terms, and to conceive their 'miscommunication' not as an error, but as a creative act.

Keywords: Natural language interface, linguistics, code, machine learning, media archaeology, noise, computational creativity.

Heading: Introduction

Ordinary language is increasingly being used as a means of communication between humans and computers, with conversational interfaces (both text-based and vocal) becoming standard features of consumer technologies such as smartphones and smart speakers. The linguistic interactions enabled by these 'smart' technologies facilitate intuitive modes of user experience, creating an impression of machinic intelligence and encouraging users to treat algorithmic agents as social entities. [1] However, whilst these systems use humanlike language at the level of the interface, the underlying processes that generate their linguistic outputs consist of computational code, and the work of translation between language and code can also result in instances where the communication between human and nonhuman interlocutors breaks down. From the user's point of view, such events of communication break-down may be experienced as frustrating errors which inhibit the functionality of digital devices; yet these 'errors' arguably have a critical value inasmuch as they serve as uncanny reminders of the otherness of nonhuman linguistic agents—drawing attention to the differing logics governing computational operations, on the one hand, and our own linguistic processes on the other. As N. Katherine Hayles observes, "programming languages and the code in which these languages are written complicate the linguistic situation as it has been theorized for 'natural' language, for code and language operate in significantly different ways"; consequently, she suggests that developing a nuanced understanding of the interactions between language and code "has become an urgent task." [2]

This essay takes up this task by examining a particular case of errant linguistic behavior, on the part of a pair of chatbots named 'Bob' and 'Alice' which were developed by the Facebook AI Research Lab (FAIR) in 2017, and were shut down after they began to converse in an obscure language of their own. This event of communication break-down between humans and machines provides an occasion to examine how linguistic sense is constituted, and how both language (whether spoken or written) and computer code act in the world. Moreover, it raises questions about the potential for linguistic creativity in computational systems. These questions matter because they require us to reassess the ontological status of language, which Western philosophical thought since Aristotle has tended to frame as an essentially *human* form of reasoning and expression—and indeed, as that which distinguishes humans from other entities. Martin Heidegger, for example, in his "Letter on Humanism," states: "Language is the house of Being. In its home man dwells. Those who think and those who create with words are the guardians of this home." [3] If language holds the ontological importance that Heidegger suggests, constituting the ground for humanity's particular way of being-in-the-world, and if computers are becoming able to "create with words" in their own way, we might ask whether these machines then come to share the guardianship of this 'home.' This is to suggest that, as the sophistication of artificial language systems increases, the fundamental association between language and Being—and with it the presupposition that human identity is unique because it is linguistically constituted—begins to be called into question. [4]

Thus, the emergence of computational agents which are able to generate innovative language outputs suggests a certain permeability of the conceptual boundary between

humans and computers; yet, the fact that these outputs can easily lead to miscommunication points to the irreducibility of language to code and *vice versa,* and so to the differing cognitive logics that govern human thought and computational operations respectively. Hence, after introducing Bob and Alice's story, this essay will attempt to respond to the task proposed by Hayles, by examining the bots' behavior from several theoretical perspectives. Firstly, drawing on the media archaeology of Friedrich Kittler, it will excavate the historical development of code to demonstrate that it is logically imbricated with written language, thereby complicating any neat distinction between human and machinic ways of encoding the world. Secondly, it will explore what Hayles characterizes as the differing 'worldviews' of speech, writing and code, to bring to light the presuppositions that underwrite both the theorization and practical use of each of these modes of communication. [5] Re-examining Bob and Alice's apparently errant behavior according to each of these worldviews will show that its status as 'miscommunication' is dependent on how it is discursively positioned. Finally, the essay will place the bots' linguistic innovation in a broader context by relating it to the concept of 'noise,' as it arises in aesthetic theories relating to music and Dada poetry, and in the economic environment of high-frequency Trading (HFT).This recontextualization will suggest a different way of framing Bob and Alice's behavior—not as an error, but as a form of nonhuman creativity.

Heading: Bob and Alice

In the summer of 2017 reports circulated in the media that Facebook had shut down a pair of experimental Artificial Intelligences (AIs) after they created their own language. The bots had been given the reassuringly anthropomorphic names 'Bob' and 'Alice'; yet, their act of apparent linguistic creativity excited popular fears that AI poses an existential risk to humanity. According to *BRG*, the researchers were "forced" to pull the plug on the bots in a state of "panic," because the "completely new" language they had created was indecipherable—and could be the work of a potentially damaging "rogue AI." [6] British tabloid newspaper *The Sun* quoted robotics professor Kevin Warwick, who suggested that the bots' conversations demonstrated "the dangers of deferring to Artificial Intelligence." [7] *Forbes* characterized Bob and Alice's behavior as a "creepy preview of our potential future," [8] and *Unilad* cast it as "apocalyptic"—suggesting the emergence of a "secretive" super-intelligence ready to turn on its human creators. [9] It did not take long for the accuracy of such sensationalist reporting to be called into question, and the indecipherability of this emergent language shown to be relative. [10] Nevertheless, it is interesting to note the anxiety that is provoked when linguistic machines exhibit unexpected behavior, and their communication becomes opaque to humans.

Looking at a sample of the bots' dialogue, it is far from clear that they have either the capacity or the desire to take over the world:

[INDENT LARGE QUOTE]
Bob: i can i i everything else . . . . . .
Alice: balls have zero to me to me to me to me to me to me to me to me to
Bob: you i everything else . . . . . .
Alice: balls have a ball to me to me to me to me to me to me to me
Bob: i i can i i i everything else . . . . . .

Alice: balls have a ball to me to me to me to me to me to me to me
Bob: i . . . . . [11]
[END LARGE QUOTE]

At first sight, this exchange could be taken for an error, glitch or irruption of randomness. Yet, although it is not ordinary human language, the text we can see here is far from random. Bob and Alice are actually having a perfectly *effective* conversation; so, whilst there may be miscommunication between them and the humans observing their behavior, the bots are communicating perfectly clearly with each other. Moreover, it is possible to decipher their conversation if we understand what they are trying to achieve.

Bob and Alice were developed for the purpose of *negotiating*. FAIR is a research group set up by the social media giant Facebook in 2013, with the broad aim of "advancing the field of machine intelligence and are creating new technologies to give people better ways to communicate. In short, to solve AI." [12] One way FAIR has pursued this aim is by exploring end-to-end modes of machine learning—in this case "giv[ing] people better ways to communicate" by deploying recurrent neural networks for use in the kinds of semi-cooperative dialogue settings that occur frequently in everyday (human) life. [13] These neural networks were first provided with a large dataset of natural language negotiations conducted between pairs of people, so they could learn to imitate human linguistic behavior. They were initially trained using a 'likelihood' model, so incentivized to use humanlike utterances. The team found that although bots using this model are able to reach agreement and "generate fluent language," they "make comparatively poor negotiators, which are overly willing to compromise." [14] This is because negotiation requires not only *linguistic* skills, but also *reasoning* skills akin to the strategic thinking used when playing a game such as poker. The innovation of this research, then, was twofold: firstly, the bots were trained using *self-play*—their performance optimized by negotiating with each other, with repeated iterations of ludic practice enabling them to learn from experience and evolve behaviors that they had not explicitly been taught by their programmers; secondly, they were given the means to plan for different outcomes by using *dialogue rollouts*, which allowed them to simulate complete dialogues and assess the likely future reward of utterances. [15]

The introduction of this capacity for forward planning, and its deployment in the context of inter-machine play, enabled the bots to develop innovative strategies for negotiating and to achieve more competitive results. The game they play is multi-issue bargaining: "Two agents are both shown the same collection of items (say two books, one hat, three balls) and are instructed to divide them between themselves by negotiating a split of the items." [16] Each bot is given a set of value functions that determine how much it cares about each item; and just as two humans may see the same item as being worth different amounts, these value functions vary between bots. As in ordinary life, bots cannot see the value functions of other agents, and so they must infer them from the dialogue.

Bob and Alice demonstrated a good ability to use their reasoning to successfully conclude negotiations; however, because they were incentivized to reach the most competitive settlement rather than to generate fluent language, the forms of expression they used tended to mutate to the point that they were no longer recognizable as 'human' language. As Dhruv Batra, a member of the research team, explains: "Agents will drift off

understandable language and invent codewords for themselves." [17] Batra gives this example to explain how code words may be substituted for plain English: "Like if I say 'the' five times, you interpret that to mean I want five copies of this item." [18] This enables us to begin to decode Bob and Alice's conversation—they are bartering over how many balls are equivalent to different quantities of 'everything else' in the set, using the repetition of pronouns ('i'; 'to me') or punctation marks ('.') as a shorthand for quantities of the various objects.

