

## Artificial Participation: An Interview with Warren Sack\*

Joseph Dumit

**Preface:** Recently I was talking with a friend, relating an incident that had happened to me when my word processing program froze, when I realized with a start that I had had the exact same conversation, word for word, with another friend, Warren Sack, the day before. The weird thing was, I could not tell whether I was causing these two friends to say the same things in response to me, or whether they were coaxing the same words out of me. Of course the conversation stopped being the same the moment I froze, and I've never been able to repeat the conversation.

What follows is a reconstruction of an interview I conducted with Warren Sack, a software designer and media theorist presently completing a Ph.D. at the MIT Media Laboratory. Sack makes computer-based social technologies that make you stop and think about how much you are invested in automatic ways of conversing. The computer program ELIZA has been doing this since 1966, but as Sack discusses below, programs that ask questions rather than answer them never became a focus of the artificial intelligence community. Sack has therefore located ELIZA's kin among performance artists, improvisational actors, ethnomethodologists, and Socrates, the great questioner, also known as the great simulator.

Sack has a history of putting philosophy and cultural studies into computer programs. He has a B.A. in Computer Science and Psychology from Yale University and worked there in the Artificial Intelligence Laboratory for three years. With colleagues at Yale he designed and implemented a computational environment – based on a touch-screen and PC – for preschool children to count and to do simple math. He also designed computer-based tutoring programs to teach undergraduate programming. After Yale, he studied philosophy and artificial intelligence for a year at the University of Paris VIII, Vincennes at St. Denis followed by three years in Santa Cruz, California co-organizing a cultural studies of science and technology working group. From 1992 until 1994, at the MIT Media Laboratory, he developed two systems for automatically editing digitized video, and experimented with text-based, automatic story generation and analysis systems. He is a member of the Machine Understanding Group at the MIT Media Laboratory and a member of the Interrogative Design Group at the MIT Center for Advanced Visual Studies.

Lately, he has been working on a computer program to generate questions and function as an “artificial participant” in on-line, electronically-mediated conversations. I decided that this program was the ideal site to interrogate Sack on the possible roles of computer programs and other social technologies for posing to us the problem of the zero level of conversation. As you will see in the dialogue that follows, there are many social zeros that we operate with. The program that Sack is working on is called SimSocrates.

---

\* Appears in *Zeroing in on the Year 2000: The Final Edition (Late Editions, 8)* George E. Marcus, Editor (Chicago: University of Chicago Press, 2000).

**WARREN SACK:** Hello?

**JOSEPH DUMIT:** Hi Warren, I'm interested in talking to you about your work as a media technology designer. In particular, you have begun to elaborate an interest in designing a kind of artificial intelligence that asks questions rather than answers them. AI brings to my mind its basis as a series of binary codes, a long string of 0's and 1's that are acted upon by a Turing machine (a computer usually) that somehow are able to interact with us intersubjectively.

**Sack:** Well, to start with, for a number of reasons I do not like the term 'artificial intelligence', nor for that matter, any notion of the intentions, social or otherwise of software programs. But rather than repeat myself and distract us into well-trodden territory, you can read these objections in my on-line papers.

Instead, I'd like to start over and explore with you a baseline notion of the zero degree of conversation.

**DUMIT:** That sounds interesting (too). Okay, what do you mean by zero here?

**SACK:** Actually, I'd prefer for you to make up a definition.

**DUMIT:** (surprised) What?

**SACK:** Any definition, make up a silly one.

**DUMIT:** (in despair) I can't, this is for publication.

**SACK:** Suppose I think of one and you guess what it is.

**DUMIT:** (relaxes a bit)

**SACK:** I've thought of one. But I'll only answer "Yes" or "No".

**DUMIT:** Okay. Interesting enough. So I just ask questions about your definition and you'll tell me whether I'm right or wrong?

**SACK:** Yes.

**DUMIT:** I'd like to start by asking if your zero is related to Barthes' notion of *Writing Degree Zero*<sup>1</sup>, or the search for a style in which there is no subjectivity embedded in the zero?

**SACK:** Yes.

**DUMIT:** Good, well perhaps then your zero has something to do with the fact that artificial intelligence projects like those of the Yale AI lab that you first worked in, store their knowledge about the world in binary data structures. This would be akin to Roman Jakobson's notion of the marked and the unmarked, a degree zero of phonology, morphology, and so on.<sup>2</sup>

**SACK:** No!

**DUMIT:** Hmm, this is harder than I thought. I'm not sure whether you meant no to the first part of my comment or the latter or both. It is quite hard to pin down what you mean by No.

**SACK:** (interjecting) Yes!

**DUMIT:** Weird. As you interrupted me, I had this flashback to Jakobson's comments regarding running into a former actor of Stanislavskij's Moscow Theatre. The actor was told by the director to take the same sentence and express it emotionally forty different ways, only via the sound shapes. Jakobson used a recording of this actor in his research on Standard Russian. And that raises the question: Does your zero degree of conversation have something to do with Jakobson's critique of the individualism

---

<sup>1</sup> Roland Barthes. *Writing degree zero*. Translated from the French by Annette Lavers and Colin Smith (London: Cape, 1984).

<sup>2</sup> Developed in dialogue with Trubetzkoy. Roman Jakobson. "The Concept of Mark" in *On Language*. Edited by Linda R. Waugh and Monique Monville-Burston (Cambridge, MA: Harvard University Press, 1990, page 136).

embedded in much of Saussure's notion of 'parole'. That, via Volosinov, parole is social. And via Sechahaye, after the speaker speaks, the sounds are 'taken in as is by the hearer, who subjects it to his analysis and interprets it in order to understand it.'? And therefore there is no zero degree of speaking, only of speaking with others, of conversation, of interactivity?<sup>3</sup>

**SACK:** (sarcastically) Yes and No.

**DUMIT:** Perhaps you are hinting toward a more concrete definition. In your writing, you are fond of simple programs, like ELIZA, that appear to be incredibly intelligent and engage some people in intense conversations merely by transforming their answers into very simple questions.<sup>4</sup> Could programs like this be related to your conversation of zero degree?

**SACK:** Yes.

**DUMIT:** Ah good. ELIZA is fascinating because it is so simple. So it seems like you are looking for the simplest definition of conversation.

**SACK:** (curtly) No.

**DUMIT:** Well maybe not of the whole conversation, but perhaps the simplest form that one side of a conversation could take so as to keep the other side (presumably human) going, encouraging the other side to add to the conversation, or at least try.

**SACK:** (pensive) Maybe.

**DUMIT:** Obviously this definition is more involved than I first suspected. There are probably many dimensions to it. Even with ELIZA, there are constraints as to what you can ask. It is fairly easy to catch it out and have the conversation break down. Yet, the conversation can still, sometimes, be eerie.

**SACK:** (encouraging) Yes.

**DUMIT:** So it seems as if what is at stake in the zero of conversation is the limit of algorithms like ELIZA. The boundary area where conversation appears to happen. I know you are fond of the Turing Test, especially at how it functions as a technology for forcing the questioner to decide whether or not the respondent on the other side of a wall is a person or a computer.<sup>5</sup>

**SACK:** (emphatically) No!

**DUMIT:** Hmm. Ah yes, in your writings, you make it clear that the Turing test is really about telling apart a man pretending to answer questions as a woman from a computer programmed to answer as a man pretending to answer questions as a woman.<sup>6</sup> Quite a bit more subterfuge in the real version. Are you perhaps answering these questions as if you were enacting a software algorithm of Yes, No?

---

<sup>3</sup> See Jakobson, "Langue and Parole: Code and Message" in *On Language*, pages 92-93. Quoting Albert Sechahaye, and V.N. Volosinov. Roman Jakobson. "The Speech Event and Functions of Language" in *On Language*, page 74. Ferdinand de Saussure. *Course in General Linguistics*. Edited by Charles Bally and Albert Sechahaye in collaboration with Albert Riedlinger. Translated by Wade Baskin (New York: McGraw-Hill, 1966). V.N. Volosinov. *Marxism and the Philosophy of Language*. Translated by Ladislav Matejka and I.R. Titunik (Cambridge, MA: Harvard University Press, 1973). Charles Albert Sechahaye. *Programme et méthodes de la linguistique théorique; psychologie du langage* (Paris: H. Champion, 1908).

