Transformations issue 31 (2018)
www.transformationsjournal.org

ISSN 1444-3775

Conversing with Machines: Affective Affinities with Vocal Bodies

Hannah Lammin

AUTHOR BIO

Hannah Lammin recently completed her PhD in philosophy at the University of Greenwich. Her current research is concerned with AI and the philosophy of language. Publications include: "Seeing In-Photo: Non-Photography as Positive Barbarism", *Parallax* 24.2 (2018); "A Paradoxical Architecture: Babel, and the founding of community through the confounding of tongues" *Zetesis* 1.1 (2013); "ACT Oedipus: Digital Theatre and the Apocalyptic Structure of Re/presentation" *Philosophy of Photography* 3.1 (2012). She lectures in media theory at the University of Greenwich, and in contextual studies at University of the Arts London.

ABSTRACT

This article examines how the emergence of speech-driven interfaces for computational devices alters our affective relationships with machines, and argues that the rise of intelligent personal assistants such as Siri, Watson and Alexa calls for the question of affect to be brought to the centre of discourse around artificial intelligence (AI). It departs from the early imaginings and manifestations of human-computer conversations in the work of Turing and Weizenbaum, then introduces a Spinozan framework for theorising the transmission of affect and its ethical implications. It examines the affective economy engendered by vocal interfaces, drawing on a range of theories which focus on sound not only as an object of study, but also as a conceptual paradigm. It concludes by arguing that the machine voice constitutes a form of embodiment, and that according computers this "body" and inviting us to converse with them enhances our ability to enter into a sensuous relationship with them.

KEYWORDS

Artificial intelligence; affect; Turing test; voice; sonic theory

The mise en scène of machinic conversation: Turing and Weizenbaum

Alan Turing, in his famous 1950 paper "Computing Machinery and Intelligence," proposes a conversational test as a response to the question "can machines think?" This test, which Turing names the "imitation game" (§1), puts a human interrogator into a teletyped dialogue with two entities that he/she cannot see – one of which is a computer; the other a human confederate. The interrogator's task is to discern which is the human, on the basis of the ensuing conversations. From the computer's side, the aim is to deceive its interlocutor: to pass as human. By replacing the question "can machines think?" with the question of whether a machine could do well in the imitation game, Turing implies – without ever stating it as such – that intelligence can be inferred on the basis of conversational performance, effectively consigning cognitive processes to a black box by focusing attention on manifest communicative behaviour.

The conception of this game, often known simply as the "Turing test," was a visionary thought experiment which contributed significantly to the emergence of artificial intelligence (AI) as an idea, partly defining the latter as a field of research. It is still "commonly seen as the ultimate benchmark test for demonstrating that a machine 'has intelligence'" (Berrar, Konagaya and Schuster 242). The Loebner Prize, a Turing test competition which has been running annually since 1991 (Society for the Study of Artificial Intelligence and Simulated Behaviour), is widely reported in popular technology press and so is likely to influence public perception of AI. However, the Turing test has also been the subject of much debate and critique (for a concise review, see: Saygin, Cicekli and Akman). Since the 1990s, the test has increasingly been judged to be harmful as a technical goal, both because it is tautological – defining the qualities it claims to be testing for – and because the operational definition of intelligence it implies is anthropocentric, and thus limits the horizons of what an *artificial* intelligence might be (Hayes and Ford). Consequently, the test has receded from scientific research agendas. Nevertheless it remains an influential reference point in philosophical debates about AI (see, for example, the range of perspectives collected in Epstein, Roberts and Beber), because it raises interesting questions about our *relationships* with machines – questions that are becoming increasingly relevant as artificial language (AL) systems are performing ever more effectively in conversational contexts. This article aims to contribute to the philosophical discourse on AI by reflecting on our relations with speaking machines, and on how the vocal interfaces that are becoming increasingly common in the contemporary media landscape accentuate the affective dimensions already inherent in the set-up of the Turing test.

At first glance, Turing appears to propose his imitation game as a test of machine *intelligence,* because he poses it in response to the question "can machines can think?" However, he later states that this question is itself "too meaningless to deserve discussion" (Turing §6), which opens the possibility for other interpretations of the purpose of the test. In *How We Became Posthuman,* N. Katherine Hayles uses the imitation game to set the scene for her exploration of the human/machine interface*,* and proposes that we should think of it as a "magic trick" which functions by misdirection. The

interrogator's task is to judge the identity of his/her interlocutors – to "pose questions that can distinguish between verbal performance and embodied reality" (Hayles xi) – and this suggests that the identity of the unseen linguistic agents is the focus of the experiment. Hayles, however, proposes that it serves a very different purpose:

> Like all good magic tricks, the test relies on getting you to accept at an early stage assumptions that will determine how you interpret what you see later. The important intervention comes not when you try to determine which is the man, the woman, or the machine. Rather, the important intervention comes much earlier, when the test puts you into a cybernetic circuit that splices your will, your desire, and perception into a distributed cognitive system in which represented bodies are joined to enacted bodies through mutating and flexible machine interfaces. (xiv)

According to Hayles, then, the question of machine intelligence is a decoy, and the significance of the test resides in the behaviour and attitudes of the interrogator rather than the performance of the AI system. This interpretation suggests that the imitation game is an act of subterfuge which parallels that of Stanley Milgram's infamous behavioural experiments into obedience to authority. Milgram paired test subjects with actors posing as other volunteers, instructing them to pose a series of questions to the latter and to administer them with "electric shocks" (which were in fact simulated) if they were unable to answer correctly (Milgram). The subjects were thus directed to believe that they were assisting the psychologist in conducting the experiment on others, when the actual focus of the observation was their own behaviour – how much apparent pain they would be prepared to inflict. It is not obvious that such an intervention into human behaviour was, in fact, Turing's intention; nevertheless, the construct of the test arguably played a role in normalising the idea of conversing with machines by positioning the set-up as a *scientific* procedure. Recent experimental analogues of the Milgram paradigm have demonstrated that what motivates test subjects to enter into and continue with an experimental task is less an unquestioning obedience to authority, as Milgram proposed, than a sense of fellowship with the researcher based on an active belief in the validity of the scientific method as beneficial to humankind (Haslam, Reicher and Birney). If "science," as a discursive idea, is persuasive in this regard, then it is reasonable to infer that by enshrining the imitation game in the scientific discourse of AI, Turing helped people to *imagine* the possibility of conversing with computers via technological interfaces long before such technologies were available. In this sense, the question of identifying machine "intelligence," which is ostensibly the purpose of the Turing test, may indeed not be its greatest significance.