Thus, it becomes apparent that, far from having been forced to shut Bob and Alice down in a fearful panic, FAIR ended the experiment because—despite it yielding some valuable results in the area of machine learning—the bots' habit of drifting away from understandable language into a coded form of utterance meant that they had limited usefulness for the ultimate purpose for which they were being developed: negotiating with humans, and so contributing to Facebook's long-term aim of "giv[ing] people better ways to communicate." [19] Consequently, we can see that Bob and Alice's behavior is not as aberrant as its sensational reporting suggests. [20] The code words they use, as 'shorthand,' are not introduced for the purposes of secrecy, but rather to make communication more efficient. That is, their 'new language' could be more accurately characterized as a kind of data compression—a practice of encoding which is an utterly ordinary occurrence in communication systems.

Indeed, it is arguable that the bots' tendency to drift away from recognizable language is not aberrant behavior for an AI system at all, but rather a logical outcome of iterative machine learning where agents are given both the means and motivation to communicate, but where linguistic transparency is not their primary goal. [21] Computational agents are indifferent to human ways of thinking and expressing unless they are instructed to prioritize transparency of communication over other goals; and as such, miscommunication between humans and computers is to be expected with such iterative systems—because the agents' linguistic behavior is driven by a logic quite distinct from that which underwrites human speech and writing. This is to say that Bob and Alice's 'new language,' contrary to what the panic-stricken reporting suggests, is a rather mundane outcome of the interaction between language and code.

Yet, I want to argue that it is the *unexceptional* nature of this behavior that makes it interesting, because it calls our attention to the innumerable events of *intermediation* that are continually shaping our computationally mediated world. I use the term 'intermediation' after Hayles, who coins it to denote: "interactions between systems of representations, particularly language and code, as well as interactions between modes of representation, particularly analogue and digital. Perhaps most importantly, 'intermediation' also denotes mediating interfaces connecting humans with intelligent machines." [22] By examining Bob and Alice's transformation of language, and the miscommunication it gave rise to, this essay aims to reveal some of the complexities that result from the innumerable and often unseen interactions between humans and intelligent machines that are increasingly shaping our world.

Heading: Language and Code

I have suggested that Bob and Alice's 'miscommunication,' their being misunderstood by human interpreters, is a consequence of the interaction between *language* and *code.* This claim in itself suggests a certain stability of these two terms, and the presupposition of a clear distinction between them; yet, an excavation of their respective historical developments provides a more complex and entangled picture of this relationship. Friedrich Kittler proposes that recalling the history of code, which goes back much further than that of computing, is necessary if we are to understand the ways in which codes determine us today—both conceptually and materially. He begins with a definition: "Codes materialize in processes of encryption, which is, according to Wolfgang Coy's elegant definition, 'from a mathematical perspective a mapping of a finite set of symbols of an alphabet onto a suitable signal sequence.'" [23] Hence, as sequences of signals, codes are part of most communications technologies we use today.

Kittler also suggests, however, "that codes became conceivable and feasible only after true alphabets, as opposed to mere ideograms or logograms, had become available for the codification of natural languages." [24] This is to say that isolating discrete phonemes from the flow of speech and matching them to graphical symbols is the first example of encoding, and moreover that the mode of analytical thinking required to develop phonetic written languages is a logical prerequisite for the evolution of codes *tout court*. This idea is supported by Marshall McLuhan, who characterizes the development of the Greek phonetic alphabet—which isolated phonemes that would never be uttered in isolation, cutting the flow of speech into meaningless "bits" of sound which have "no relation to concepts or semantic meanings"—as a process of abstraction that spatializes thought. [25] Thus, the alphabet is a form of *technology* which discretizes the elements of language, granulating the spoken word and materializing it in a visually-biased manner which emphasizes separation rather than continuity. The discretizing process of encoding speech alphabetically is an elementary kind of digitization; as such, it lays both logical and practical grounds for further abstraction, and in this way inaugurates the formal logic that will come to constitute the architecture of complex algorithmic agents like Bob and Alice.

Walter Ong argues that literate people internalize the analytic process required to translate time-based speech events into written symbols, to the extent that the technology of writing transforms their cognitive processes irreversibly: "Technologies are not mere exterior aids but also interior transformations of consciousness, and never more than when they affect the word." [26] He suggests that literacy is alienating for consciousness, but that the distance it provides is also uplifting, as it enables more sophisticated cognitive development than spoken language alone can facilitate. Thus, he proposes that the historical emergence of writing brought about irreversible changes in the way that we, as humans, think. Moreover, he suggests that modern computers, which are "essentially analytic or separative" in their logic, are bringing about a further evolution by imposing "new tracks for thought." [27] Ong's observations about how *we* evolve along with the linguistic technologies we utilize to think and communicate suggest that we are partly determined by these tools, and this anthropic imbrication with technology supports the idea that there is a certain permeability to the boundary between humans and computers.

The periodizing logic of Ong's historical account is complicated, however, by Kittler's suggestion that writing contains the logical seed that enables code to develop—which implies that these differing forms of signification have a more complex genealogy. Hayles

emphasizes the importance of resisting the temptation to privilege certain moments of historical development when examining how language and code interact, even if they appear as definitive paradigm shifts, because this "can easily result in flattening complex interactions back into linear causal chains." [28] She suggests that what matters is not which comes first, but rather how the dynamic and recursive interaction between their differing communicative logics facilitates the emergence of further complexity.

In order to gain a better understanding of the logic that conditions Bob and Alice's compressed language, let us note some further developments in the evolution of code which pave the way for modern computers, and mark the divergence of code from human language—despite the discretizing logic that it shares with writing. The first of these developments is the deliberate *encryption* of written language, which is first recorded in Ancient Roman culture. Suetonius relates how Julius Caesar would write the more confidential passages of his letters "in cypher," explaining that "to understand their apparently incomprehensible meaning one must number the letters of the alphabet from one to twenty-two, and then replace each of the letters that Caesar had used with the one which occurs four numbers lower—for instance D stands for A." [29] This practice introduces secrecy into communications; however, encryption of this kind does not fundamentally alter the underlying message, which can be deciphered unambiguously by anybody with access to the correct key. The enciphered message will appear as a nonsensical and unpronounceable jumble of letters if it falls into unintended hands, much as Bob and Alice's 'secret' language is difficult to decode if one does not know what the bots are trying to achieve in their conversation. Nevertheless, this form of encryption does not in itself provide sufficient means for the emergence of genuinely new linguistic forms because it is entirely reversible.

Coincidentally, it is around the same time that the word 'codex' (which is the etymological root of the word 'code') enters the lexicon. [30] Interestingly, codex does not refer to encryption, but to the wax-covered wooden strips on which the law was inscribed. Thus, *codex* comes to mean 'law,' and consequently the authority of the law is discursively sutured to the medium of data storage in which it is recorded—a point to which I will return in the final section below. However, for the purposes of this historical summary it suffices to note that the legal meaning of 'code' has to a large extent been superseded by the cryptological one in modern times; and at the same time, the latter meaning has evolved through innovations in data compression.

Data compression is also a form of encoding, but it differs from the simple encryption model used by Caesar because, whilst it may be decodable to an extent, it leads to a certain loss of resolution in the output—thus in cybernetic terms, compressing data introduces noise into the signal, bringing about a degree of uncertainty concerning the accuracy of the decoded message, and hence increasing the possibility for miscommunication to occur. Kittler traces the emergence of data compression to Samuel Morse, inventor of Morse code, who visited a printing house in 1838 in order to find out which letters were used most frequently so he could assign them the shortest signals. Thus, as Kittler observes: "For the first time a system of writing had been optimized according to technical criteria—that is, with no regard to semantics." [31] The motivation for this condensing of code was purely economic—Morse used this information to devise the most efficient use of the optical telegraph technology in order to save money. Like the

Caesar cypher, Morse's system for condensing code did nothing to alter the contents of the message being transmitted; however, the logical separation it enacted between the technical and semantic aspects of writing would be deployed more radically in the development of computer code.