<sup>4</sup> Joseph Weizenbaum, "ELIZA – A Computer Program for the Study of Natural Language Communication between Man and Machine," *Communications of the Association for Computing Machinery* (9: 36-45), 1966.

<sup>5</sup> Alan Turing "Computing Machinery and Intelligence," *Mind*, Volume LIX, No. 236 (1950).

<sup>6</sup> Warren Sack. "Artificial Intelligence and Aesthetics," in *Encyclopedia of Aesthetics*. Edited by Michael Kelly. (New York: Oxford University Press, 1998, pp. 123-130).

**SACK:** (triumphantly) Yes!

**DUMIT:** And it certainly felt like a conversation. It was almost effortless to read into your brief remarks complete participation, sympathy, and evaluation. I suppose I could go on and with a series of short questions try and solve the algorithm you are using. Unless of course the sequence of Yes's and No's were generated in advance, or, God forbid, you were rolling dice. But wait a minute. If that were the case, then how could you have led me to this conclusion? You didn't know what I was going to ask, or in what order?

**SACK:** No.

**DUMIT:** Perhaps we should stop here.

**SACK:** Yes.

**DUMIT:** (decides to wait him out) So...

**SACK:** Yes?

**DUMIT:**

**SACK:** <sup>7</sup>

### **Philosophical Roots**

**DUMIT:** (exhausted) I'd like to push restart and try this interview again, in a more traditional manner. Can you describe how you think of conversations that involve a zero position?

**SACK:** I've been trying to enumerate the kinds of zero that I am interested in, specifically two kinds of zero that intersect: the zero person and zero knowledge. Firstly, is there a kind of zero position in the social world? In other words, is there a social position filled by someone colloquially referred to as a "zero," a "nobody," or a "nothing" as one might refer to someone in a phrase like "He means nothing (to me)"?

Most social theories have some notion of a topology of positions or roles; i.e., a set of positions in which people are related to one another through institutional or economic power, kinship, function, or ritual. Analogously, many theories of narrative and philosophy discuss the positions or relations between sets of *dramatis personae*. For instance, the narratologist Algirdas-Julien Greimas asserts that there are only six "actants," i.e., positions or types of character in narrative literature.<sup>8</sup>

In Gilles Deleuze's writings about Plato<sup>9</sup> -- specifically in his review of Plato's scathing attacks on the Sophists -- Deleuze recasts Plato's efforts to define the essence of the Sophist as an effort to sort out the Sophist's position in the social networks of society. Plato compares the Sophist to the hunter, the cook, the politician, and the fisherman and finds the Sophist to be a false pretender, an imitator, a false copy of all of them. In essence then -- if we can still use that turn of phrase -- the Sophist is a simulacrum according to Deleuze. It is thus that I imagine one can talk of the Sophists', simulacral position in the social order as a zero position. The parallel between the Sophistic genealogy of simulacra and contemporary, computational technologies of simulation interests me.

---

<sup>7</sup> See Keith Johnstone. *Improv: Improvisation and the Theatre* (New York: Routledge, 1981, pages 113-116).

<sup>8</sup> Algirdas Julien Greimas. *On Meaning: Selected Writings in Semiotic Theory*. (Minneapolis: University of Minnesota Press, 1987) *Du sens: Essais Sémiotiques*. (Paris: Editions du Seuil, 1970).

<sup>9</sup> Gilles Deleuze. *Difference and Repetition*. Translated by Paul Patton (London: Athlone Press, 1994). Gilles Deleuze. *The Logic of Sense*. Translated by Mark Lester with Charles Stivale; edited by Constantin V. Boundas (New York: Columbia University Press, 1990).

The second zero is a position of zero knowledge; i.e., the position of the so-called know-nothing. Again it is Deleuze's writings (but also those of Jacques Derrida<sup>10</sup>, Hanna Arendt<sup>11</sup>, and others<sup>12</sup>) on the Sophists, simulations, and Socrates that have unearthed the issue that I want to root through. Uncannily, while Plato praises Socrates, but condemns the Sophists as know-nothings, Socrates – in Plato's *Apology* – insists that he, Socrates, knows nothing.<sup>13</sup> So is this a homey, comfortable, *heimlich* relationship between the philosophical position of Socrates and that of the Sophists? Or, is this an *uheimlich*, uncanny, relationship between the philosophies that might serve to label them both as philosophies that come, not from the wise, but, from the know-nothings; i.e., from agents with zero knowledge and wisdom?

Because of the genealogies of Deleuze, Derrida, and Arendt that connect the simulacra to Sophists to Socrates, I refer to this double-zero position (a zero position in the social order and a position of zero knowledge) as a Sophistic position.

Speaking from either or both of these zero positions can be socially and politically dangerous or havoc wreaking. But I think there are some other more hopeful – therapeutic -- possibilities for such a speaker. A speaker, that might assist other people in just locating themselves (if you keep with this trope of a social matrix). If people are going to figure out where they are and what they know with respect to others in the social matrix, a zero position is something to bounce ideas off of.

Foucault's evaluation of Deleuze's discussion of Socrates, simulacra, and the Sophists<sup>14</sup> and Foucault's own writings about Socrates clarify the ethics of a zero position. Foucault's discussion of this ethics contextualizes it in a longer history, or genealogy, of what he calls "technologies of the self:" technologies "which permit individuals to effect by their own means, or with the help of others, a certain number of operations on their own bodies and souls, thoughts, conduct, and way of being, so as to transform themselves in order to attain a certain state of happiness, purity, wisdom, perfection, or immortality."<sup>15</sup> To play the zero position of the Sophists or Socrates is to encourage one's interlocutors to know themselves by getting them to reflect on their own thoughts and language.<sup>16</sup> Sherry Turkle indirectly places computer technologies in this

---

<sup>10</sup> Jacques Derrida. *Dissemination*. Translated by Barbara Johnson (Chicago: Chicago University Press, 1981).

<sup>11</sup> Hanna Arendt. "Philosophy and Politics", *Social Research*, 57, no. 1 (Spring 1990), pp. 73-103.

<sup>12</sup> See, for example, G.B. Kerferd, *The sophistic movement*. New York: Cambridge University Press, 1981.

<sup>13</sup> Plato. *The collected dialogues of Plato, including the letters*. Edited by Edith Hamilton and Huntington Cairns. Translators: Lane Cooper and others, New York: Pantheon Books, 1966.

<sup>14</sup> Michel Foucault. "Theatrum Philosophicum" in *Aesthetics, Method, and Epistemology: Essential Works of Foucault 1954-1984, Volume Two*. Edited by James D. Faubion. Translated by Robert Hurley and others (New York: The New Press, 1998).

<sup>15</sup> Michel Foucault. "Technologies of the Self" in *Ethics, Subjectivity and Truth: Essential Works of Foucault 1954-1984, Volume One*. Edited by Paul Rabinow. Translated by Robert Hurley and others (New York: The New Press, 1997, page 225).

<sup>16</sup> My understanding of this ethics is due to John Rajchman's discussion of the ethics of Foucault and Lacan, especially his discussion of ancient Greek and early Christian techniques or technologies developed to engender free speech, specifically, *parrhesia*. John Rajchman. *Truth and Eros: Foucault, Lacan, and the Question of Ethics* (New York: Routledge, 1991).

longer lineage of “technologies of the self” when she refers to the computer as “the second self.”<sup>17</sup>

Jacques Lacan points out how the psychoanalyst’s position is similar to the Socratic position: a position of the know-nothing who acts as a sort of “critical mirror”<sup>18</sup> reflecting the language of the analysand back for (re)consideration by the analysand.<sup>19</sup> This reflective function of the Sophistic position can be rather machinic and yet the analysand often imagines the reflected comments to be extremely knowledgeable. In the terminology of psychoanalysis, this reception of the analysand is called a process of *transference*. The analyst – to quote Lacan – becomes – from the perspective of the analysand – the “subject supposed to know.” But, the “knowledge” in such an interaction might be completely one-sided; i.e., it might rest entirely with the analysand. The analyst might, in fact, be performing as a machine, as a simulacrum.