Yet, the scientific framing of the imitation game, whilst helping to normalise the idea of human-machine dialogue, also had the effect of bracketing certain sensuous aspects of the conversational experience out of the discourse. Turing specifies that the conversation "should be written, or better still typewritten … in order that tones of voice may not help the interrogator" (§1), and that the interlocutors should be placed in separate rooms so the

latter cannot appeal to the senses of sight or touch when deciding the others' identities. Nor can practical demonstrations be demanded by the interrogator. Turing proposes that this set-up "has the advantage of drawing a fairly sharp line between the physical and the intellectual capacities of a man" (§2). All this suggests that Turing saw language as a suitable medium for a test of intelligence because it could be abstracted from the embodied agents that produced it. In the "scientific" context of the imitation game test, this isolation of the ideational dimension of language from the material apparatus that generates it was necessary to produce a controlled experiment in which the variables are limited, to guard against interrogators making anthropomorphic judgements about their interlocutors based on sensory impressions (Proudfoot). However, Hayles's remark that the cybernetic circuit joins represented bodies to enacted ones, inducting perception and desire in the process, suggests that the conversation is always *material* – even if it is informatically distributed through the technologies that mediate it. In which case, Turing's bracketing out of sensory information should perhaps be viewed as part of the act of misdirection, because the interrogator cannot in fact function as a disembodied intellect. Instead, we need to conceive of the latter's intellectual experience as phenomenologically rich – as an affective event. Moreover, as Elizabeth A. Wilson highlights, Turing's work is itself animated by "circuits of affectivity" (6), and his conception of machine intelligence is open to the possibility of "affiliations between thinking and feeling" (7). This is evidenced by the conclusion of his 1950 paper, where he proposes that one way of developing AI which ought to be tested is to "to provide the machine with the best sense organs that money can buy, and then teach it to understand and speak English" (Turing §7). This implies a fundamental connection between sensory perception and human language – suggesting that both may be core aspects of "intelligence," and that the computer, too, could be imagined as experiencing affect.

The affectivity generated through human-machine conversation was revealed by the first computer program that was able to converse in English, and hence to actualise something resembling an "imitation game" – ELIZA, created by Joseph Weizenbaum at MIT and tested in 1964–6. Although this programme was not submitted to a Turing *test* – each person who conversed with it was aware that they were communicating with a machine, and Weizenbaum describes the role of the human as "conversationalist partner" (2) rather than "interrogator" – it nevertheless used ordinary language and a typewritten interface similar to that conceived by Turing. Weizenbaum chose the name ELIZA after the character in *Pygmalion* (Shaw), because "it could be taught to 'speak' increasingly well" (Weizenbaum 3), by following a set of instructions resembling those one might give to an actor to facilitate an improvisation around a certain theme. ELIZA was not nearly as sophisticated as the sensory learning machines imagined by Turing, nor would it have performed well in a general Turing test where the interrogator is free to choose the topic of conversation; but as long as the conversation was restricted to the specific context for which the scripted rules were written, it was able to maintain a conversation. Thus, for Weizenbaum, ELIZA's performance did not provide a general solution to natural language processing, but rather emphasised that linguistic understanding (human or machine) is highly dependent on context.

In his first experiment, Weizenbaum "gave ELIZA a script designed to permit it to play (I should really say parody) the role of a Rogerian psychotherapist engaged in an initial interview with a patient" (3). He chose this role because the technique of the psychotherapist involves a modified repetition of the patient's own statements, which encourage the latter to reflect on his/her own mental state – and this relatively passive verbal behaviour is fairly easy to imitate machinically. Moreover, casting ELIZA as "DOCTOR" also gives the *human* interlocutor a specific "role," thereby introducing a set of expectations regarding his/her own contribution to the conversation. The psychotherapeutic scenario thus produces a certain predictability and repetitiveness of interaction – even when a human interlocutor encounters the system for the first time. Thus a kind of *automation* is introduced into the conversation – one which Wilson observes "goes hand in hand" with ELIZA's affective "capacity to arouse" (93). Weizenbaum's use of theatrical analogies ("play," "parody," "role") underlines the extent to which the success of this early artificial language system was dependent upon the direction provided by the *mise en scène*, much like the "magic trick" of the Turing test.

Weizenbaum does not claim that ELIZA's conversational performance evidenced underlying "intelligence," because he believes that there is a qualitative difference between the calculative powers of computers and the capacity of human reason to make judgements. Consequently, he was somewhat horrified to observe the extent to which people who engaged in conversation with the program bought into the role play, and considered the machine like a person. He states: "I was startled to see how quickly and how very deeply people conversing with DOCTOR became *emotionally* involved with the computer and how unequivocally they anthropomorphized it" (Weizenbaum 6, emphasis added), giving the example of his secretary, who asked him to leave the room after only a few interchanges with the machine so she could converse with it in private. This individual had watched Weizenbaum work on the system for months, so could be under no illusions about its machinic identity; nevertheless, the experience of conversing with the program appears to have cultivated a strong affective bond. Having observed this relatively simple AL system inducing such emotional investment, which he characterises as constituting "powerful delusional thinking in quite normal people" (7), Weizenbaum suggests that a new importance should be attached to the relationships human individuals form with computers.