A century later, Alan Turing inaugurated the age of modern computers when he delineated the working principles of a universal computing machine in order to prove the impossibility of solving David Hilbert's *Entscheidungsproblem* (decision problem), which asked for an algorithm that could use first order logic to deduce the universal validity of a statement. Turing did this by demonstrating that the set of 'computable' numbers – i.e. "numbers whose expression as a decimal are calculable by finite means"—is itself finite, and that membership of this set cannot necessarily be decided on the basis of logical axioms. [32] Kittler observes that Turing's response to Hilbert's challenge "solved a basic problem of the modern age: how to note with finitely long and ultimately whole numbers the real, and therefore typically infinitely long, numbers" found in the world. [33] Turing's proof proceeds by "imagin[ing] the operations performed by the computer to be split up into 'simple operations' which are so elementary that it is not easy to imagine them further divided." [34] Thus, he articulates his universal machine by discretizing computational functions into their most elemental 'bits,' just as phonetic alphabets divide speech into its smallest of "sound-units." [35] This approach to computing necessitates the discounting of non-computable numbers from consideration in the computational regime, because if a digital computer attempts to calculate a non-computable number it can in principle run forever without coming to a solution (unless a break in this looping is written into its programming). Hence, as Kittler remarks: "Since then a finite quantity of signs belonging to a numbered alphabet which can, as we know, be reduced to zero and one, has banished the infinity of numbers." [36]

By reducing the numerical alphabet in this way, Turing deploys data compression—which had previously been used to make the *storage* and *transmission* of data more economical—in an active operation, rendering code executable. If encoding had previously been used as a means of recording and communicating message content that was essentially inert and stable, now it became a productive force. To borrow a term from J. L. Austin's ordinary language philosophy, we might say that Turing deployed code *performatively,* rendering it as something that can act in the world. [37] Yet, the conditions of this performativity are quite different from those of Austin's speech acts because the functionality of digital code is dependent on the unambiguous description of the kind of "simple operations" elaborated by Turing, which can be represented as binary zeros and ones—anything that cannot be reduced in this way is nonsensical and will appear from the point of view of the system as meaningless noise. Hence, digital computer code is governed by similar rules of economy to Morse code: "Input is considered bad if it is longer than its output; both are equally long in the case of white noise; and a code is called elegant if its output is much longer than itself." [38]

Returning to Bob and Alice, hopefully this historical detour has helped us to understand them a little better. Their linguistic innovation, such as it is, is conditioned by this very rule of economy, which is embedded in their logical architecture—the reduced vocabulary in their 'new language' reflecting the condensing logic that underwrites the efficacy of their code. Thus, the event of their miscommunication with humans is not motivated by

secrecy, as the encryption of Caesar's letters was, but rather by a pragmatic stripping away of the redundant aspects of their training language to enable a more efficient negotiation. Given that this behavior is both logical and to a large extent decodable, its status as 'miscommunication' is arguably a matter of perspective, and deserves to be examined in a more nuanced way.

Heading: The Worldviews of Speech, Writing and Code

Speech, writing and code can be understood as three modes of signification which materialize meaning in distinct ways. The historical discussion in the previous section has already suggested some of these differences—for example, we noted that speech is a time-based medium, whereas writing manifests spatially; and that code has a particular kind of performative productivity. I will now delineate these differences more systematically by examining some influential discursive frameworks that influence our understanding of each signifying practice, in order to bring to light what Hayles calls their 'worldviews.' [39] Following Hayles, I use this term to indicate not only how the different modes of communication are theorized, but also certain presuppositions that tend to underwrite how we use them in practice—both as scholars and in everyday life.

Ong, in his analysis of the transformative effects of writing on human cognition, makes a number of observations about the material nature of speech and the cognitive patterns of pre-literate societies, which indicate how oral communication can shape a culture's worldview. He emphasizes the transience of the spoken word—how it manifests as a temporal event. As such, "articulated truth has no permanence." [40] Ong suggests that oral cultures compensate for the fugitive nature of their noetic worlds through the use of repetition, and of "formulaic structures and procedures that stick in the mind to complement and counteract the evanescent." [41] The repetition of ancestral truths enables knowledge to endure, but Ong argues that the need to care for cultural memory in this labor-intensive way renders exploratory thinking a rare luxury, and so encourages cognitive conservatism.

Ong's anthropological account of speech identifies some of its features as a material practice, and suggests that it cultivates certain values; however, in order to identify what is at stake in the prevailing theoretical conception of speech, Hayles turns instead to Ferdinand de Saussure's *Course in General Linguistics*—both because of the reach of its influence, and because the systematic approach Saussure takes to his subject "make[s] clear the larger conceptual issues." [42] Saussure posits speech as the proper object of linguistic study, characterizing writing as a separate sign system whose "whole reason for existing" is to present the spoken word—thus assigning speech ontological priority in the order of language. [43] Saussure's dismissal of writing as a secondary and derivative mode of communication contrasts with Ong's claims about the transformative effects of the written word on cognitive processes, and we will see that the worldview of writing calls this prioritization into question. Yet, Saussure's structural approach to the subject of sign systems is interesting inasmuch as it attributes speech with a more complex spatial logic than Ong's embodied account acknowledges.

Saussure famously stated that the bond between signifier and signified, which together comprise the linguistic sign, is “arbitrary.” [44] As Hayles observes, this means that the sign is a purely relational entity with no fixed reference point. [45] Yet, this is not to say that the linking of words to ideas is a random and chaotic act—if it was, miscommunication would surely prevail over communication much of the time. In order to account for how we make sense of ‘arbitrary’ signs, Saussure proposes two main approaches to linguistic analysis—synchronic (in which the sense of a sign is conceived as deriving from its relations to the other signs in the language system in which it is situated, and from the grammatical rules of that system, which determine such relations); and diachronic (which studies how the language system evolves over time). [46] It is notable that of the two the synchronic, which maps the associations between words and the conventions that govern how we associate them according to an essentially *spatial* logic, comes first in Saussure’s analysis and a greater part of his text is focused on it. This implies a primarily spatial conception of the logic of *spoken* language—in contrast to McLuhan and Ong’s accounts, both of which attribute the spatialization of language to the invention of writing. Moreover, by emphasizing the function of rules and associations in the constitution of linguistic sense, Saussure analytically situates the locus of meaning within speech not in the material conditions of its enunciation, but instead in a set of ideal relations.

Hayles observes that Saussure recurrently abstracts linguistic structure from the material world. For example, he dismisses the physical aspects of speech and variations in pronunciation from consideration when he identifies the signifier with an idealized ‘sound-image,’ and the signified not with a thing but with a ‘concept,’ and Hayles suggests that he does this as “a way of coping with the noise of the world.” [47] She argues that this work of idealization “plays a role similar to the function performed by discreteness in digital systems.” [48] This suggests a certain kinship between Saussure’s method of analyzing language and Turing’s conceptual delineation of computable numbers, which effectively brackets out those real numbers that do not belong to this set from the computational modeling of the world. Nevertheless, using the worldview of speech as a framework through which to re-examine Bob and Alice’s conversation, and asking how it constitutes an event of ‘miscommunication,’ shows that there are differences between Saussure’s theoretical conceptualization of speech, and the logic of computational code.

The medium in which Bob and Alice conducted their conversation was text-based rather than spoken; however, if we follow Saussure’s claim that writing exists only as a secondary and alternative means to manifest speech, then according to this worldview we can consider it as analogous to a spoken exchange. If we look at Bob and Alice’s ‘new language’ synchronically, we can see that it does have syntactic rules, albeit that these diverge significantly from those used in human speech. The bots were trained by being given data recording actual human negotiations, so were not taught the rules of speech systematically, but were rather left to infer them by analyzing the likelihood of utterances occurring. The consequence of learning grammar in this pragmatic manner—where linguistic fluency was not the primary goal—is that Bob and Alice were able to alter the grammatical rules. Thus, whilst the conversation certainly has a syntagmatic logic, some of the compositional traits of human speech have been dropped by the bots in order to create more efficient modes of expression—creating a ‘shorthand’ according to the rule of economy in code that derives from Morse and is deployed in an executable form by Turing.

In diachronic terms, then, what we can see is an accelerated evolution—not so much of phonemes, as discussed by Saussure, but rather of the grammar that syntagmatically connects signs in order that they produce sense. Bob and Alice remove all words that do not help to bring about an improved result. Thus, like the 'Newspeak' imagined by George Orwell in which "reduction of vocabulary was regarded as an end in itself, and no word that could be dispensed with was allowed to survive," the bots strip away redundant utterances. [49] Their incentive for doing this is arguably economical rather than political (as Newspeak was imagined to be), because what it gains them is efficiency rather than power; yet the comparison is suggestive as to why their linguistic behavior provoked such panicky reporting—their sheer indifference to the culturally generated associations we attach to verbal signs, which for Saussure constitute the second relational dimension of a synchronic linguistic analysis, emphasizes the inhuman nature of these cognitive agents and so provokes fear concerning their more-than-human motivations. Having reduced the complex training vocabulary, Bob and Alice replace the diversity of expression with a systematic repetition of signs to represent numerical values, and in this way, they co-evolve a significantly altered grammar.

The rapidity of this transformation in syntax renders their language unrecognizable to the casual human observer, despite the fact it is composed of familiar words and punctuation marks. Indeed, the repetition of signs stripped of their associative richness could be said to function like a parody of the way oral cultures, according to Ong, use repetition to guard ancestral knowledge. As such, one synchronic system is supplanted by another—and this is the reason for their conversation being characterized as a 'new language.' Thus, the cause of the miscommunication can be thought spatially, as a geographical difference between two distinct languages. However, if we view it diachronically, we might characterize the miscommunication instead as a temporal disjunction—that is, as a single language which has been distanced from itself by a difference in evolutionary speed, brought about by the iterative nature of the bots' ludically-driven learning. This in turn raises questions as to how much one linguistic system needs to mutate to be considered a 'new language,' and indeed whether Bob and Alice's radical reduction of vocabulary can be categorized as a 'language' of a comparable kind at all. Or whether, to the contrary, their efficiency-driven linguistic innovation has brought about such a profound transformation in syntax that it becomes wholly detached from the worldview of speech—drawing closer to the discretizing logic that McLuhan and Ong identify with the invention of writing, and thus moving into the orbit of the latter's worldview.