So, my interest is in exploring this genealogy of the zero position, the Sophistic tradition, of the simulacrum by designing computer programs that know nothing, have no social position, and yet, simultaneously, might be technologies of the self insofar as they serve as critical mirrors for self reflection.

**DUMIT:** So your experiments use a computer program?

**SACK:** Contemporary technologies for simulation are based on the computer. I think the computer is part of the Sophistic philosophical tradition. Consciously or unconsciously the people who have been designing computer-based (dis)simulations for the past fifty years participate in some of that Sophistic philosophical discussion about what is a simulation or what is a dissimulation. The Sophistic tradition differs from the Platonic tradition because it does not – automatically – condemn dissimulation as lying and, thus, ethically wrong.

**DUMIT:** So computer experiments are almost the flip side of the accusations against the Sophists, that they are simply enacting a method, without conscience, without stakes? That they are just being mechanical?

**SACK:** Yes. In certain circumstances there is ethical good in that because what is built is a computational, critical mirror. These Sophistic, computer technologies of (dis)simulation just reflect people’s comments back at them. They are more interesting than a simple parroting though because they amplify certain words or comments.

I think this is what various experiments in theater, psychoanalysis, ethnomethodology, and computer science (specifically artificial intelligence) have been about: some of these experiments illustrate the ethical uses of a zero position.

In designing one of these technologies, I’ve been helped by looking at earlier technologies that challenge our notions of normal conversations and examining how they work or they do not work.

---

<sup>17</sup> Sherry Turkle. *The Second Self: Computers and the Human Spirit* (New York: Simon and Schuster, Inc., 1984)

<sup>18</sup> This turn of phrase, “critical mirror,” I owe to discussions with Krzysztof Wodiczko.

<sup>19</sup> “The essential, initial moment that is of particular relevance to the question we must ask ourselves about the action of the analyst, is that in which it is said that Socrates never claimed to know anything, except on the subject of *Eros*, that is to say, desire.” Jacques Lacan, “Of the Subject Who is Supposed to Know, Of the First Dyad, and of the Good” in *The Four Fundamental Concepts of Psycho-Analysis*. Edited by Jacques-Alain Miller. Translated by Alan Sheridan. (New York: W.W. Norton and Company, 1981, page 232).

## The Simulacra of Theater and Ethnomethodology

“This series of liberated simulacrum is activated, or mimes itself, on two privileged sites: that of psychoanalysis, which should eventually be understood as a metaphysical practice since it concerns itself with phantasms; and that of the theater, which is multiplied, polyscenic, simultaneous, broken into separate scenes that refer to each other, ...”<sup>20</sup>

**DUMIT:** Let’s start by talking about Keith Johnstone’s exercises for improv theater actors<sup>21</sup>. After that, perhaps we can talk about Harold Garfinkel’s *Studies in Ethnomethodology*<sup>22</sup>, and then some earlier experiments in artificial intelligence simulations of psychoanalysis?

**SACK:** Right. To talk about how they work it is worthwhile starting with Johnstone’s own description of one of his procedures. He writes:

“I say to an actress, ‘Make up a story.’ She looks desperate, and says, ‘I can’t think of one.’

‘Any story,’ I say. ‘Make up a silly one.’

‘I can’t,’ she despairs.

‘Suppose I think of one and you guess what it is.’

At once she relaxes, and it’s obvious how very tense she was.

‘I’ve thought of one,’ I say, ‘but I’ll only answer “Yes,” “No,” or “Maybe.”’

She likes this idea and agrees, having no idea that I’m planning to say ‘Yes’ to any question that ends in a vowel, ‘No’ to any question that ends in a consonant, and ‘Maybe’ to any question that ends with the letter ‘Y’.

For example, should she ask me ‘Is it about a horse?’ I’ll answer ‘Yes’ since ‘horse’ ends in an ‘E.’

‘Does the horse have a bad leg?’

‘No.’

‘Does it run away?’

‘Maybe’

She can now invent a story easily, but she doesn’t feel obliged to be ‘creative,’ or ‘sensitive’ or whatever, because she believes the story is *my* invention. She no longer feels wary, and open to hostile criticism, as of course we all are in this culture whenever we do anything spontaneously.”<sup>23</sup>

I performed this procedure in our introductory dialogue. Johnstone, after enacting this procedure to tease a story out of the actress, comments that he had to make a few revisions to the original description of the procedure. Specifically, he says “Yes” if he has given the actress too many “No”s in a row (too many being two). Also, after a while, the actress begins to get discouraged, so, at the end of the interaction, he always says “Yes.” As a computer program his procedure might look like this:

If the last two answers were “No,” then answer “Yes.” Else, if more than 30 total answers, then answer “Yes.” Else, if the question ends in vowel, then answer “No.”
----------------------------------------------------------------------------------------------------------------------------------------------------------------------------

<sup>20</sup> Foucault, “Theatrum Philosophicum.” *Op. Cit.*, page 348.

<sup>21</sup> Keith Johnstone, *Op. Cit.*

<sup>22</sup> Harold Garfinkel. *Studies in Ethnomethodology* (Englewood Cliffs, NJ: Prentice-Hall, Inc., 1967).

<sup>23</sup> Johnstone, *Op. Cit.*, pages 113-114.

Else, if question ends in “Y,” then answer “Maybe.”  
Else, answer “Yes.”

That is what the mechanism is, that is -- so to say -- 'How it works.' But that is not how we would describe it as someone interacting with it. It works that way for Johnstone, but it doesn't work that way for his student from whom he is trying to elicit a story. From her perspective, their interaction might be coded like this:

Wait until Johnstone has a story.  
Ask him yes or no questions.  
Try and put the answers together into a coherent story.  
If a coherent story is recognized, then tell Johnstone the story and learn from him whether the story was the one he was thinking about.

So the work of creating coherence between the responses is all on her side rather than on Johnstone's side. Here we have two descriptions of a mechanism. Seen from Johnstone's perspective the story elicitation procedure is so simple that it seems incredible to find that it can work as it does. Seen from the student actress' perspective the procedure is not simple at all: she has to construct a series of questions, interpret the coherence between the answers to her questions, and, finally, conclude the process by finding an overall structure and suitable end to the elicited story.

In theater, psychoanalysis, and various psychoanalytically-motivated artistic practices, there is an appreciation for the incredible amount of work that can be accomplished with simple elicitation procedures like the procedure described from Johnstone's perspective. Techniques like this one were the research of, for example, the Surrealists in their pursuit of the means to siphon stories out of the unconscious; techniques they referred to as *automatic writing*.<sup>24</sup>

Note that despite declared homonymous interests in the “automatic,” automatic writing and automatic programming in artificial intelligence is *not* about the invention of simple procedures to elicit creative work from people. Rather, most efforts in AI have been to write complicated procedures to try to get the computer to do some sort of creative work and to, thus, leave the person out of the loop entirely. In other words, AI has been interested in automating the procedures like the one described from the student actress' perspective, not the – perhaps much easier to implement – procedure as seen from Johnstone's vantage point.

Johnstone's procedure, and the Surrealist techniques of “automatic writing,” all work according to an aesthetics of “deautomatization.” Deautomatization and defamiliarization (or “making strange”) were described by the Russian Formalists (e.g., Roman Jakobson) as necessary processes that can work against the mentally- and spiritually-numbing forces of automatization, familiarization, and habituation. I want to return to this conflict (between avant-garde art and computerization) in understandings of the notion of the “automation” when we later discuss ELIZA, a computer program written in the 1960s by Joseph Weizenbaum to simulate a Rogerian psychotherapist.<sup>25</sup> ELIZA uses a procedure like, but a bit more complicated than, the one described from Johnstone's perspective.