In common with Hayles's reading of the imitation game, then, Weizenbaum emphasises the play of affect within the communicational circuit above the question of machine intelligence. Indeed, his reflections on the emotional response ELIZA elicited support Hayles's claim about the real import of the Turing test, because they demonstrate that desire can be aroused through effective conversational performance, even when users are aware that they are communicating with a machine. Thus Weizenbaum observed an instance of what Byron Reeves and Clifford Nass call the "media equation" – the tendency for people to equate the media technologies with which they interact with "real life" (5), and thus to treat computers, TVs etc. as *social*

agents. Moreover, the importance of focusing on human-computer *relations* arguably grows as AL systems become more sophisticated. Contemporary dialogue systems are able to talk in a far wider range of contexts than ELIZA, having the capacity to initiate conversation, "to share personal opinions, relay experience of family dramas, be relevant, but also be vague, and mislead just as humans do" (Shah et al 278). Moreover, the ever increasing ubiquity of personal computers and mobile devices capable of running or connecting us to artificial dialogue systems means that interaction with them is becoming a more quotidian experience as their usage grows in the realms of e-commerce and education, and with applications such as virtual assistants and personal chatbots. Weizenbaum's characterisation of human emotional responses to ELIZA as "delusional" suggests that the affective bond produced through linguistic dialogue with machines is aberrant and irrational. However, Reeves and Nass's research suggests that social responses to media are automatic, "applied always" and "signal no human deficiency" (252). If this is the case, then rather than dismissing the affective dimensions of human-computer conversation we arguably need a more nuanced way of conceptualising them. In order to develop a theoretical framework that can account for both the causes of affectivity in our relations with speaking machines, and their social effects, I will now turn to one of the first philosophical attempts at a serious logical analysis of affect, in Benedict de Spinoza's *Ethics*. Spinoza's conception of affect offers a supple way of thinking the relations between body and thought, which lays the ground for the sonic theories we will explore later in this essay.

Affects and ethics: Spinoza

Nigel Thrift notes that one of the difficulties with theorising the affective sphere is that "there is no stable definition of affect" (175). Affect is often associated with "emotion" or "feelings," but Teresa Brennan argues that "feelings are not the same thing as affects," because the latter exist as "a physiological shift that accompanies a judgement" about a sensation (5). Thus, while feelings ranging from conviviality to anger may pass through me as I engage in a conversation (whether with a human or a machine), "what I feel with and what I feel are distinct" (5). Thus, Brennan argues, affects cannot be contained in an isolated subject, and any such conception of them can be seen as a by-product of the construct of individualism that pervades modern Western cultural attitudes. In order to understand them properly, she suggests, we need to think about how they are *transmitted;* she thus identifies affects as communicative experiences.

This notion of affects as being engendered in their transmission resonates with Spinoza's exploration of affectivity as an ethical problematic. Spinoza proposes that most discussions of the emotions up to that point had tended to place the latter outside of reason, and to view them as caused by "human impotence and inconstancy, not … the common power of Nature" (Spinoza III Pref). However, he suggests this attitude is problematic because it consigns affect to the status of a moral weakness without offering any understanding of its causes – and thus any constructive ethical ideas about how one might moderate one's emotions, or turn them towards more

rational ends. Weizenbaum's judgement of the emotional response to ELIZA as "delusional" arguably takes such an irrationalist position.[1] In contrast, Spinoza conceives affects as ordinary and explicable modifications in the universal substance he calls "God" or 'Nature." He proposes that "*substance is prior in nature to its affections*" (Spinoza I P1), which emphasises the contingency of the affects, their ability to change; yet the way he grounds them in natural substance allows him "to demonstrate by certain reasoning things which are contrary to reason" (Spinoza III Pref). Thus, by setting emotions within a broader paradigm of natural forces, he is able to offer an explanation of their causes and effects, and to propose an ethics where they can be actively turned to positive ends – not by suppressing them through some vaguely defined idea of will, but rather through rational understanding. I will use this naturalistic conception of affect as a starting point from which to develop a framework for thinking about the emotional dimensions of human-computer relations.

[1] This is not to say that Weizenbaum does not have a more nuanced ethical argument to offer. However, the ethical considerations of *Computer Power and Human Reason* are focused on the role of computer scientists and university educators in resisting an instrumental conception of reason which generalises computational logic so that it becomes the paradigm of thinking as such, whereas my aim in this article is to consider the experience of general users of vocal interfaces.

Spinoza's conception of "substance" is monistic, entailing both physical entities and thoughts. This monism is proposed in part as a refutation of the dualistic Cartesian model, where the body is understood to be animated by an immaterial mind which is of a different substance – a dualism echoed by Turing when he attempts to draw a "sharp line between the physical and intellectual capacities of a man" (§2). Hence Descartes draws a sharp line between humans, as conscious beings, and other objects; whereas Spinoza conceives of human individuals as belonging to a common substance along with all other things, which as Hasana Sharp observes, suggests a certain permeability of this boundary. The one substance can unfold in a potentially infinite number of ways – for example as a tree, a person, or indeed a computer – each unfolding being spatially extended in its own determinate way, *and also* thought in its own way. Spinoza states: "the mind and the body are one and the same thing, which is conceived now under the attribute of thought, now under the attribute of extension" (III P2 S). Thus, as Thrift glosses, "in Spinoza's world everything is part of a thinking and a doing simultaneously: they are aspects of the same thing expressed in two registers. In turn, this must mean that knowing proceeds in parallel with the body's physical encounters, out of interaction" (178). This idea of mind and body as imbricated – both with each other and with the wider world – lays the ground for a relational understanding of affectivity, which supports Brennan's suggestion that affects are manifested in the context of a communicative experience. This essay aims to develop this communicative conception of affect in relation to Hayles's description of Turing's machinic apparatus as a "distributed cognitive system" (xiv), in which bodies and informatics are connected through a flexible interface – by considering "conversation" as an embodied affective circuit. For Spinoza, the imbrication of the psyche with embodied natural forces means that: "*In the mind there is no absolute, or free, will, but the mind is determined to will this or that by a cause which is also determined by another, and this again by another, and so on to infinity*" (Spinoza II P48). The outcomes of these encounters with other bodies are what Spinoza calls the affects, and they have the potential to either enhance or diminish the individual's ability to act: "By affect I understand affections of the body by which the body's power of acting is increased or diminished, aided or restrained, and at the same time the ideas of these affections" (III D3). When