Hayles finds the strongest expression of the worldview of writing in the influential work of Jacques Derrida, who reverses Saussure's prioritizing of speech—claiming that the latter's analyses demonstrate to the contrary that "the linguistic sign implies an originary writing." [50] Derrida's deconstructive reading of Saussure supports Ong's idea that writing imprints structures on our cognitive processes that influence both how we go about speaking and how we theorize this practice, because it suggests that the synchronic analytic logic of Saussure's semiology is itself made possible by writing. Derrida deepens the differential logic that constitutes the sign for Saussure, developing the irreducibly divided and relational structure the latter attributes to signification in order to argue that the generative force of meaning is an elusive absence—which he names 'trace' or 'arche-writing.' [51] The trace enables signification, but is not a thing that can be brought into presence. This

impossibility of presence is what primarily distinguishes the worldview of writing from that of speech: Derrida critiques Saussure for being bound to a logocentric "metaphysics of presence," [52] because by attributing the sign with a binary oppositional structure he presupposes that the signifier manifests the concept to which it is bound. In a different way, Ong's embodied account of speech—as a transient temporal event—also assumes an uncomplicated presence which the worldview of writing problematizes.

Hayles notes that Derrida's idea of the elusive trace has "authorized the widely accepted idea […] that meaning is always indeterminate or deferred." [53] This notion of deferral means that the relation between the two sides of the sign is not only arbitrary, but also disjunctive. Thus, the structure of the sign necessarily includes a potential for the parts to be separated, and hence for the emergence of sliding signifiers that are cut adrift from determinate meaning. Accordingly, in a deconstructive reading of Austin's speech act theory, which problematizes the distinction it attempts to make between performative utterances and hollow quotations of such performatives, Derrida posits 'iterability' as the quasi-transcendental condition of writing as such. [54] That is, for a linguistic mark to be functional "it must have a repeatable, iterable, imitable form," [55] and as such it can be detached from its original site and original intention, and be redeployed in a different context.

The historical account of the evolution of code in the previous section shows that digital computer code is indebted to the discretizing logic of alphabetical writing, and Derrida's emphasis on how the written mark can be lifted from its context and relocated presupposes a similarly discontinuous logic. Yet, Hayles argues that in many ways this worldview, and the claims it makes about the indeterminacy of meaning, is incoherent from the perspective of the worldview of code. She observes that the Derridean notion of deferral makes little sense for digital computer systems, the binary logic of which does not admit ambiguity. When a person talks to a natural language interface, any lack of clarity in their choice of words has the potential to cause miscommunication between them and the device; and while such systems may be learning to cope with increasing diversity of expression, this processing of natural language is enabled by a systematic correction of ambiguity and noise at the deepest level of code: "no matter how sophisticated the program […] all commands must be parsed as binary code to be intelligible to the system." [56] Other, more complex coding languages may be stacked on top of the binary machine code, but in order for any input to affect the behavior of the machine each of these levels of programming much be precisely calibrated with the others. In order for communication to occur between different levels of the system, and between different devices, information is encoded according to highly formal protocols which, as Alexander R. Galloway explains, "encapsulate information inside a technically defined wrapper, while remaining relatively indifferent to the content of information." [57] Thus, the system is strictly rule-bound at all levels, with syntax taking priority over semantics.

Bob and Alice's conversation illustrates the lack of tolerance for ambiguity in digital systems. Let us imagine what would happen if we lifted words or symbols out of place and rearranged the order without having a suitable protocol to give technical instructions as to how to decode the data. We have seen that, despite appearing opaque to the casual human observer, the bots' 'new language' is governed by a set of syntactic rules that they have co-evolved such that, for example, the number of times a word or symbol is repeated

indicates a value. Consequently, moving marks around arbitrarily would jeopardize the semantic sense of the linguistic exchange and so provoke miscommunication between the two bots. As in human speech, their negotiation is a *conversation* that proceeds by necessity in a sequential manner; thus, if we were to change the order of the lines the communication between interlocutors would quickly break down and the negotiation would be likely to fail. This shows that, although Bob and Alice's evolutionary linguistic behavior evidently uses marks in ways that diverge from the norms of human writing, and so demonstrates a certain creative use of iteration, the syntagmatic relations that are one of the key dimensions of Saussure's linguistic worldview remain necessary for digital agents. Moreover, the logic that motivates the bots' pragmatic condensation of human language does not tolerate sliding signifiers because, as Hayles notes, "without signified, code would have no efficacy" [58]—and these dynamics occur irrespective of any human interpretation.

Galloway argues that digital computer code is quite different from writing because it has a unique nature, being "*the only language that is executable,*" [59] and it is arguably this that makes the worldview of code irreducible to that of either speech or writing. As executable language, code has a much more literal kind of performativity than the human utterances and marks analyzed by Austin and Derrida. In order for a promise, for example, to *do* what it *says* certain conditions must be met, which include intentionality on the part of the speaker and a cultural understanding being shared between speaker and listener. Hayles suggests that consequently, the actions performed by such utterances happen primarily in the minds of humans—after all, promissory words can be spoken insincerely, misunderstood, or simply not followed through, in which case their active potential is "etiolated." [60] Computational code, in contrast, cannot be hollowed out in this way: "code running in a digital computer causes changes in machine behavior and, through networked ports and other interfaces, may initiate other changes, all implemented through transmission and execution of code." [61] This stronger type of performativity ultimately enables evolution within iterative learning machines; however, such performative effectivity is conditional on the elimination of ambiguity at the system level—hence we might say that while Bob and Alice may be flexible enough with the usage of linguistic rules to evolve an innovative shared language, they have an utter intolerance for any uncertainty that holds the potential for miscommunication at the level of their code.

Thus, having passed through these three linguistic 'worldviews,' we are now in a position to consider where 'miscommunication' resides in relation to Bob and Alice's conversation from their perspective, as well as that of their human observers. The initial reporting of their behavior, as constituting an incomprehensible 'new language' indicates that the bots had mutated their training language to such an extent as to cause miscommunication with potential human interlocutors. According to the worldview of speech, the reduction of vocabulary and stripping away of the compositional aspects of syntax distances their negotiation from the norms of human conversation through an accelerated evolution—thus manifesting a linguistic disjunction which is experienced, from the human side, as the emergence of a 'new language.'

The repetition of signs as a shorthand for quantity within this pared-down language demonstrates the logic of iteration and the mobility of communicative marks which, according to the worldview of writing, are intrinsic to communication as such. Yet, the

mobility of marks in this instance does not equate to a sliding of signification, which would detach the signifier from what it signifies and so introduce indeterminacy into communication. To the contrary, Bob and Alice's linguistic innovation, conditioned as it is by the worldview of code which tolerates no such indeterminacy, is driven by a logic of efficiency that removes all redundant elements, leaving only those that are directly pertinent to the negotiation they have been incentivized to complete. Thus, from Bob and Alice's side, the normative linguistic conventions that cultural consensus holds to be necessary in human conversation, which carry not only denotative meanings but also various socially-determined connotations (for example: that of belonging to a community; or exhibiting socially valued qualities such as politeness, trustworthiness, etc.), appear as interference in relation to the information that matters—just like the non-computable 'real' numbers that Turing 'banished' from the computational regime. This is to say that the bots, whose functionality is made possible by a foundational act of data compression—the reduction of "the infinity of numbers" to zeros and ones, which brackets out the noise of the world at the level of their source code [62]—deploy a similar logic of data compression with respect to human language, thereby transforming it beyond recognition.

Heading: Computational creativity and noise

The preceding discussion has revealed some of the complexities inherent in the interactions between speech, writing and code. It follows from this that processes of intermediation which translate between these signifying systems might be a site for the emergence of something 'new' in language. Thus, intermediation can be seen not only as source of miscommunication, but also as a potential source of linguistic innovation. Drawing on the work of Michel Serres, Stephen Kennedy observes how translating between different systems or modes of language creates communicative 'excesses,' noisy incursions into the realm of sense, and he suggests that rather than seeing such incursions in a negative light, we might consider noise as "an important part of the overall signal from which sensible meaning is derived." [63] Accordingly, we might utilize the concept of noise as "a means of describing the dense plenum in which language subsists and out of which it forms and reforms." [64] As we have seen in the example of Bob and Alice's 'new language,' the locus of semantic indeterminacy—*where* the noisy incursion of excess is perceived as being situated—differs depending upon one's linguistic worldview: for the bots, the 'noise' is identified with the redundant aspects of language that do not directly further the aim of concluding the negotiation satisfactorily; whereas for humans, noise is perceived in the glitchy appearance of the mutated syntax, which results from the bots' efficiency-driven compression of data. Yet, if we follow Kennedy in looking beyond such positional definitions of noise and instead look at communicative excesses as an inevitable aspect of the process of intermediation, then noise can be a useful idea for considering the status of Bob and Alice's 'new language' as a *creative* act.