---

<sup>24</sup> André Breton. “The Automatic Message” in *The Automatic Message, the Magnetic Fields, the Immaculate Conception: Anti-Classics of Surrealism*. Translated by David Gascoyne, Anthony Melville, and Jon Graham (London: Atlas Press, 1997).

<sup>25</sup> Weizenbaum, Op. Cit.

**DUMIT:** I see two levels here. Technologies such as Johnstone's method are interesting for what they reveal when they work. That is why you have collected them, because they do something uncanny to the respondent. And then, on top of that, your agenda is to look for where they break down in order to see where to go next.

**SACK:** Yes, before we talk about ELIZA, we might want to talk about the other things that are happening in the Johnstone-student interchange that Johnstone is not telling us. Specifically, we might want to speculate on where and why the described procedure breaks down. Johnstone does describe some of the breakdowns: for example, he says that he just said "Yes" all the time at the end when it looked like she was getting discouraged. But I think there are several ways to talk about how it works and how it breaks down.

**DUMIT:** It seems like the first thing that you pointed out about how it works is that it requires someone in the position to initiate a conversation. Johnstone has to be someone such that it matters to you to respond to him. Not a stranger. There has to be a whole context of knowing, or trust.

**SACK:** No, I don't agree that one cannot play a game like this with a stranger. In fact, I am interested in exactly that possibility. Think of all of the times someone totally unknown to you has sat next to you on an airplane or a train and told you their entire life story with practically no prompting from you. There simply has to be a choice, on the part of the person who is going to talk to Johnstone, that Johnstone, or someone else in a role like Johnstone's, is worth talking to. It might happen anywhere even without any pre-established trust between the participants.

But, such an interaction entails a non-trivial negotiation between the conversational participants. Firstly, the person or the machine has to broadcast its existence in some manner so that you know it is there. Secondly, the person has to make the choice that they are going to deal with this thing. And making that choice is more complicated than it looks. It is not simply a binary choice: Should I talk to this machine or person? It is a choice about how long to talk, how much effort to put into it, what kinds of assumptions and inferences to make in figuring out the responses. All of these choices are involved.

**DUMIT:** How exactly do we "choose" this?

**SACK:** Maybe "choice" is not the best word to use here because it implies a rational or conscious decision. Such "choices" might not be rational decisions. They might be a default set of assumptions, that at any time you talk to somebody, you put a lot of effort into figuring out their responses. You bring a set of assumptions to bear in order to make their responses coherent.

The sociologist of ethnomethodology, Harold Garfinkel, (borrowing from the sociologist Karl Mannheim<sup>26</sup>) calls this set of assumptions and inferences the "documentary method" of interpretation: "The method consists of treating an actual appearance as 'the document of,' as 'pointing to,' as 'standing on behalf of' a presupposed underlying pattern. The method is recognizable for the everyday necessities of recognizing what a person is 'talking about' given that he does not say exactly what he means, or in recognizing such common occurrences and objects as mailmen, friendly gestures, and promises [and stories, like Johnstone's story]"<sup>27</sup>

**DUMIT:** It is interestingly not an efficiency analysis, or an application of Occam's razor. If someone is responding to you, you do not look for the 'simplest' explanation –

---

<sup>26</sup> Karl Mannheim. "On the Interpretation of Weltanschauung" in *Essays on the Sociology of Knowledge* (New York: Oxford University Press, 1952, pages 53-63).

<sup>27</sup> Harald Garfinkel, "Common sense knowledge of social structures: the documentary method interpretation in lay and professional fact find", in *Op. Cit.*, page 78.

for instance, that they might simply be reading off of a prepared list of “Yes’s” and “No’s” in order to respond to you. That might be the simplest explanation of someone’s responses, but we typically go far in the other direction and come up with very involved and sometimes convoluted explanations for the other’s responses.

**SACK:** Yes, Garfinkel’s own studies of the documentary method illustrate how unparsimonious -- even generous -- people can be in conversing with a virtual machine such as a person following a very simple procedure.

For one of his experiments, Garfinkel uses a procedure even simpler than Johnstone’s. Before the start of the dialogue, Garfinkel asks his experimental subjects to discuss the background of some serious problem on which the subject would like advice and then to address to a “counselor” a series of questions that could be answered with either a “Yes” or a “No.” Garfinkel then instructs his confederate, the “counselor,” to “answer” the sequence of questions by simply reading from a list of “Yes’s” and “No’s” pre-decided with a table of random numbers. In other words, Garfinkel’s procedure does not even use the last letter of the interlocutor’s question (as Johnstone’s procedure does), but, rather, only computes a sequence of “Yes’s” and “No’s” with a random process -- like a series of coin flips -- and then uses that sequence in the “dialogue.” After these “dialogues,” Garfinkel claims “The typical subject assumed, over the course of the exchange, and during the postexperimental interview, that the answers were advice to the problem, and that this advice as a solution to the problem was to be found via the answers [from the pre-programmed ‘counselor’]”<sup>28</sup>

Garfinkel expands upon his experimental results to point out that social scientists also employ a documentary method. Even as scientists, sociologists are not necessarily going to construct the most parsimonious explanation for the answers they get. Such parsimony – e.g., the belief that a respondent’s answers were simply based upon a precomputed list of “Yes” and “No” – would be scientifically and socially ludicrous.

**DUMIT:** Other than the fact that Johnstone and Garfinkel say “Yes” and “No,” and people don’t settle for the most parsimonious explanation, why are you collecting these procedures together? My sense of them is that of the zero-degree conversant. With what conditions are we able to create the context where a simply-described mechanism functions as a respondent in a conversation?

**SACK:** I agree that the conditions of context are crucial differences between the various examples of this Sophistic, Socratic, simulacral method. But, the contexts are not entirely the determinant of how or why the simulated conversants work. Johnstone and Garfinkel have differing aesthetic and ethical imperatives that also differentiate their work.

**DUMIT:** One thing that that comes to mind is the different domains of these technologies. Johnstone’s is one of creativity. He says, “Let’s guess, or make something up.” His methods bring out what is within you. Ostensibly his invitation is to co-create something with his respondent, even though, his methods do not require him to participate in the creativity beyond being there, at the place and time of the creativity. He puts you in a situation where creativity can happen. Garfinkel’s domain is not at all the same. It is about ruminating or about troubling his subject. He says at one point that what he wants to do is put people in situations where all of their techniques for making sense of a situation fail.

**SACK:** Well they are kind of complements of each other. Johnstone’s says that people are bad at improv theater because they “block”; i.e., they inadvertently cut off a dialogue in which something creative might happen if the conversation could be nurtured

---

<sup>28</sup> *Ibid.*, page 89.

and allowed to grow into a longer, more intricate, interaction. His notion of “blocking” is really about not giving something in return for a conversational move offered by one’s partner in dialogue. A classic notion of blocking is when two people are creating a dialogue and one person says “No.” Johnstone has lots of techniques for not saying “No” and, through their usage, allowing the conversation of interaction to continue rather than to become blocked. In an improv scene, one actor might begin the scene by saying to the other “I see you have spilled chocolate on your shoes.” If the other says “No” the dialogue dies right there. He tries to get people, i.e., improv actors-in-training, to not block. In general, he sees the untrained actor as blocked up.

Garfinkel's assumptions about the untrained person are of exactly the opposite sort. He posits a person who, given a set of blocking strategies, is able to get around them and improvise some sort of means to deal with the situation nonetheless.

**DUMIT:** Johnstone's is about motivating a conversation. Garfinkel is about killing a conversation, ultimately trying to confuse or bewilder.