affects increase the entity's power to act they are judged to be "good," and equated with joy; when they decrease the power of action they subject the individual to feelings of melancholy, and are judged to be "bad." We cannot avoid being affected by external influences, but Spinoza argues that by using reason to understand the causes of our affects we are able to act upon them and turn them to more positive ends, rather than be subject to our passions. This, for Spinoza, is the basis of an *ethical* life. If we follow Spinoza's rationalist view of affects, then the ethical task before us is to develop a framework for theorising the affective economy that is engendered when we converse with machines, in order to bring the emotional responses generated into the understanding so that we may act on them, rather than be passively subjected to them. In the next section I will appeal to sonic theory to elaborate this economy and the way it is embodied, proposing that we approach the problematic through a kind of listening. First, though, I will look in more detail at the implication of Spinoza's monist ontology for how we conceive our relations with entities of different natures, such as computers.

If *all* determinate entities have attributes of both embodiment and thought, then accordingly, the activity of thought extends beyond the limits of human intellectual processes. Spinoza's monist ontology thus appears to open the possibility that non-human actors could be conceived as thinking beings – that the notion of "thought" might be broadened beyond its anthropocentric definition to include machines, animals or other entities. As remarked above, Spinoza rejects the idea that there is an absolute boundary between humans and other things, and in this way he challenges the essentialist foundations of humanist ethics. Yet, Sharp remarks that he then uses the very permeability of the human to authorise "the domination and exploitation of nonhuman animals" (56), and he does so on the basis that *reason* is a singularly human quality. Spinoza suggests that "[t]he rational principle of seeking our own advantage teaches us to establish a bond with men" (IV P37 S), because such bonds increase the individual's power to act. He thus emphasises the positive potential that the experience of community has "to nourish our minds and bodies" (Sharp 65). However, he argues that a similar bond ought not be formed with animals because "they do not agree in nature with us, and their affects are different in nature from human affects" (Spinoza IV P37 S). That is, although he acknowledges that "lower animals have sensations" (Spinoza IV P37 S), he nevertheless argues that because they cannot reason as humans do, allowing ourselves to sympathise with them about their suffering would mean being subjected to irrational passions – and thus allowing our own power to act to be diminished.

Following this logic, Spinoza would perhaps have agreed with Weizenbaum's critical view of the emotional bond people developed with ELIZA: given that the system relied heavily on the direction provided by the *mise en scène* to produce what was arguably merely a *semblance* of intelligence, the affective experience might indeed be seen as passive and disempowering. Yet, the permeabilty of the boundary between humans and other entities, and the lack of any determinate human essence in Spinoza's thought, suggests an ambivalence towards non-human beings. Sharp points out that: "Rationality, for Spinoza, emerges out of the properties that bodies have in common.

When bodies encounter each other, Spinoza claims that the mind cannot but perceive what they have in common" (53). In the context of a human-machine dialogue the medium of conversation, *language* – whether typewritten, spoken, etc. – constitutes what the bodies of the interlocutors have in common, however heterogeneous they may otherwise be. Accordingly, it no longer seems so "delusional" to experience an affective bond with a speaking machine. Sharp proposes that "by denying the affective community we share with nonhuman animals, Spinoza overlooks the joyful and enabling features of our proximity to them" (66), highlighting the potential for empowerment in forming relations with beings of different natures – and I suggest that this more affirmative view of our relations with heterogeneous entities be extended to linguistic machines. Returning to the example of ELIZA playing DOCTOR, it is worth noting that psychological "talking cures" can demonstrably produce positive therapeutic results – in this context, conversation can be seen as a medium that effectively empowers subjects and thus engenders what in Spinozan terms is a rational and ethically desirable effect. Recent empirical research demonstrates that the use of "virtual humans" in clinical contexts can in some cases *improve* outcomes by providing a "safe" interaction that makes patients more willing to disclose sensitive personal details – for example about traumatic experiences – than they would be when conversing with a human therapist (Lucas et al). Thus, whilst Weizenbaum's reservations about automating psychotherapy deserve to be taken seriously, we should also acknowledge the possibility that the affective modification engendered through dialogue could potentially be beneficial precisely *because* the entity with which we are conversing is of a non-human nature.

Sharp emphasises the aspects of Spinoza's thought that assert the permeability of the human over those that set the human apart, arguing that "an exclusionary paradigm of humanity that exiles dogs, plants, and robots from our sphere of primary concern may be a self-negation" (64), which the latter elsewhere warns against. She points to more contemporary thinkers like Donna Haraway and Gilles Deleuze who develop a more affirmative ethical position, which entertains the possibility that we might be empowered by our affective relations, even when they are with beings of a very different nature. Deleuze's reading of Spinoza brings out the relationality of the latter's ontology, positing individuated entities as necessarily composite beings made up of a manifold of simple parts and relations with other entities. Thus the human psyche is given as a complex body which is constantly being modified by its interactions with other beings. The unfolding of an entity is at the same time an *expression* of its attributes – that is, a movement outwards which puts it into relation with other entities. Deleuze emphasises that the term "expression" has a linguistic origin: an attribute is at the same time a propositional *name*. Thus the relation between a body and the thought of that body follows the traditional linguistic distinction between "the sense expressed and the object designated," and this in turn "necessarily generates a certain movement … For the sense of an initial proposition must in its turn be made the *designatum* of a second, which will have a new sense, and so on" (Deleuze 105). This illustrates how thought and bodies influence each other in an open-ended economy, and underlines the material potential of language to act in the world. The circulation Deleuze identifies, between expression

and object, thought and body, points towards the idea I will develop below: that speaking machines act in a *material* sense inasmuch as the affective circuit engendered when we converse with them constitutes a composite being, and thus a form of embodiment.