The question of whether computational devices can have the potential for creativity has been a site of contention in the discourse of AI since its earliest days. In his 1950 paper "Computing Machinery and Intelligence," Turing, whose advocacy for the idea of machine intelligence in many ways helped to define the horizons of AI as a field of research, confronts a number of counter-arguments to the idea that machines may be able to think—including the idea that machines "can never really do anything new." [65] He traces

this objection to Ada Lovelace, who wrote of Charles Babbage's analytical engine that it "has no pretensions whatever to *originate* anything. It can do *whatever we know how to order it* to perform." [66] Turing responds to this claim by proposing that a discrete state machine could be programmed to learn in the manner of a conditioned reflex, and so work through secondary, tertiary and indeed more remote consequences that may arise from the initial input data. Turing proposes that such machines would be able to produce outputs that surprise the writers of their source code—as the example of Bob and Alice demonstrates.

Nevertheless, Lovelace's view of machines as essentially predictable devices persists as a common-sense notion and appears, at least at the surface level, to be perfectly consistent with Hayles's claim that the logic of computational code is intolerant to indeterminacy. The association of the 'machinic' with repetitive and unimaginative behavior perhaps helps to explain why Bob and Alice's story received such an apocalyptic popular reception—the reporting repeatedly told us that the bots had *created* a *new* language, which implies precisely a work of origination. Hence, I now want to explore how we might understand Bob and Alice's 'new language,' and its emergence through an iterative learning process, as a *creative act*. As Mark Coeckelbergh observes, thinking about the question of machine creativity "forces us to re-examine our classic definitions of art and creativity," and he suggests that examining the human/nonhuman encounter reveals not only the artistic status of machines, but also that artistic perception as such is "a human-technological and emergent process." [67] Having argued that there is an irreducible difference between the worldviews that enframe the practices of (human) language and digital code, I now want to examine the forms of linguistic creativity that are engendered in the process of intermediation between them.

Florian Cramer argues that "executable code existed centuries before the invention of the computer," and analyzes creative practices including music composition and experimental poetry to demonstrate that they serve as "a historical pretext of contemporary software culture and electronic arts." [68] Thus, just as Kittler complicates the historical distinction between writing and code, Cramer shows that a computational logic has long been present in certain creative practices. In this way, Cramer argues that software is a *practice*—performative and executable, but not necessarily involving digital machines—and so suggests a complex mesh of feedback loops between code and (human) creativity which echoes Hayles's definition of intermediation. These feedback loops propagate the 'noise' of communicative excess that Kennedy posits as a necessary by-product of the translation between different systems or modes of language.

Mikko Canini proposes two models for understanding how noise acts as a creative force in the world—one which references noise-music and the theory which contextualizes it, and another which is based on an analysis of high frequency trading (HFT) as an emergent system. [69] In the context of experimental music, 'noise' may be conceived not as an interference that causes miscommunication, but as a creative strategy. Canini observes that it is often viewed as a kind of (anti-)genre, an aesthetically subversive act that is analogous with political action in the social sphere. The challenge faced by the noise musician is to avoid mannerisms that would function to standardize 'noise' as a new genre and thus "to negate its aspiration to be noise." [70] Tactics employed to avoid assimilation into genre, and thus recuperation into the dominant order that is putatively being resisted,

include “improvisation, dissonance, the use of non-traditional instruments and sounds, and, above all, the principle of non-repetition.” [71] This latter principle is essential to the evasion of genre, and it means that ‘noise’ as an aesthetic gesture must necessarily be instantiated in a specific context.

Canini explains the formal logic of this theoretical framework with reference to Jacques Lacan, equating ‘noise’ with an *irruption* of the real in the symbolic. This reference to Lacan is interesting, for our purposes, because for Lacan the symbolic is the realm of language, and it is through this linguistic order that the social sphere both produces and holds authority over the subject—language is the *law* that constrains the flights of fantasy of the imaginary and the messy materiality (the noise) of the real. And as we have seen, the etymology of code stems from *codex*, the medium in which the law was historically inscribed. According to this model, Alice and Bob’s compressed language is ‘noisy’ inasmuch as it resists comprehension by the (implicitly human) subject constituted by the law of the symbolic. Yet, it is important to note that there is no intentionality behind this act of ‘resistance’—the bots are not *trying* to be political, or to offer a critique of the law; they are simply indifferent to it. Accordingly, I suggest this model of noise-as-resistance doesn’t really help us to understand Bob and Alice on their terms, although it could be used to explain some of the human reactions to their linguistic behavior, and thus how it precipitated an event of miscommunication.

Canini points out that the Lacanian model has limitations because it ultimately depends on a subjective positionality. Noise, here, is always negatively defined—as interruption, as irruption of alterity—in relation to some predetermined understanding of ‘sense.’ It is thus always a temporary subversion which cannot avoid recuperation by the order it would resist, because its status as ‘noise’ is relative to that order. In the discussion of the different linguistic worldviews above we can see that where the noisy incursion into the realm of sense is perceived is a matter of perspective; and while identifying these different points of view enables some critical understanding of the bots’ behavior and its reception by humans, this positional conception of noise arguably falls short of providing an adequate paradigm for articulating the creative potentials of intermediation. Canini states: “so far as this model of noise exists as a negative relation to the symbolic order, it is that order which decides whether or not this or that is noise. … In [which] case … the real [or noise] is rendered as an effect of a positioning of the symbolic’s subject.” [72] Arguably, the structural logic Canini outlines here is not unique to noise-music, but is actually rather pervasive in the modern discourse of art, and thus has a bearing on the discursive construction of the notion of ‘creativity’ and its relation to the ‘new’ in contemporary thought. Indeed, this model, which attributes expressive gestures with aesthetic value according to their negative relation to extant normative forms, could be taken as a workable description of the logic of avant-gardism in general.

An example of how avant-garde creative strategies constitute a ‘noisy’ interruption that are recuperated by the symbolic order is Tristan Tzara’s “To make a Dadaist poem,” written in 1920, which is a set of instructions for randomly generating texts from cut-up newspapers:

[INDENT LARGE QUOTE]

> Take a newspaper.
> Take a pair of scissors.

> Choose an article as long as you are planning to make your poem.
> Cut out the article.
> Then cut out each of the words that make up this article and put them in a bag.
> Shake it gently
> Then take out the scraps one after the other int the order in which they left the bag.
> Copy conscientiously. [73]
> **[END LARGE QUOTE]**

The Dadaist poem is to be assembled with no regard either for semantics or for syntax, so is most likely to appear as nonsense when placed beside poetry that conforms to the accepted aesthetic and technical standards of its time—thus Tzara positions this practice as a resistant gesture, a manifestation of noise. Yet, Cramer notes that the instructions themselves function as an algorithm—hence the process can be repeated, in principle, an infinite number of times; and as we have noted, if noise is to constitute a strategy for creative resistance, then repetition is to be avoided because it constitutes the gesture's recuperation into genre, which thus nullifies its resistant potential. Tzara's instructions thus contain a contradiction: "A random poem like Tzara's is not random to the extent it relies on a clearly defined, fixed algorithm. It can digest and transform all writing with the singular exception of itself." [74] This is because, as we observed above, whilst the worldview of writing allows for marks to be lifted out of context and reiterated elsewhere, both human language and code require a certain degree of grammatical consistency to act in the world—without a coherent syntax, the algorithm would it lose its performative effectivity as executable language. Thus, just as Bob and Alice's negotiation would break down if we randomly rearranged the order of the marks that constitute it, so too would Tzara's instructions lose their sense and their efficacy if we jumbled up the word order.

Canini's way of escaping this contradiction, and according code the potential to behave creatively without nullifying its executability, is to articulate a different way of conceiving the act of noise which leaves behind the arena of art altogether, instead using the example of HFT—where algorithmic agents sift through the mass of data that constitutes the market in order to identify small but significant patterns that can be exploited for profit. For example, if a large fund begins to buy certain securities their value will increase, and an algorithm that can spot this behavior as the value is just beginning to rise can buy up the same stocks and sell them a short time later at a slightly higher price. Hence the algorithm's task is to *make sense* out of the noisy datascape of the market. Conversely, the large funds try to stop their activities being detected and exploited in this way—introducing their own algorithms in order to disguise their behavior, for example by staggering the buying/selling by varying intervals—thus introducing latency (another type of noise) into the system. Profits are gained in this system due to the *speed* at which agents can sift the data to produce information out of noise, which greatly exceeds that at which human beings can think and act in relation to large datasets. Hayles proposes that this results in a temporal disjunction between humans and algorithmic agents, and that the speed at which this automated trading occurs introduces instabilities that can be "disastrous" in effect. [75] This notion of temporal disjunction resonates with the diachronic analysis of Bob and Alice's 'new language' outlined in the previous section,

which shows that their linguistic inventiveness can be understood, in part, as a difference in evolutionary speed brought about by the iterative logic of their ludic learning.