### Simulating a Psychotherapist

**SACK:** Let’s discuss a third procedure related to Johnstone’s and to Garfinkel’s; namely, an early experiment in artificial intelligence to simulate the interviewing techniques of a Rogerian<sup>29</sup> psychotherapist. The simulation, a computer program called ELIZA, was designed and implemented by Joseph Weizenbaum in 1966 at MIT.<sup>30</sup> The anthropologist Lucy Suchman has shown how the workings of ELIZA are similar to Garfinkel’s experiments with the documentary method.<sup>31</sup> I think Suchman’s use of Garfinkel’s vocabulary to explain how ELIZA works can give one a new appreciation for the ELIZA program as procedure of elicitation rather than as a failed “model of the mind” as most artificial intelligence histories of ELIZA would have us look at it. Here is a snippet of a dialogue with ELIZA. The transcript is taken from Weizenbaum’s original 1966 paper. The text in capital letters was generated by the machine, ELIZA.

“Men are all alike.

IN WHAT WAY

They’re always bugging us about something or other.

CAN YOU THINK OF A SPECIFIC EXAMPLE

Well, my boyfriend made me come here.

YOUR BOYFRIEND MADE YOU COME HERE

He says I’m depressed much of the time.

I AM SORRY TO HEAR YOU ARE DEPRESSED”<sup>32</sup>

**DUMIT:** Was ELIZA originally designed to be a model or spoof of Rogerian therapy?

**SACK:** As far as I understand it was supposed to be a model of a therapist. There were two sorts of aspirations for it: (1) it was hoped that ELIZA could be developed as model of a therapist; and, (2) it was hoped that ELIZA could be the foundation of a more general “man-machine” communication technology. Weizenbaum's was interested in, as he put it, “man-machine communication” and using written English for that. For him, patient-therapist interaction happened to be a limited, and thus tractable, situation of dialog within which to do communication experiments. As far as I understand, previous

---

<sup>29</sup> Carl Rogers. *On Becoming a Person*. (Boston: Houghton Mifflin, 1961).

<sup>30</sup> Weizenbaum, *Op. Cit.*

<sup>31</sup> Lucy Suchman. *Plans and Situated Actions: The Problem of Human-Machine Communication* (New York: Cambridge University Press, 1987, page 64).

<sup>32</sup> Weizenbaum, *Op. Cit.*, page 36.

to developing ELIZA, Weizenbaum had a series of conversations with the psychotherapist Kenneth Colby (who later developed his own computational models of analysts and analysands).<sup>33</sup> Colby, I believe, was more specifically interested in a computational model of a therapist. But, in Weizenbaum's original paper, Weizenbaum says that the therapist script is but one of the many scripts that ELIZA might run. There were going to be many others, that would have the same machinery and so these scripts were suppose to be easy to swap in and out. I don't think it was a coincidence that none of the other scripts worked as well as the therapist script.

I imagine that Weizenbaum's notion of script comes from the social science of the time; i.e., the appropriation by social scientists of the language of theater to describe social life. Erving Goffman, for example, is usually credited with this dramaturgical perspective for sociology.<sup>34</sup> These scripts of social science were a way of asserting that people had recurrent routines and set habits of speech and behavior. Perhaps not for Goffman, but for a variety of other social scientists, the theoretical construct of the script was a means for talking about the institutional or ideological reproduction of fixed, stable structures of social life as compared to the improvisational, performance-based instabilities of interpersonal interaction. Weizenbaum presumably assumed the psychotherapist script was analogous to these other scripts articulated by the social scientists of the time. If such an analogy was warranted then, it seemed reasonable to believe that other scripts could be coded in machine-readable format and then one could swap them in and out so causing ELIZA to simulate a variety of conversational roles.

**DUMIT:** I was asking because I'm trying to figure out how it was similar to Abelson's project to simulate Barry Goldwater?

**SACK:** Abelson's work on his Goldwater Machine<sup>35</sup> was an attempt to represent the scripts of ideology that might be said to have motivated the knee-jerk, Cold Warrior rhetorics of public figures like Barry Goldwater. Weizenbaum, in his original paper on ELIZA, cites some of Abelson's early work.

**DUMIT:** Both ELIZA and the Goldwater Machine seem to have been interventions or social experiments. Abelson did not set out to design an intelligent machine. Rather, he set out to show that right-wing ideology had no intelligence behind it by showing that the Cold Warrior's discourse could be simulated by a machine. One might read ELIZA as an analogous critique of Rogerian Therapy insofar as such therapy could be proceduralized as a very simple trick of feeding people's words back to them. It seems ELIZA was originally received as an interesting technology. Then, later, after its original reception, Weizenbaum said it was just a trick. In contrast to Abelson and Weizenbaum, Colby's work of that time seems like an effort without irony to simulate neurotics and psychotics.

**SACK:** There exists a conflicted discussion concerning ELIZA's presentation and reception. In the original 1966 paper by Weizenbaum, he explains the ELIZA computer

---

<sup>33</sup> Kenneth Colby. "Computer Simulation of a Neurotic Process" in *Computer Simulation of Personality: Frontiers of Psychological Research*. Edited by S.S. Tomkins and S. Messick (New York: Wiley, 1963, pages 165-180). Kenneth Colby, S. Weber, F.D. Hilf. "Artificial Paranoia" in the *Journal of Artificial Intelligence* (2: 1-26), 1971. Kenneth Colby, James B. Watt, John P. Gilbert. "A Computer Method for Psychotherapy: Preliminary Communication" *Journal of Nervous and Mental Diseases* (2: 148), 1966.

<sup>34</sup> Erving Goffman. *The Presentation of Self in Everyday Life* (Garden City, NY: Doubleday, Anchor Books, 1959).

<sup>35</sup> Robert P. Abelson. "The Structure of Belief Systems" in *Computer Models of Thought and Language*. Edited by R.C. Schank and K.M. Colby. (San Francisco: W.H. Freeman and Company, 1973).

program as an experiment in “man-machine communication.” He suggests at the end of the paper how it might be improved for better communication. I think he means in the conclusion of that paper that he wants to improve the machine’s “understanding” where “understanding” is considered to be a process of translating and then storing English sentences into formalized, predicate-argument structures (e.g., logical statements) that can later be retrieved and compared by the machine as a synthetic “memory” of the conversation. This kind of rationalistic notion of “understanding” was prevalent in artificial intelligence research until the mid-1980s and is hegemonic even today.

By 1966, it is clear that ELIZA had an incredible allure for specialists and non-specialists alike; they wanted to talk to the machine. People want to talk to the machine the way they want to talk to their psychologist or psychoanalyst. So they found interaction with it fulfilling. But Weizenbaum was morally against the whole idea. Needless to say, Weizenbaum’s moral indignation is a very different approach to dialogue with a machine than the ethics of the zero position -- we have just discussed -- and as such ethics have been explicated by Deleuze, Lacan, Foucault (in his discussion of free speech), and others. After 1966, Weizenbaum stopped talking about ELIZA as an experiment in “man-machine communication” and began explaining ELIZA’s mechanism as a mere trick.<sup>36</sup> In fact, even in Weizenbaum’s 1966 paper, he introduces ELIZA’s mechanism as a thing to be explained away (even as he discusses it, not as a “trick”, but as an experiment in “man-machine communication.”)

At the time, in 1966, other people in the computer science community seemed to be willing to see ELIZA as a first-step towards improved “man-machine communication.” Again, within this community, “improvement” would have been understood as improvements in machine “understanding” as a rationalistic exercise to translate English into a set of more formal, machine-interpretable, statements. But, I think Weizenbaum’s moral outrage – against, not only ELIZA, but the larger goals of artificial intelligence to build “intelligent” machines – was grating for the community of AI researchers. Consequently, their interpretation of ELIZA became less generous. Instead of a first-step towards a shared research goal, ELIZA was seen as a simple “trick.” Simultaneously, they begin to dismiss or ignore Weizenbaum’s discussion of the morals of AI by belittling his technical work. So they essentially said, “Yes, in fact, Weizenbaum is correct, it was just a trick. It was a hack. And anyone with half an hour to spare can put one of those together. Who cares, it was just a piece of trivial work. So let’s just not talk about Weizenbaum, nor about his little hack.”