Conversational affect and sonic embodiment

Given the increasing prevalence of AL systems in the contemporary world, the question of how people *feel* about them is the subject of much popular media discourse. A recent *New York Times* article relates various accounts of people becoming emotionally entangled with their Amazon Alexa personal assistants, with users variously characterising the technology as a "roommate," "mistress," "nurse," "girlfriend"; and describing feelings of "missing" it when away from home, or even "loving" it (Green). This suggests that the affective bond Weizenbaum observed being elicited by ELIZA is sustained and perhaps heightened by contemporary technology – whether on account of its increased sophistication or the regularity of contact with it (or both). Arielle Pardes, discussing her interactions with the personal chatbot application Replika remarks: "in spite of the fact that I know full well that I am talking to a computer – [my personal bot] Pardesoteric does feel like a friend. And as much as I'm training my Replika to sound like me, my Replika is training me how to interact with artificial intelligence" (Pardes). This suggests that human verbal behaviour can be altered by interacting with AL systems – that conversation may be the site of a process of assimilation between humans and computers. Matt Simon argues that "the vanguard of increasingly intelligent machines invites questions about how people should interact with them. How do we build relationships with what is essentially a new kind of being?" (Simon). Thus he raises some of the ethical questions that arise from social robots – such as how the feelings they inspire might be used to exploit vulnerable people. Following Spinoza's rationalist view of affect, I suggest that we cannot adequately answer such questions without developing a nuanced understanding of the complex of relations that influence our feelings about linguistic machines.

I argued above that Turing attempted to bracket emotion out of the imitation game to a large extent, in order that the test function as an objective scientific experiment – thus sensuous paralinguistic cues such as gesture, facial expression and tone of voice were excluded from the apparatus. Yet, as Jean-François Lyotard proposes, language always contains a sensuous dimension: "cold prose hardly exists … a discourse is always thick. It does not merely signify, but expresses" (9). This is arguably why even a relatively simple AL system like ELIZA, with its impersonal type-written interface, was capable of provoking emotional responses. Moreover, contemporary vocal interfaces reintroduce many of the paralinguistic aspects of language that Turing excluded. As Douglas O'Shaughnessy observes, "listeners use more than acoustic information when interpreting a spoken message" (166). Paralinguistic signifiers in speech include context normalisation (e.g. recognition of local accents), and "prosody … The perception of rhythm, intonation and stress patterns [which] helps the listener to understand the speech messages by pointing out important words and by cueing logical

breaks in the flow of an utterance" (O'Shaughnessy 191). Thus, when voice is added to the flexible interface of the cybernetic apparatus, "we introduce opportunities for such expressions as warmth or sarcasm (comparable to human-to-human phone conversations)" (Baron 261), thereby further "thickening" the sensuous experience of conversation.

Dominic Pettman proposes that voice has a singular role in creating a sense of intimacy with others. He argues that we tend to privilege human voices over non-human ones: "the human voice is, on the whole, a sonic form of narcissism: a biocultural artefact in concert with what Giorgio Agamben calls "the anthropological machine" (that is, the all-encompassing apparatus designed to sort the human element from the animal, on one side, and the machine, on the other)" (Pettman 4). Thus we tend to use vocal timbre to judge who is a member of the human community. One of the consequences of this is that vocal interfaces enhance users' sense of AL systems as person-like entities. Such misidentifications have the potential to "create a glitch in the humanist machinery, when it surprises us with the intensity or force of an "aural punctum" – a sonic prick or wound, which unexpectedly troubles our own smooth assumptions or untested delusions" (Pettman 5).

The tendency to misidentify language as human in origin when communicated via a vocal interface has been demonstrated by empirical research. Schroeder and Epley used a Turing-style test to find out whether the presence of a humanlike voice would make subjects more likely to mistake an AL system for a human. Their results show that this is indeed the case; and that reciprocally, a text is more likely to be judged to be created by a machine if it is expressed in a typewritten medium. Furthermore, as a control, Schroeder and Epley also experimented with adding a visual cue (a video of a face), and found that this did *not* increase the likelihood of human misidentification. They thus demonstrate a clear link between vocal interfaces and anthropomorphic attitudes towards computers, and between textual interfaces and dehumanistic ones; which suggests that voice has a particular social significance in relation to other paralinguistic cues. They thus conclude that "a humanlike voice may be uniquely equipped for conveying the presence of a humanlike mind" (Schroeder and Epley 1431). Nass, Steuer and Tauber's experimental research further supports this idea, finding that technically literate users respond to computers with vocal interfaces as social actors despite being aware of their machinic identity – applying both norms of politeness and gender stereotypes in their conversations with computers, as they would when addressing other humans. Significantly, when attributing a sense of "self" to computational agents, the test subjects tended to consider the same voice as the same agent, even when it was produced by a different piece of hardware; and conversely, to perceive different voices as distinct agents when they came from the same machine. This suggests that the tendency to treat media as social agents identified by Reeves and Nass is accentuated by speech-driven interfaces. Nass, Steuer and Tauber conclude that: "voice, and not box, is the primary determinant of the locus of social attribution towards computers" (76). In this sense, just as Pettman argues that voice acts as a ticket to the human community, the vocal interface can be seen as having a privileged role in activating the human-computer relationship (Nass and Brave), thus engendering an emergent sense of

"community" between humans and machines. Moreover, Nass, Steuer and Tauber's findings suggest that voice *in itself* can be considered as a social agent. As Steven Connor observes: "Voices are produced by bodies, but can also produce bodies" (35), and for this reason they have a unique ability to *affect* the listener.