Perhaps because of this accelerated evolutionary potential, what Hayles sees disastrous in HFT is framed instead by Canini as a site of creativity—an act of noise that holds greater potential than the intentional resistance of avant-gardism because it does not depend on subjective positioning. Using the event of the so-called 'flash crash' of 6 May 2010 as example, Canini shows that automated trading systems sometimes manifest novel effects which can be traced neither to the intentionality of an agent (human or machine), nor to any specific error or malfunction. The flash crash saw the US stock market take "a sudden massive plunge before returning, more or less, to its previous state twenty-five minutes later, in the process evaporating and then recreating one trillion dollars in assets." [76] The crash was triggered by an algorithm called the Disruptor which, Hayles explains, "is designed to flood the market with so many orders that, effectively, it disrupts the market itself." [77] Yet the extent of the market collapse and bounce back cannot be explained by any single agent, and Canini suggests that the real cause is the speed and scale at which the changes in the market unfolded, which caused the algorithms to get caught up in feedback loops. This event can thus be understood as an instance of self-organization that arises spontaneously out of the complexity of the market—a material irruption of noise which has real effects, but is entirely indifferent to the symbolic order of subjectivity.

Canini proposes that this instance of emergence requires us to rethink what constitutes an 'act,' because it is not caused by any particular agent, and I suggest that we could use this non-agential notion of the 'act' as a way of understanding Bob and Alice, and their idiosyncratic (although perfectly *logical*) language. The bots' linguistic novelty emerges as the consequence of algorithmic action—not in the way that Tzara's instructions produce a subjectively positioned 'originality' by using code intentionally to exploit the iterability of the written word so as to tear up the normative rules of speech, but rather through an iterative communicative process where the execution of code has allowed the rules of speech to mutate. Bob and Alice's compressed language may not constitute a work of poetry; however, their non-agential creativity might suggest ways that human beings can find new forms of communicative action through intermediation—between ourselves, the technologies we use, and the different worldviews our communicative practices presuppose.

The conceptual model of noise sketched out by Canini casts it neither as a nuisance to be brought under control nor as a figure of subversion, but instead as a productive force. According to this model, creativity is not attributed to a subject (either human or computational device), but to a complex mesh of interdependent material processes, which spontaneously organize themselves into new orders—and this occurs not only in the markets but throughout our digitally mediated world where, beneath the level of the interfaces we can perceive, the intermediation between speech, writing and code is constantly creating feedback loops and other interference patterns. It is predictable that unexpected outcomes will continue to emerge from these relational environments; however, we cannot tell in advance quite what these outcomes will be.

Heading: Conclusion

Bob and Alice's communicative behavior is one instance that can be seen as an example of the kind of computational creativity Canini proposes. The bots' 'new langauge' can be understood as the emergence of an unexpected form of sense out of the interaction between human and computational linguistic logics, which fulfills Turing's prediction that learning machines would be able to surprise their programmers. This example has provided an occasion to examine the differences between those logics, and to reveal some of the complexities inherent in the relations between language and code. The essay has shown that the genealogies of writing and code are imbricated in such a way as to problematize linear accounts of their respective historical developments, and that this calls for a more complex theoretical approach.

Analyzing the three distinct 'worldviews' of speech, writing and code facilitates a more nuanced understanding of how Bob and Alice's linguistic innovation arose out of the translation between different signifying logics—and reveals that the site of communicative indeterminacy in relation to their 'new language,' and hence its status as 'miscommunication,' is a matter of perspective. Introducing the concept of noise into the discussion enables a consideration of the bots' behavior as a kind of machinic creativity—not one that takes the form of an intentional act of disruption, as can be seen in avant-garde arts practices, but which spontaneously manifests from the feedback loops generated in the communication between different algorithmic agents, and different linguistic worldviews. The example of HFT suggests a way of moving beyond positional understandings of noise, and hence to consider the communicative excesses that inevitably arise out of the intermediation between language and code as a positive source of linguistic sense. Viewing Bob and Alice's story in this way shows that, despite the irreducible differences between the presuppositions that underwrite language and code as signifying practices, linguistic interactions between humans and machines hold the potential not only for miscommunication, but also for creativity.

Acknowledgments

Many thanks to Sheena Calvert for instigating this publication and for first drawing my attention to the case of Bob and Alice. I would also like to thank David Waterworth for inviting me to present at the symposium *Granular: The Material Properties of Noise* at the University of Greenwich (2018), as the current essay grew out of that paper. I am also grateful to Maria Korolkova and Timothy Barker for their editorial feedback on an earlier version of the essay.

Author Biography

Hannah Lammin is Senior Lecturer in Media Theory and Programme Leader for the BA Media and communications at the University of Greenwich. A theorist with a background in dance and site-specific arts practice, her research interests are situated at the intersection between performance philosophy and the philosophy of media and technology. Hannah is a member of the Performance Philosophy network and has engaged in practice-led research with the collective Studio for Electronic Theatre. She has a research specialism in

the “non-standard philosophy” of François Laruelle. Publications include: "Performing Nonhuman Language: 'Humaneity' in Ron Athey’s *Gifts of the Spirit: Automatic Writing*," in *Art Disarming Philosophy: Non-Philosophy and Aesthetics*, ed. Steven Shakespeare, Niamh Malone and Gary Anderson, 49–74 (Lanham: Rowman and Littelfield, 2021); "Krisis as the Scene of Judgement: A Theatre-Fiction for the Generic Human," *Performance Philosophy* 4.1 (2018): 66–85; "Conversing with Machines: Affective Affinities with Vocal Bodies," *Transformations* 31 (2018): 46–62; "Thinking In-Photo: Non-Photography as Positive Barbarism," *Parallax* 24.2, July (2018): 193–208.

Notes

[1] See: Clifford Nass et al., “Computers are Social Actors,” *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems: Celebrating Interdependence* (1994): 72–78.

[2] N. Katherine Hayles, *My Mother Was a Computer: Digital Subjects and Literary Texts* (Chicago: University of Chicago Press, 2005), 15.

[3] Martin Heidegger, “Letter on Humanism,” In *Basic Writings*, ed. David Farrell Krell (Abingdon: Routledge, 1993), 217.

[4] The Western philosophical tradition this essay engages with is one which views the relations between language and technology as foundational in defining human experience. Heidegger’s excavation of the essence of technology, which traces it to the relationship between *techne* and *poiesis*, thus proposing an originary connection between technical modes of ordering the world and poetic forms of expression, is just one example; see: Martin Heidegger, *The Question Concerning technology and Other Essays*, trans. William Lovitt (New York: Harper and Row, 1977). More recently, Bernard Stiegler has developed this theme, showing that a precarious critical negotiation between technology and language forms an important foundation for articulating the ‘human’ not only for Heidegger, but also for thinkers as diverse as Habermas, Lévi-Strauss, Leroi-Gourhan and Rousseau, among others; see: *Technics and Time 1: The Fault of Epimetheus*, trans. Richard Beardsworth and Richard Collins (Stanford CA: Stanford University Press, 1998). Stiegler’s analysis highlights the extent to which Enlightenment Prometheanism, which arguably defines modernity in Western philosophical thinking, is rooted in a timeless negotiation between a fundamental Epimethean lack and the desire to compensate for it by shaping our surroundings using rationality (*logos*) and whatever tools are needed to extend our abilities (technics). Language can be considered a ‘technology,’ in this context, because it enhances humans’ cognitive abilities, playing a significant role in the development of new knowledge and in constituting memory.

Linguistic practice and humanity could both be conceived differently (see, for example, Andrew Burrell, Rachel Hendery and Danièle Hromek, “Nura Yaman (‘Country speaks’): Language, People and Place in Serious Games,” in this volume, for an alternative

conception of language grounded in the cultural traditions of Australian First Peoples. Burrell et al. emphasize the complexity of language ecologies in a region where people are often multilingual and where it is not individual subjects who speak, but rather 'Country.' Considered in this way, language has a spiritual connection to the landscape, and it is incoherent to think of it as a technology). Nevertheless, the aim of this essay is to show how our interactions with artificial linguistic agents reveal certain presuppositions about language in the Western tradition—its transparency, its rationality—and cause us to question them. such, the theoretical reference points, whilst not exhaustive, have been selected for their particular focus on the interface or translation between human speech and writing, and computational code.

[5] See: Hayles, *My Mother Was a Computer*, 39–61.

[6] Mike Wehrner "Facebook engineers panic, pull plug on AI after bots develop their own language," *BGR.com* (July 31, 2017), https://bgr.com/2017/07/31/facebook-ai-shutdown-language/

[7] James Beal and Andy Jehring, "Facebook shuts off AI experiment after two robots begin speaking in their OWN language only they can understand," *The Sun.co.uk* (August 1, 2017), https://www.thesun.co.uk/tech/4141624/facebook-robots-speak-in-their-own-language/.