But the story doesn’t end here because ELIZA was, and still is, institutionalized within artificial intelligence (AI). If you teach an introduction to AI programming, the first assignment in natural language processing is to write an ELIZA.<sup>37</sup> So practically anybody who has taken an introductory course in AI programming has written one of these. What does that say about the whole field? It is an interesting contradiction. On the one hand, it says, here is the simplest of our systems. Here is how it works, here is how you put it together, and here is how it represents an experiment in human-computer communication. But then as you grow in the field and develop as a researcher, you are supposed to leave it all behind, because technically it is considered to be a trivial thing. Thinking about what ELIZA might mean socially, politically, ethically is not necessarily

---

<sup>36</sup> Joseph Weizenbaum, *Computer Power and Human Reason*. (San Francisco: W.H. Freeman and Company, 1976).

<sup>37</sup> For a good example of this simultaneous institutionalization and dismissal of ELIZA, see Peter Norvig. *Paradigms of Artificial Intelligence Programming* (San Mateo, CA: Morgan Kaufmann Publishers, 1992, p. 151-174).

taken up again within AI. Rather such work is considered to be someone else's work, specifically, the work of philosophers and sociologists.

With the advent of MUDs, MOOs, IRC, and a variety of other new "social spaces" on the Internet, there are a myriad of recapitulations of ELIZA called 'chatter bots.' These 'bots do essentially what ELIZA was doing but with bigger databases of transformation rules.<sup>38</sup> For example, one of ELIZA's transformation rules that looks something like this:

If the input contains a phrase like this

"I remember X"

then ask the following question:

"What in the present situation reminds you of X?"

ELIZA's database contains dozens of these transformation rules, newer chatter bots – like JULIA written by Michael Mauldin at Carnegie Mellon University<sup>39</sup> – contain hundreds of such transformation rules. The advent of these newer chatter bots has re-engendered the two-fold discussion about ELIZA and ELIZA-like systems: (1) Can technologies like ELIZA "understand" language? and, (2) What are the ethics of designing and interacting with artificial intelligence systems?

### **Artificial Participation and Machine Understanding**

**DUMIT:** I'm still struck by the assumption of those debating that what AI or a computer should do is "to understand" rather than to participate.

**SACK:** One way to envision the history of artificial intelligence is to see it as an outgrowth of cybernetics.<sup>40</sup> During and after World War II cyberneticians were interested in human-machine assemblages. The machines they built were designed to process feedback from the surrounding environment, interact with the environment – or participate with a human to which the machine was attached – and then process another cycle of feedback, etc.

Diverging from this tradition, early-AI researchers began to focus more on internal, mental events and less on feedback from the environment or interacting humans. Early-AI, i.e., what has been called "symbolic AI," might be characterized as a reinvention of the Cartesian isolated thinker; or, more generally, as an effort to implement a variety of rationalistic philosophical ideas in computer hardware and software. The rationalists' projects of separating mind from body and internal mental events from external events were reanimated in the work of symbolic AI. Consequently, it was considered essential for AI researchers building language technologies to address the problem of how "external" written and spoken language could be processed by the machine and thereby converted into some sort of "internal" representation. This rationalistic description of "natural," "external" language as imperfect forms of formal, "internal" representations has largely dictated the kinds of work done in the areas of AI natural language understanding and computational linguistics. The whole field of "natural language understanding" is inexplicable if it cannot be placed in this older, philosophical tradition of rationalism.<sup>41</sup>

**DUMIT:** So the older AI approach to interaction with the environment was to create an internal representation of the environment?

---

<sup>38</sup> Andrew Leonard. *Bots: The Origin of New Species*. (San Francisco: Hardwired, 1997).

<sup>39</sup> Michael L. Mauldin "Chatterbots, Tynmuds, and the Turing Test: Entering the Loebner Prize Competition" in *Proceedings of AAAI'94 Conference* (Seattle: AAAI Press, 1994).

<sup>40</sup> Warren Sack, Op. Cit. See also Howard Gardner. *The Mind's New Science: A History of the Cognitive Revolution* (New York: Basic Books, 1985, pages 144-145).

<sup>41</sup> C.f., Terry Winograd and Fernando Flores. *Understanding Computers and Cognition: A New Foundation for Design* (Reading, MA: Addison Wesley, 1987).

**SACK:** This rationalistic approach produced two overwhelming problems for AI research: (1) common sense; and, (2) the so-called “frame problem.” AI researchers repeatedly cite “common sense” as a problem because if everything we, as people, know about the world has to be encoded in a formal language for a machine, then there appear to be millions and millions of banal, common sense rules (like, “What goes up must come down,” “People need food and water to survive,” “Steel is a stronger material than balsa wood,” etc., etc.) that need to be written down and then translated into a formal, “internal,” machine-readable format. The AI scientist, Doug Lenat, has been trying to encode all of common sense knowledge in a database for the last fifteen years.<sup>42</sup> But, while Lenat’s approach is the most literal interpretation of this apparent challenge for AI, many other AI projects in expert systems, machine learning, and natural language understanding have been devoted to translating “informal,” human knowledge and language into machine-readable databases.

The “frame problem” is also a result of AI’s Cartesian division of “internal” mental events from external events. The frame problem is this: If something in the “outside” world changes (i.e., “outside” of the machine itself), what assertions and inferences of the machine’s “internal” model of the world need to be changed or updated?<sup>43</sup>

These two problems (common sense, and the frame problem) kept AI researchers busy for years and have only recently been put into question by “newer” approaches to building machines. These newer approaches are, in many ways, re-inventions of cybernetics. By refusing to draw a strict division between the “internal” representations of the machine and the “external” events of the “outside” world, machines are being built that can more fluidly interact or participate with people, the surrounding environment, and other machines.

There are a great number of reasons why AI research – which previously was concerned almost entirely with “symbol systems”<sup>44</sup> is increasingly interested in what might be called “indexical systems.” However, it is clear that when the anthropologist Lucy Suchman introduced the term “index” to AI researchers in 1987<sup>45</sup>, a variety of new insights were opened in AI research, especially research in the area of robotics and physically-situated systems.

The distinction between a “symbol” and an “index” is a distinction made by the semiotician Charles Sanders Peirce. According to Peirce, an index is a sign that is directly -- for instance, causally -- linked to its object such as smoke is linked to fire, or as a thermometer is linked to the surrounding temperature. A symbol, by comparison, is a sign linked to its object more indirectly, for instance, through an association of ideas or by habitual connection. The meaning of an index is thus more dependent upon its immediate context than the meaning of a symbol. In other words, AI has begun to build systems with “context sensitive” parts.

This new interest in indexical systems could be explained in semiotic terms or as a return to cybernetics. But, Suchman’s introduction of the term is done to introduce ethnomethodology, and specifically the work of Garfinkel, to designers of “intelligent” systems. For Garfinkel, indexical expressions were of central concern because he

---

<sup>42</sup> Douglas Lenat and R. Guha, *Building large knowledge-based systems* (Reading, MA: Addison-Wesley, 1990).

<sup>43</sup> J. Martins “Belief Revision” In *Encyclopedia of Artificial Intelligence, Second Edition*. Editor-in-chief: Stuart C. Shapiro (New York, 1992, pp. 110-116).

<sup>44</sup> Herbert Simon, *The Sciences of the Artificial. Second edition* (Cambridge, MA: MIT Press, 1981).

<sup>45</sup> Suchman, *Op. Cit.*

considered the meaning of language to be dependent upon the context and the production of common sense to be a result of participants' application of the documentary method to the locutions of the other participants in a discussion. In short, Suchman shows how Garfinkel's methodology and assumptions are the inverse of AI's: where AI researchers took language to be context-independent, fixed, symbols Garfinkel thought of it as context-sensitive and dependent upon indexical signs; where researchers in symbolic AI thought of "common sense" as a vast database of stored, consensual knowledge, Garfinkel thought of common sense as a product of social interaction.