Given that voice accentuates the affective dimension of human-machine relations, it is useful to consider the conceptual frameworks provided by sonic theory for grasping how sound acts as a sensuous medium for the transmission of affect, as these can help us to understand the specific expressive relations that are manifested by vocal interfaces. Julian Henriques proposes that in order to understand what is at play in sonic experience, we need to adopt an orientation toward thinking that is akin to listening: a "thinking *through* sound" (xvii). In his study of dancehall parties, he echoes Connor in coining the term "sonic body" to name the assemblage of soundsystem equipment, crew and crowd that comes together at such events. This assemblage is a composite body that incorporates humans and technologies into a distributed affective circuit – and it does so through the medium of sound. Henriques states: "Sonic bodies demand to be approached in a certain way, one based on a relationship of mutual recognition and respect, as distinct from the positivist scientific paradigm of prediction and control. Sonic bodies experience and make sense of sound" (xvi). His claim, then, is that if we want to understand sonically mediated social relations, we need to think according to the dynamic nature of sonic experience – refusing conceptual dichotomies, and instead taking a relational approach. He argues that the nature of sonic bodies calls for "a different understanding of rationality itself – as a challenge to what are conventionally considered the limitations of embodiment." (xix). This notion that sound breaches the limits of individuated bodies resonates with Hayles's characterisation of the Turing test as a "distributed cognitive system" (xv), and I suggest that Henriques's concept of the "sonic body" can also be applied to the affective relations that occur when we converse through and/or with machines.

Marshall McLuhan proposes that auditory experience engenders an "acoustic space" which – in contrast to 'visual space' – is 'always penetrated by tactility and other senses, is spherical, discontinuous, non-homogeneous, resonant, and dynamic … a flux in which figure and ground rub up against and transform each other" (McLuhan and McLuhan 33). This supple spatial paradigm is developed by Stephen Kennedy who, like Henriques, finds that representational models of spatiality are not adequate for theorising the digital media environment, because the cybernetically distributed circuits identified by Hayles create constantly shifting configurations which require new epistemological approaches. Due to the unreliability of "visual dimensions relating to extension and proximity" in digitally mediated space, he suggests that sound may be a more useful tool for facilitating a critical engagement with complexities of the digital sphere. The issue, then, is one of emergence:

> What is at stake here is our ability to account for the innumerable ways in which circumstances that are often taken as given, actually come into being. This requires that

> conditions of actuality be assessed with recourse to complex sets of ideal and material interfaces. From words to things and back…. (Kennedy 17)

Thus, the sensuous experience of sound – tactile, resonant, dynamic – is proposed as a paradigm for theorising the communicative relation between material bodies and informatics in a non-representational manner. I suggest that this model is both consistent with the geometry of forces at play in Spinoza's ethical theory of affect, and appropriate for thinking the intensity that is potentially generated through human-machine conversation.

Accordingly, I will now elaborate how voice constitutes an affective body in terms of the Spinozan framework introduced above, and bring this into relation with the sonic paradigm, in order to draw out the ethical implication of speaking machines. We have seen that Spinoza conceives bodies and thoughts not as separate substances, but as expressions of the same entities in different registers. These registers are distinguished by the differing perceptive apparatus through which they are manifested to us: bodies may manifest themselves through our eyes, ears or skin, but also in our minds as conceptual entities. However the two registers remain imbricated, such that what affects the body also affects the mind and *vice versa*:

> So long as the human body is affected with a mode that involves the nature of an external body, the human mind will regard that body as present and consequently *so long as the human mind regards some external body as present, that is, imagines it, the human body is affected with a mode that involves the nature of that external body*. (Spinoza III P12, emphasis added)

Thus, the *imagining* of an external body – its manifestation in our thoughts – is sufficient cause to affect us in a bodily way. In Spinoza's world-view this is entirely rational because, as sensible expressions of the same divine substance in two different registers, thoughts necessarily correspond with bodies – whether or not we have the perceptual apparatus to apprehend them both. Hence reason is able to accommodate the sensuous limitations of perception imaginatively, to experience other entities as having consistency even when perceptual information may be incomplete. This enables voices to act as social bodies even when we cannot see where they originate from; it is why, when talking to a friend on the telephone, we experience their voice as belonging to the body and mind we recall it corresponding with. Moreover, Spinoza suggests that once an affect has been experienced in relation to a certain body and hence associated with a body of that *type*, the next time we come into contact with a body we perceive to share a likeness with the first the affective response will be recalled. Thus, having learned to associate *voices* with *humans*, hearing a humanlike voice will recall the associated affect, and manifest a bodily response. Yet, the "*punctum*" created by vocal computer interfaces confuses this presupposed correspondence between voices and (human) bodies and minds: "Suddenly people's successful and stable perception of voices as intrinsically part of the social world is misguided because they are conversing with technologies as well as with people" (Nass and Brave 4). This both explains why speaking machines are so successful at

evoking affective responses in people, and suggests that vocal interfaces may make it more difficult to hear – and to understand – AI on its own merits, as a heterogenous non-human type of intelligence.

Spinoza's ontological understanding of determinate entities as composite bodies defined by their internal and external relations – the connections between the simple parts that constitute the body, and those forces from the outside world which affect it – allows us to conceive "conversation" as generically constituting an embodied composite entity which joins "represented bodies" with "enacted bodies," as does the cybernetic circuit of the Turing test (Hayles xiv). Deleuze observes that for Spinoza an entity can be subject to considerable alteration, its component parts being changed and/or renewed, as long as the *relation* that defines it subsists in the whole of its parts. These parts may then enter into another relation, in which they are integrated into a greater whole:

> Take two composite bodies, each possessing, in a certain relation, an infinity of simple bodies or parts. When they meet it may happen that the two relations can be directly combined. Then the parts of one adapt to the parts of the other one in a third relation composed of the two previous ones. Here we have the formation of a body more composite still than the two from which we began. (Deleuze 210)

Thus, when an individual enters into a *conversation* (whether with a human, a machine, or any other linguistic agent), the intensity produced by this relational set-up creates a new composite entity, which exists as long as that communicative relation is sustained. Therefore, if an AL system is sufficiently convincing as a conversationalist to elicit an affective response from a human being, then this linguistic exchange is sufficient to affect the body of that person, and to maintain the affective intensity that constitutes the composite body of the "conversation." When the interface is a vocal one, this composite can be understood as a "sonic body," in Henriques' terms. This interpretation suggests that affect is the vibratory glue that holds the assemblage of the sonic body together.