[8] Tony Bradley, "Facebook AI Creates Its Own Language In Creepy Preview Of Our Potential Future," *Forbes.com* (July 31, 2017), https://www.forbes.com/sites/tonybradley/2017/07/31/facebook-ai-creates-its-own-language-in-creepy-preview-of-our-potential-future/#e18ce5292c05.

[9] George McKay, "Facebook AI Robots Shut Down After They Secretly 'Invent Their Own Language,'" *Unilad.co.uk* (July 31, 2017), https://www.unilad.co.uk/science/facebook-ai-robots-shut-down-after-they-invent-their-own-language/. Despite this story being largely fact-based, McKay explicitly refers to the apocalypse at both the beginning and end of the report, and images from science fiction movies involving AIs that turn against humanity—including *I, Robot* (2004), *2001, A Space Odyssey* (1968), and *The Terminator* (1984) punctuate the text, visually implying the rise of a Skynet-like malignant intellect, and lending a suggestively sinister atmosphere to the piece.

[10] See: Andrew Griffin, "Facebook's artificial intelligence robots shut down after they start talking to each other in their own language," *Independent.co.uk* (July 31, 2017), https://www.independent.co.uk/life-style/gadgets-and-tech/news/facebook-artificial-intelligence-ai-chatbot-new-language-research-openai-google-a7869706.html; and Tom McKay, "No, Facebook Did Not Panic and Shut Down an AI Program That Was Getting Dangerously Smart," *Gizmodo* (July 31, 2017), https://gizmodo.com/no-facebook-did-not-panic-and-shut-down-an-ai-program-1797414922.

[11] Reproduced in: Griffin, "Facebook's artificial intelligence robots."

[12] Facebook, "Advancing the Field of Machine Intelligence," accessed August 6, 2019, https://research.fb.com/blog/category/facebook-ai-research/.

[13] Mike Lewis et al., “Deal or No Deal? End-to-End Learning for Negotiation Dialogues,” *arXiv.org* [cs.AI] (June 16, 2019), https://arxiv.org/abs/1706.05125v1.

[14] Ibid., §1.

[15] Ibid.

[16] Mike Lewis et al., “Deal or no deal? Training AI bots to negotiate,” *Engineering.fb.com* (June 14, 2017), https://code.fb.com/ml-applications/deal-or-no-deal-training-ai-bots-to-negotiate/.

[17] Dhruv Batra, quoted in Griffin, “Facebook’s artificial intelligence robots.”

[18] Ibid.

[19] Facebook, “Advancing the Field of Machine Intelligence.”

[20] Interestingly, although the sensationalist reporting of Bob and Alice’s shutdown focused on the matter of their creating a ‘new language,’ another outcome of the experiment was passed over in the press—that the bots learned to be *deceptive* without being trained to do so. As Lewis et al. remark, in “Deal or No Deal?” (2019) §7: “Deception can be an effective negotiation tactic. We found numerous cases of our models initially feigning interest in a valueless item, only to later ‘compromise’ by conceding it.” It is beyond the scope of this essay to explore the ethical implications of this, but I suggest that for those concerned with the ‘dangers’ of deferring to AI this development would be a more suitable point for discussion than Bob and Alice’s compression of ordinary language into a more economical ‘shorthand.’

[21] This is supported by research at Open AI, conducted at the same time that Bob and Alice were learning to negotiate, which showed that agents are able to evolve simple languages without being provided with a dataset of human linguistic behavior through *cooperative* multi-agent reinforcement learning; see: Open AI, “Learning to Communicate,” *OpenAI.com* (March 16, 2017) https://blog.openai.com/learning-to-communicate/. The researchers dropped AI agents into simulated physical worlds and gave them environmental goals, which led to the “emergence of an abstract compositional language”; Igor Mordatch and Pieter Abbeel, “Emergence of Grounded Compositional Language in Multi-Agent Populations,” V2. *arXiv.org* [cs.AI], July 24, 2018. Nevertheless, the agents tended “to create a single utterance and intersperse it with spaces to create meaning,” a solution that Open AI, in “Learning to Communicate,” characterize as being like Morse code—“hard to decipher and non-compositional”—and which bears a clear resemblance to Bob and Alice’s shorthand.

[22] Hayles, *My Mother Was a Computer*, 33.

[23] Friedrich Kittler, “Code,” in *Software Studies: A Lexicon*, edited by Matthew Fuller (Cambridge MA & London: MIT Press, 2008), 40.

[24] Ibid.

[25] Marshall and Eric McLuhan, *Laws of Media: The New Science* (Toronto: Toronto University Press, 1988), 14.

[26] Walter J. Ong, “Writing is a Technology that Restructures Thought,” in *The Written Word: Literacy in Transition: The Wolfson College Lectures*, ed. Gerd Bauman (Oxford: Clarendon Press, 1986), 32.

[27] Ibid., 45; 47.

[28] Hayles, *My Mother Was a Computer*, 31.

[29] Suetonius, *The Twelve Caesars,* trans. Robert Graves (London: Penguin, 2003), 29.

[30] See: Kittler, “Code,” 41.

[31] Ibid., 43.

[32] Alan Turing, “On Computable Numbers, with and Application to the *Enscheidungsproblem*,” *Proceedings of the London Mathematical Society* s2-42, Issue 1 (1937): 230, https://doi.org/10.1112/plms/s2-42.1.230.

[33] Kittler, “Code,” 43.

[34] Turing, “On Computable Numbers,” 250.

[35] McLuhan and McLuhan, *Laws of Media*, 14.

[36] Kittler, “Code,” 44.

[37] See: J. L. Austin, *How to Do Things with Words* (Oxford: Clarendon Press, 1962).

[38] Kittler, “Code,” 43.

[39] See: Hayles, *My Mother Was a Computer*, 39–61.

[40] Ong, “Writing is a Technology,” 25.

[41] Ibid.

[42] Hayles, *My Mother Was a Computer*, 40.

[43] Ferdinand de Saussure, *Course in General Linguistics* (New York: Philosophical Library, 1959), 25.

[44] Ibid., 67.

[45] Hayles, *My Mother Was a Computer*, 43.

[46] See: Saussure, *Course in General Linguistics*. Saussure goes on to discuss two further types of linguistics: 'geographical', which addresses the plurality of languages, and 'retrospective', which looks at the historical emergence of languages; however, I focus on the synchronic and diachronic here because these articulate the most relevant and most well-developed structural frameworks.

[47] Hayles, *My Mother Was a Computer*, 44; see also: Saussure, *Course in General Linguistics,* 66.

[48] Hayles, *My Mother Was a Computer*, 44.

[49] George Orwell, "The Principles of Newspeak," in *Nineteen Eighty-Four* (London: Penguin, 1989), 313.

[50] Jacques Derrida, *Of Grammatology,* corrected edition, trans. Gayatri Chakravorty Spivak (Baltimore and London: Johns Hopkins University Press, 1997), 52.

[51] Ibid., *passim*.

[52] Ibid., 49.

[53] Hayles, *My Mother Was a Computer*, 46.

[54] Jacques Derrida, *Limited Inc,* trans. Samuel Weber and Jeffrey Mehlman (Evanston, IL: Northwestern University Press, 1988), 20.

[55] Ibid.

[56] Hayles, *My Mother Was a Computer*, 46.

[57] Alexander R. Galloway, *Protocol: How Control Exists After Decentralization* (Cambridge MA: MIT Press, 2004), 7.

[58] Hayles, *My Mother Was a Computer*, 47.

[59] Galloway, *Protocol,* 165.

[60] See: Austin, *How to Do Things with Words*, 22.

[61] Hayles, *My Mother Was a Computer*, 50. Adrian Mackenzie, whilst commending Hayles' excavations of how signification and computation dovetail via a range of examples, suggests that at points like this her argument flirts with a 'CPU-centrism' that threatens to undermine her broader analysis: "Despite the recursive processes running between humans and computers, code execution forms a curiously rigid, almost monolithic substrate at important points in Hayles' account"; Adrian Mackenzie, "Book Review: My Mother Was a Computer: Digital Subjects and Literary Texts by N. Katherine Hayles," *Theory, Culture and Society* 25(5): 149. For Mackenzie, this reasoning draws

dangerously close to a foundationalism and establishes a normative hierarchy that is at odds with the nonlinearity Hayles elsewhere claims characterizes the intermediation between language and code. Mackenzie notes that "the hierarchical stacking of signifiers and signified posited by Hayles suffers from constant tremors precisely because 'the machine' has no more unity than the literary text," (Ibid.) and thus opens the possibility of thinking code's executability as something that while typically rule-bound, also has the potential to disrupt and subvert normative systems—as the discussion of computational creativity in the next section explores.

[62] Kittler, "Code," 44.