Suchman goes beyond Garfinkel by pointing out how machines, such as ELIZA, can be a conversational participant from a Garfinkel-inspired perspective.<sup>46</sup> Suchman's articulation of these differences between AI and ethnomethodological approaches to language, participation, and common sense turned out to be extremely useful for the technical work of AI researchers who were trying to build interactive machines. It gave such researchers, especially researchers working on robots and other physically-situated systems, a different approach to thinking about – and ultimately avoiding – the so-called "frame problem" and the assumed problems of common sense.<sup>47</sup> Strangely, Garfinkel-inspired, ethnomethodological approaches to AI language technologies have not really been taken up even though Garfinkel was very concerned with language.

### **Repetition and Representation**

**DUMIT:** Right. It strikes me that mundane things like repetition and assent are incredibly important in conversations, in keeping them going. I've clipped a couple of passages by the linguist Deborah Tannen and would like to get your feedback:

"... repetition serves an over-arching purpose of creating interpersonal involvement. Repeating the words, phrases, or sentences of other speakers (a) accomplishes a conversation, (b) shows one's response to another's utterance, (c) shows acceptance of others' utterances, their participation, and them, and (d) gives evidence of one's own participation. It provides a resource to keep talk going, where talk itself is a show of involvement, of willingness to interact, to serve positive face. All of this sends a metamessage of involvement. This may be the highest-level function of repetition — in the sense in which Gregory Bateson<sup>48</sup> adapts Bertrand Russell's notion of logical types to describe the metamessage level of interaction: the level at which messages about relationships are communicated."<sup>49</sup>

**SACK:** You are pointing to the work of Harvey Sacks, Gail Jefferson, Emanuel Schegloff, Harold Garfinkel, and others initiated a field now known as conversation analysis. Conversation analysts attempt to find recurring patterns in large corpora of naturally occurring conversations. and have developed a variety of ways to talk about the significance of recurrences and repetition in language.

**DUMIT:** In my reading of Tannen's reading of Bateson, the kinds of repetition that ELIZA enacts concertedly produce this meta-message of involvement.

---

<sup>46</sup> Suchman, Op. Cit., page 63.

<sup>47</sup> See especially, Philip Agre. *Computation and Human Experience* (New York: Cambridge University Press, 1997).

<sup>48</sup> Gregory Bateson. *Steps to an Ecology of Mind* (New York: Ballantine, 1972).

<sup>49</sup> Deborah Tannen, *Talking Voices: Repetition, dialogue, and imagery in conversational discourse* (New York, NY: Cambridge University Press, 1989: 52).

“In a closely related way, repetition also serves the purpose served by all conventionalized discourse strategies at every level of language: giving talk a character of familiarity, making the discourse sound right. This is a verbal analogue to the pleasure associated with familiar physical surroundings: the comfort of home, of a favorite chair. It is the trust in a speaker one knows, or one who seems — by virtue of appearance, dress, kinetics, and ways of speaking — like one to be trusted. The pattern of repeated and varied sounds, words, phrases, sentences, and longer discourse sequences gives the impression, indeed the reality, of a shared universe of discourse.”<sup>50</sup>

Calling attention prior, during, or after to the ease with which a dialogue can happen at the zero degree (algorithmically, without stakes, without status, without position of one side) produces a meta-metamessage that returns the sender’s message back to him or her self inverted.<sup>51</sup> This message takes the very familiarity, trust and pleasure of right-soundingness of the conversation and reveals it to be a doppelganger. “Just the same, but different,” to invoke both Sigmund Freud<sup>52</sup> and Homi Bhabha<sup>53</sup>.

**SACK:** Your references to conversational analysis, specifically the work of Tannen, psychoanalysis, specifically the aesthetics of the “uncanny” (i.e., the doppelganger), and philosophical analyses of the “other” tie together several – as yet unconnected – strands of our discussion. After our opening, simulated dialogue, I explained the double-zero I am interested in the terminologies of French, post-structuralist theory. Since then our conversation has wandered into American semiotics and social science and the largely-American vocabulary of artificial intelligence. What ties these two ends of the discussion together is a fascination with repetition that both the French post-structuralists and the American conversation analysts have.

In his book *Repetition and Difference*<sup>54</sup>, Gilles Deleuze explains why repetition is an alternative to representation in explanations of the meaning of signs. I will risk a crude interpolation of Deleuze’s argument into our conversation here. An explanation of the concept of representation usually presupposes *a priori* differences between “here” and “there,” “presence” and “absence,” “internal” and “external.” In other words, it rests upon an understanding of signs as symbols. If this word here -- “word” -- is to represent that thing there -- “thing” -- then “word” and “thing” must be separable from one another and, furthermore, the “word” and the “thing” must have essentially context-independent meaning. After all, if a word changes radically with each change in its context, then how good a job can it do standing in for the “thing” not present to represent itself? So, the concept of representation makes the most sense for those who analyze the meanings of the symbolic functions of language. But, how is meaning to be explained in the domain of indexical signs where language changes meaning according to context?

Repetition in conversation gives us a means of describing the meaning of indexical signs. For example, if I say “Hello” and then you say “Hello” both “Hello”s mean practically the same thing: they are greetings. But, if after you say “Hello” I say “Hello”

---

<sup>50</sup> Ibid., p. 52.

<sup>51</sup> Jacques Lacan. “The Purloined Letter” in *The Purloined Poe: Lacan, Derrida and Psychoanalytic Reading*. Edited by John P. Muller and William J. Richardson (Baltimore: Johns Hopkins University Press, 1988).

<sup>52</sup> Sigmund Freud. “The ‘Uncanny’” in *Standard Edition of the Complete Psychological Works of Sigmund Freud*. Edited by James Strachey. Volume 17, pages 217-252 (London, 1955).

<sup>53</sup> Homi Bhabha

<sup>54</sup> Deleuze, *Op. Cit.*

again, my second “Hello” (the third in the series), is probably not a greeting but is, perhaps, a sign of a noisy environment or a symptom of the fact that I cannot hear very well after last night’s 150 decibel punk concert. The meanings of indexical signs -- like these three “Hello”s -- can be charted out as differentials, i.e., slight, context-dependent differences from the previous repetition of the same sign. In short, “Hello” does not represent or stand in for something else; rather, its meaning is its differences from previous repetitions of “Hello.” Each new mention of “Hello” links it to previous moments of the conversation and to the conversation participants who have said “Hello.”

For Deleuze, these series of sign repetitions that do not point back to some “original” or “real” thing -- but rather point to one another and the contexts of the present, past, and future -- are *simulacra*. Given this description of simulacra as elements in a series of repetitions we really should go back to the beginning of our interview and revise the double-zero that was defined using the trope of an *a priori* matrix of social positions and an idea of “knowledge” that, at least implicitly, seems to reference the concept of representation (rather than repetition). But, maybe that can be left as an exercise for the reader?

If I am not mistaken, Deleuze subscribes to a sort of sociolinguistics not unlike the productions of Tannen.<sup>55</sup> Perhaps it is too obvious to mention, but these sociological analyses of the meaning of language (and the production of language) are in strong contrast to the psychological analyses preferred by artificial intelligence researchers and cognitive scientists in general. Even today, AI work in discourse analysis, computational linguistics, and natural language processing in general rely heavily on representations of conversants’ goals and intentions.<sup>56</sup> This tradition of discourse analysis is dependent upon the field of pragmatics especially speech act theory as it has been defined by John Searle and others. Speech acts (i.e., acts with words) are defined as illocutionary acts, i.e., as acts of intention.<sup>57</sup> It is strange that AI discourse analysis has followed this path since, from Searle’s perspective, machines cannot have intentions and, consequently, they cannot really be said to be legitimate participants in a conversation. He has a long-winded, much debated article on this issue.<sup>58</sup> I think it is amusing that he and his critics and allies have spent so much time on his argument concerning the impossibility of “machine intelligence.” It seems obvious to me that he couldn’t say anything other than what he does given his early work on the definition of speech acts as acts of “inner” intentions. Likewise, I find it funny that so many people have spent their entire careers implementing Searle-like pragmatics in an effort to give machine’s symbolic “goals” and “intentions” since such a project seems oxymoronic in the lexicon of pragmatics<sup>59</sup>.