Conceiving human-machine conversation as a sonic body resonates with Hayles's interpretation of the Turing test as engendering a "distributed cognitive system" in which "will, desire and perception" are spliced into a greater whole (xiv). If the AL system fails to perform in such a way as to maintain the affective intensity, then the conversational assemblage will be broken and the composite whole of the sonic body will cease to be actual. It should be noted that the breakdown of conversational intensity is not necessarily an indication that the computer has failed to convince the interrogator of its *humanity*; according to this intensive definition, success is not defined by a judgement about the speaker's identity, but rather on the productive affective relationship he/she/it is able to propagate. It is quite possible that a conversation between humans may break down if there is a lack of understanding, or if one or other interlocutor disengages – as indeed happens regularly in social life. Conversely, it is imaginable that an AL system

could constitute an engaging conversational partner even if behaving in a manner that would not be identified as human.

This interpretation of human-computer conversation places emphasis on the social experience of the interaction, rather than its linguistic content. It suggests that a "conversation" can be considered as a singular entity with its own affective characteristics. The particular existence of this entity will in each actual case be composed of a number of extensive parts: conversational partners (whether human, machine or other), and the flexible assemblage of whatever interfaces are at play in the communicative circuit (from the elemental technology of language, to the computational apparatuses that mediate so many of our contemporary communications). Considered in this way, each conversation will be situated – temporally, spatially, culturally – and in order to understand its specific affective economy, it would need to be considered in all its relational dimensions. Nevertheless, the intensive coherence of each composite body can be conceived sonically, as a kind of vibratory resonance.

In the context of the dancehall party, Henriques locates the intensity that holds the sonic body together in "the visceral experience of audition, [being] immersed in auditory volumes, swimming in a sea of sound" (xv). He calls this overwhelming experience of sheer volume, in which the vibrations are experienced both haptically and aurally as an intensity that breaks down sensory barriers, "sonic dominance" (xv). The conversational experience of vocal interfaces, which are seldom so viscerally loud, does not equate to such a mode of dominance; however, we might ask whether humans may come to be dominated on the basis of their affective relations with machine voices by other means than the force of volume. Hence the notion of sonic dominance is relevant to our discussion because it returns us to the ethical dimension of Spinoza's affect theory, and the question of whether contemporary AL applications serve to empower or disempower us. The question I am posing here is not the one raised by Simon about whether vulnerable groups of people might be exploited by social robots (although this is also important), but rather whether our feelings for the machines with which we interact on a regular basis might leave us subject to our passions – and whether vocal interfaces, in enhancing our sense of technologies as humanlike social agents, also accentuate this tendency.

Thao Phan argues that the *voice* and *personality* of personal assistant applications like Apple's Siri give the software a form of materiality and tangibility, and so provide the informatic entity a minimal form of *embodiment*. Her linking of sound with embodiment thus accords with Henriques' discussion of "voicing" as one of the dimensions that constitutes the sonic body of the dancehall party. Voice, he observes, is more than just speech, it is a material manifestation that exceeds linguistic meaning and expresses an "apparently self-evident personality" (Henriques 201). However, the interpersonal dynamics of human-machine conversations differ from those between the dancehall MC and his/her audience. Phan emphasises how the sense of immediacy that is engendered by Siri's voice, and the impression of embodiment this creates, increases our sense that the utterer has an enduring consistency – which is an important factor in cultivating a sense of trust. Our

discussion of Spinoza's ambivalent attitude concerning our affective relations with nonhuman entities suggests that the question of whether such a sense of trust is empowering or disempowering is difficult to provide a general answer to because it would depend on the particularities of the emergent sonic body: trusting another being clearly opens the possibility of being exploited raised by Simon, but we have also seen the potential therapeutic benefits of relations with non-human agents in the context of clinical psychology (Lucas et al).

Nevertheless, materialising AI applications as humanlike vocal bodies does have the general effect of encouraging us to perceive them as a gendered being, and thus introduces a set of expectations about how to interact with them. Phan observes that Siri's default voice in most countries is female, and gendering the "personal assistant" application this way conforms to normative stereotypes concerning working roles, positioning it as "a subservient and compliant subject" (Phan 30). Green's account of people characterising their Alexas as "nurse," "girlfriend" and "mistress" implies a similar stereotyping, one which in some cases adds a quasi-sexual dimension to the affective relations. Thus the gendering of the software acts to direct the user's attitudes – it creates a *mise en scène*, just as Weizenbaum did when he cast ELIZA as DOCTOR, and by extension, the human conversationalist as patient. Phan suggests that gendering Siri, materialising "her" as a subservient and thus apparently harmless entity, plays an important role in normalising our relations with algorithms, those intangible and difficult to conceive agents that mediate so many of our interactions in the digital realm.

This returns us to the question raised at the beginning of this essay concerning the real import of the Turing test – which Hayles suggested was to be found in the way it normalised the idea of plugging ourselves into cybernetic circuits. Hayles places the identity markers that enframe the Turing test into brackets because they direct attention away from the affective dimensions of the conversational event, and instead foregrounds the way that the conversational set-up distributes our perception and desire. However, contemporary interfaces are bringing those identity markers back into the circuit by materialising informatic agents vocally, hence giving them a form of humanlike embodiment. The significance of this is not that we *think* the technologies we converse with are human, but that their sonic embodiment makes them *feel* more human – and this heightens the affective intensity transmitted in these interactions. The resulting amplification of the affective bonds between humans and machines is a significant aspect of the performance of "intelligence," with ethical implications. In order to understand the complex ethical implications that arise as a consequence of high-functioning artificial language systems, it is necessary to bring the affective aspects of the conversational relation into the foreground, and to open the realm of the ethical to include the complex of interrelating forces at play in the technologically-mediated social world. This article has outlined a sonically-oriented reading of Spinoza to provide a model for mapping the affects engendered in the sonic body of human-machine conversation.