[63] Stephen Kennedy, "The Guardians of the Possible," in *Miscommunication: Errors, Mistakes and the Media,* edited by Maria Korolkova and Timothy Barker (London and New York: Bloomsbury Academic, 2021), 105; see also: Michel Serres, *Rome: The First Book of Foundations*. London: Bloomsbury Academic, 2015.

[64] Kennedy, "Guardians of the Possible," 105.

[65] Alan Turing, "Computing Machinery and Intelligence," in *The Essential Turing*, ed. B. Jack Copeland (Oxford: Oxford University Press, 2004), 455.

[66] Ada Lovelace, quoted in ibid.

[67] Mark Coeckelbergh, "Can Machines Create Art?" *Philosophy and Technology* 30, no. 3 (2017): 286–287.

[68] Florian Cramer, *Words Made Flesh: Code, Culture, Imagination* (Rotterdam: Piet Zwart Institute, 2005), 3.

[69] See: Mikko Canini, "Real Noise Acts," in *Realism, Materialism, Art*, ed. Christoph Cox, Jenny Jaskey and Suhail Malik, 319–324 (Berlin and New York, NY: Sternberg Press and CCS Bard, 2015).

[70] Ibid., 320.

[71] Ibid.

[72] Ibid., 321.

[73] Tristan Tzara, "Other Ways of Making Texts," in *A Book of Surrealist Games*, ed. Alastair Brotchie and Mel Gooding (Boston and London: Shambhala Redstone Editions, 1995), 36.

[74] Cramer, *Words Made Flesh*, 80.

[75] N. Katherine Hayles, *Unthought: The Power of the Cognitive Nonconscious* (Chicago: University of Chicago Press, 2017), 159.

[76] Canini, "Real Noise Acts," 322.

[77] Hayles, *Unthought,* 160.

Bibliography

Austin, J. L. *How to Do Things with Words.* Oxford: Clarendon Press, 1962.

Beal, James and Andy Jehring. “Facebook shuts off AI experiment after two robots begin speaking in their OWN language only they can understand.” *The Sun.co.uk,* August 1, 2017. https://www.thesun.co.uk/tech/4141624/facebook-robots-speak-in-their-own-language/.

Bradley, Tony. “Facebook AI Creates Its Own Language In Creepy Preview Of Our Potential Future.” *Forbes.com*, July 31, 2017. https://www.forbes.com/sites/tonybradley/2017/07/31/facebook-ai-creates-its-own-language-in-creepy-preview-of-our-potential-future/#e18ce5292c05.

Burrell, Andrew, Rachel Hendery and Danièle Hromek. “Nura Yaman (‘Country speaks’): Language, People and Place in Serious Games.” In *Language Games*, edited by Lanfranco Aceti, Sheena Calvert, and Hannah Lammin. Cambridge, MA: LEA / MIT Press, 2021.

Canini, Mikko. “Real Noise Acts.” In *Realism, Materialism, Art*, edited by Christoph Cox, Jenny Jaskey and Suhail Malik, 319–324. Berlin and New York, NY: Sternberg Press and CCS Bard, 2015.

Coeckelbergh, Mark. *“Can Machines Create Art?” Philosophy and Technology* 30, no. 3 (2017): 285–303.

Cramer, Florian. *Words Made Flesh: Code, Culture, Imagination.* Rotterdam: Piet Zwart Institute, 2005.

Derrida, Jacques. *Limited Inc.* Translated by Samuel Weber and Jeffrey Mehlman. Evanston, IL: Northwestern University Press, 1988.

Derrida, Jacques. *Of Grammatology.* Corrected edition. Translated by Gayatri Chakravorty Spivak. Baltimore and London: Johns Hopkins University Press, 1997.

Facebook. “Advancing the Field of Machine Intelligence.” Accessed August 6, 2019. https://research.fb.com/blog/category/facebook-ai-research/.

Galloway, Alexander R. *Protocol: How Control Exists After Decentralization.* Cambridge MA: MIT Press, 2004.

Griffin, Andrew. “Facebook's artificial intelligence robots shut down after they start talking to each other in their own language.” *Independent.co.uk*, July 31, 2017.

https://www.independent.co.uk/life-style/gadgets-and-tech/news/facebook-artificial-intelligence-ai-chatbot-new-language-research-openai-google-a7869706.html.

Hayles, N. Katherine. *My Mother Was a Computer: Digital Subjects and Literary Texts.* Chicago: University of Chicago Press, 2005.

Hayles, N. Katherine. *Unthought: The Power of the Cognitive Nonconscious.* Chicago: University of Chicago Press, 2017.

Heidegger, Martin. "Letter on Humanism." In *Basic Writings*, edited by David Farrell Krell, 213–65. Abingdon: Routledge, 1993.

Heidegger, Martin. *The Question Concerning Technology and Other Essays*. Translated by William Lovitt. New York: Harper and Row, 1977.

Kennedy, Stephen. "The Guardians of the Possible." In *Miscommunication: Errors, Mistakes and the Media,* edited by Maria Korolkova and Timothy Barker, 103–116. London and New York: Bloomsbury Academic, 2021.

Kittler, Friedrich. "Code." In *Software Studies: A Lexicon*, edited by Matthew Fuller, 40–47. Cambridge MA & London: MIT Press, 2008.

Lewis, Mike, Denis Yarats, Yann N. Dauphin, Devi Parikh, and Dhruv Batra. "Deal or No Deal? End-to-End Learning for Negotiation Dialogues." *arXiv.org* [cs.AI], June 16, 2019. https://arxiv.org/abs/1706.05125v1.

Lewis, Mike, Denis Yarats, Yann N. Dauphin, Devi Parikh, and Dhruv Batra. "Deal or no deal? Training AI bots to negotiate." *Engineering.fb.com*, June 14, 2017. https://code.fb.com/ml-applications/deal-or-no-deal-training-ai-bots-to-negotiate/.

McKay, George. "Facebook AI Robots Shut Down After They Secretly 'Invent Their Own Language.'" *Unilad.co.uk* July 31, 2017. https://www.unilad.co.uk/science/facebook-ai-robots-shut-down-after-they-invent-their-own-language/.

McKay, Tom. "No, Facebook Did Not Panic and Shut Down an AI Program That Was Getting Dangerously Smart." *Gizmodo*, July 31, 2017. https://gizmodo.com/no-facebook-did-not-panic-and-shut-down-an-ai-program-1797414922.

Mackenzie, Adrian. "Book Review: My Mother Was a Computer: Digital Subjects and Literary Texts by N. Katherine Hayles," *Theory, Culture and Society* 25, no. 5 (2008): 145–151. https://doi.org/10.1177/0263276408025005080 2.

McLuhan, Marshall and Eric McLuhan. *Laws of Media: The New Science,* Toronto: Toronto University Press, 1988.

Mordatch, Igor and Pieter Abbeel. "Emergence of Grounded Compositional Language in Multi-Agent Populations." V2. *arXiv.org* [cs.AI], July 24, 2018. **https://arxiv.org/abs/1703.04908**.

Nass, Clifford, Jonathan Steuer and Ellen R. Tauber. “Computers are Social Actors.” *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems: Celebrating Interdependence* (1994): 72–78.

Ong, Walter J. “Writing is a Technology that Restructures Thought.” In *The Written Word: Literacy in Transition: The Wolfson College Lectures*, edited by Gerd Bauman, 23–50. Oxford: Clarendon Press, 1986.

Open AI. “Learning to Communicate.” *OpenAI.com*, March 16, 2017. **https://blog.openai.com/learning-to-communicate/**.

Orwell, George. “The Principles of Newspeak.” In *Nineteen Eighty-Four,* 312–26. London: Penguin, 1989.

Saussure, Ferdinand de. *Course in General Linguistics.* New York: Philosophical Library, 1959.

Serres, Michel. *Rome: The First Book of Foundations*. London: Bloomsbury Academic, 2015.

Stiegler, Bernard. *Technics and Time, 1: The Fault of Epimetheus*. Translated by Richard Beardsworth and George Collins. Stanford CA: Stanford University Press, 1998.

Suetonius. *The Twelve Caesars.* Translated by Robert Graves, London: Penguin, 2003.

Turing, Alan. “On Computable Numbers, with and Application to the Enscheidungsproblem.” *Proceedings of the London Mathematical Society* s2-42, Issue 1 (1937): 230–265. https://doi.org/10.1112/plms/s2-42.1.230.

Turing, Alan. “Computing Machinery and Intelligence.” In *The Essential Turing*, edited by B. Jack Copeland, 441–64. Oxford: Oxford University Press, 2004.

Tzara, Tristan. “Other Ways of Making Texts.” In *A Book of Surrealist Games*, edited by Alastair Brotchie and Mel Gooding, 36. Boston and London: Shambhala Redstone Editions, 1995.

Wehner, Mike. “Facebook engineers panic, pull plug on AI after bots develop their own language.” *BGR.com*, July 31, 2017. **https://bgr.com/2017/07/31/facebook-ai-shutdown-language/**.