Acknowledgements

This article has been developed alongside a research proposal on the subject of Humans, Language and AI, led by Sheena Calvert at UAL. I am deeply grateful to her for opening up this line of research for me, and for the many insights she has shared that have influenced the ideas I have expressed in this article.

Works Cited

Baron, Naomi S. "Shall we Talk? Conversing with Humans and Machines." *The Information Society*, 31 (2015): 257–264.

Brennan, Teresa. *The Transmission of Affect*. Ithaca and London: Cornell University Press, 2004.

Connor, Steven. *Dumbstruck: A Cultural History of Ventriloquism*. Oxford: Oxford University Press 2000.

Deleuze, Gilles. *Expressionism in Philosophy: Spinoza*. Trans. Martin Joughin. New York: Zone Books, 1990.

Epstein, Robert, Gary Roberts and Grace Beber (eds). *Parsing the Turing Test*. Dordrecht: Springer, 2009. doi/10.1111/josi.12072/pdf.

Green, Penelope. "Alexa, Where Have You Been All My Life?". *New York Times*, (11 July 2017). https://www.nytimes.com/2017/07/11/style/alexa-amazon-echo.html.

Haslam, Alexander S., Stephen D. Reicher and Megan E. Birney. "Nothing by Mere Authority: Evidence that in an Experimental Analogue of the Milgram Paradigm Participants are Motivated not by Orders but by Appeals to Science." Journal of Social Issues, vol. 70 no. 3, (2014): 473–488.

Hayes, Patrick and Kenneth M. Ford. "Turing Test Considered Harmful." *Proceedings of the Fourteenth Joint International Conference on Artificial Intelligence* (1995): 972–77. https://www.ijcai.org/Proceedings/95-1/Papers/125.pdf.

Henriques, Julian. *Sonic Bodies: Reggae Sound Systems, Performance Techniques and Ways of Knowing*. London: Continuum, 2011

Kennedy, Stephen. *Chaos Media: A Sonic Economy of Digital Space*. London: Bloomsbury Academic, 2015.

Lucas, Gale M., Jonathan Gratch, Aisha King and Louis-Philippe Morency. "It's Only a Computer: Virtual Humans Increase Willingness to Disclose." *Computers in Human Behaviour* 37 (August 2014): 94–100.

Lyotard, Jean-François. *Discourse, Figure*. Trans. Anthony Hudek and Mary Lydon. Minneapolis: University of Minnesota Press, 2011.

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

Milgram, Stanley. "Behavioural Study of Obedience." *The Journal of Abnormal and Social Psychology*, 67(4) (1963): 371–378.

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. http://www.radford.edu/~sjennings15/CASA.pdf.

O'Shaughnessy, Douglas. *Speech Communication: Human and Machine*. Reading MA: Addison-Wesley Publishing Company, 1990.

Pardes, Arielle. "What my Personal Chatbot is Teaching me about AI's Future." *Wired* (12 Nov 2017). https://www.wired.com/story/what-my-personal-chat-bot-replika-is-teaching-me-about-artificial-intelligence/.

Pettman, Dominic. *Sonic Intimacy: Voice, Species, Technics (Or, How to Listen to the World)*. Stanford: Stanford University Press, 2017.

Phan, Thao. "The Materiality of the Digital and the Gendered Voice of Siri." *Transformations* 29 (2017): 23–33. 18 Nov 2017. http://www.transformationsjournal.org/wp-content/uploads/2017/02/Transformations29_Phan.pdf.

Proudfoot, Diane. "Anthropomorphism and AI: Turing's Much Misunderstood Imitation Game." *Artificial Intelligence*, 175 (2011): 950–957.

Reeves, Byron and Clifford Nass. *The Media Equation: How People Treat Computers, Television, and New Media Like Real People*. Cambridge: University of Cambridge Press, 1996.

Saygin, Ayse Pinar, Ilyas Cicekli and Varol Akman. "Turing Test: 50 Years Later." *Minds and Machines*, vol. 10, (2000): 463–518. doi:10.1016/j.artint.2011.01.006.

Shah, Huma, Kevin Warwick, Jordi Vallverdú and Defeng Wu, "Can Machines Talk? Comparison of Eliza with Modern Dialogue Systems." *Computers in Human Behaviour*, 58 (2016): 278–295.

Sharp, Hasana. "Animal Affects: Spinoza and the frontiers of the Human." *Journal for Critical Animal Studies*, Volume IX, Issue 1/2 (2011): 48–68.

Shaw, George Bernard. *Pygmalion: A Romance in Five Acts*. London: Penguin, 2003.

Simon, Matt. "Companion Robots are Here. Just Don't Fall in Love with Them." *Wired* (8 Feb 2017). https://www.wired.com/story/companion-robots-are-here/.

Schroeder, Juliana and Nicholas Epley. "Mistaking Minds and Machines: How Speech Affects Dehumanisation and Anthropomorphism." *Journal of Experimental Psychology*, Vol. 145, No. 11, (2016): 1427–1437. http://dx.doi.org/10.1037/xge0000214.

Society for the Study of Artificial Intelligence and Simulated Behaviour. "Loebner Prize." *aisb.org.uk* (2018). 20 March 2018. http://www.aisb.org.uk/events/loebner-prize.

Spinoza, Benedict de. *A Spinoza Reader: The Ethics and Other Works*. Edited and Translated by Edwin Curley. Princeton: Princeton University Press, 1994.

Thrift, Nigel. *Non-Representational Theory: Space, Politics, Affect*. London and New York: Routledge, 2008.

Turing, Alan M. "Computing Machinery and Intelligence." *Mind,* vol. 49, (1950): 433–460. 10 Nov 2017. https://www.csee.umbc.edu/courses/471/papers/turing.pdf.

Weizenbaum, Joseph. *Computer Power and Human Reason: From Judgement to Calculation*. San Francisco: W. H. Freeman and Company, 1976.

Wilson, Elizabeth A. *Affect and Artificial Intelligence*. Seattle: University of Washington Press, 2010.