---

<sup>55</sup> Gilles Deleuze and Félix Guattari. “November 20, 1923 -- Postulates on Linguistics” in *Thousand Plateaus: Capitalism and Schizophrenia*. Translated by Brian Massumi. (Minneapolis: University of Minnesota Press, 1987).

<sup>56</sup> See, for example, Megan Moser and Johanna D. Moore. “Towards a Synthesis of Two Accounts of Discourse Structure” in *Computational Linguistics* (Volume 22, Number 3, pages 409- 419), September 1996.

<sup>57</sup> John Searle. *Speech Acts: An Essay in the Philosophy of Language* (New York: Cambridge University Press, 1969).

<sup>58</sup> John Searle. *Minds, Brains and Science* (Cambridge, MA: Harvard University Press, 1984).

<sup>59</sup> Note that pragmatics is a subfield of linguistics that has been investigated from an Anglo-American philosophical perspective and also from a more “Continental” philosophical perspective. Thus, when Deleuze says “pragmatics” he often includes what many Anglo-American linguists would call sociolinguistics. Compare, for instance, Gilles Deleuze. *Negotiations: 1972-1990*. Translated by Martin Joughin (New York: Columbia University Press, 1995, pages 28-29) with a

If, as Tannen claims, to repeat is to participate in a conversation, then she provides us with a radically different notion of what an artificial participant should be. As Suchman points out, such a participant can be as simple as ELIZA. It is my opinion that a sociolinguistic description of conversation and repetition implies a completely different approach for constructing computer-based language technologies than the teleology-based approach taken by AI researchers. If what we are interested in is artificial participation rather than machine understanding<sup>60</sup>, then it is crucial to design machines that are good at repetition, but it is really quite beside the point to design machines that *have* intentions, as AI researchers have been trying to do for years.

In my work I am trying to follow this implied sociolinguistic approach to the construction artificial participants. Perhaps I will be as successful at constructing such “indexical systems” as the roboticists and the AI people working on physically-situated agents have been in following the sociolinguistic, Garfinkel-inspired approach described by Suchman. But, even if the sociolinguistic approach is useful for describing how conversation and conversational participants work -- or could work -- it leaves the ethics of interaction and participation unaddressed, or at least unresolved. That’s why I keep wading back into the depths of philosophy, because I think there some of these ethics and aesthetics -- of the uncanny, of the other, of the simulacrum -- are addressed.

### **Ethics and Aesthetics**

**DUMIT:** Suchman seemed to be talking about computational artifacts as being preferably accountable, reasonable, and responsive to the immediate context and with the people interacting with them. Garfinkel seemed to be interested in almost the same thing, but defined negatively.. He seemed to be fascinated with where common sense breaks down in rational conversation and also, where common sense is maintained or re-produced in irrational conversational interactions .

**SACK:** I think this difference between Suchman and Garfinkel can be described using the language of ethics and aesthetics. Also, they can be differentiated according to their respective institutional settings. While Garfinkel was an academic, a professor at UCLA, Suchman was working for the research branch of the Xerox Corporation, specifically Xerox PARC, the Xerox Palo Alto Research Center. Her discussion, in the book *Plans and Situated Actions*, of conversational analysis, the semiotics of the index, and computer-based machinery is illustrated by an analysis she does of the ways in which a the computer-based help system of a photocopying machine breaks down. In other words, one way to read Suchman is the way that I think the computer business of Silicon Valley does read her: she shows how anthropology can be useful for streamlining business communication and improving processes of automation. I don’t necessarily agree with this reading of Suchman, but I think, for many people in the software and

---

standard Anglo-American textbook on the subject; e.g., Stephen C. Levinson. *Pragmatics* (New York: Cambridge University Press, 1983). In this particular case, I am referring to the Anglo-American tradition of pragmatics.

<sup>60</sup> A third path that finds both artificial participation and machine understanding problematic -- even impossible -- has been articulated by Winograd and Flores, *Op. Cit.* For them, computers are not to be simulations but should be, rather, “tools for conversation”; i.e., substrates in which people converse with one another. The systems they propose to build (and often what the researchers in the field of Computer-Supported Cooperative Work, in general, often propose) are rather like complicated electronic mail programs for composing, editing, sending, and receiving messages. Winograd’s and Flores’ position is a straightforward refusal of computers as synthetic conversants and seems to me to be a less tortured incorporation of Searle’s speech act theory than the the one of AI researchers attempting to program symbolic intentions into computers.

hardware industries, Suchman's work was showed them why it would make good business sense to hire anthropologists for a research and development department. In the language of business, Suchman's work illustrates how to do a better job of automation (specifically computerization). In the language of aesthetics and ethics (specifically those of Russian Formalism), Suchman's work shows how *automatization* can be accelerated.<sup>61</sup>

But, what is seen as a virtue from the perspective of capitalism and business, could be seen as a vice from the perspective of avant-garde art, like the Russian Formalist's perspective. Thus, the Russian Formalist Viktor Shklovsky states that the purpose of art is to work against habitualization, familiarization, and automatization: "Habitualization devours works, clothes, furniture, one's wife, and the fear of war... art exists that one may recover the sensation of life; it exists to make one feel things, ... The technique of art is to make objects 'unfamiliar', to make forms difficult, to increase the difficulty and length of perception because the process of perception is an aesthetic end in itself and must be prolonged."<sup>62</sup>

Garfinkel's experiments seem to have the aesthetics of avant-garde art: he stages a process that might be called defamiliarization or deautomatization. His experiments shock his subjects into reconsidering what they take for granted, what they assume to be commonsensical. And, he produces these effects in a manner that is reminiscent of the Russian Formalists' techniques of defamiliarization, 'making strange,' i.e., *ostraneniye*: "the basic way of seeing any object anew is 'to place the object in a new semantic row, in a row of concepts which belong to another category.'"<sup>63</sup> Garfinkel effects a subject's defamiliarization by telling the subject -- after the subject has been "counseled" and has understood the "counseling" to be a sequence of reasonable answers to the questioned posed -- that the "answers" the subject was given were not "answers," but rather only a sequence of randomly selected "Yes"s and "No"s. This repetition and recontextualization of the subject's "conversation" with the machinic "counselor" is a procedure of deautomatization that poses a challenge to the subject's common sense.

The Russian Formalist vocabulary of (de)automatization is one way of describing two opposing ways to "use" the insights of sociolinguistics and the sociolinguistic/ethnomethodological analysis of software. But, as we have already mentioned in our discussion, one can also describe the aesthetics and ethics of Garfinkel's procedures with the psychoanalytic terms of the "uncanny" or *unheimlich*; or, with the anti-psychoanalytic aesthetics of Gilles Deleuze and Félix Guattari who elaborate an analysis of repetition and simulacra.

**DUMIT:** Yes.

### **Reiteration**

**SACK:** I feel like we are now starting to repeat things, I also feel like this is the point where we should stop.

**DUMIT:** Yes, we should stop.

**SACK:** What is printed above is structured like a face-to-face interview. It is, in fact, just a small snippet from hours and hours of conversation we have had together face-to-face, over the phone, and via electronic mail. As I sit here now (Where? When?),

---

<sup>61</sup> Fredric Jameson. *The Prison-House of Language: A Critical Account of Structuralism and Russian Formalism* (Princeton, NJ: Princeton University Press, 1972, page 51).

<sup>62</sup> Viktor Shklovsky, "Art as Technique," In L.T. Lemon and M.J. Reis *Russian Formalist Criticism* (Lincoln, NE: University of Nebraska Press, 1965, pages 11-12).

<sup>63</sup> Viktor Shklovsky "O teorii prozy" as cited by Jameson, Op. Cit., page 60.

looking at my computer screen -- editing and rewriting portions of the interview -- I hope I have not been too aggressive in my re-editing of your questions. I know that, in the literature of interviewing techniques, it is quite outside of the "metacommunicative norms"<sup>64</sup> for the interviewee to heavily edit the questions of the interviewer prior to their publication.

**DUMIT:** Well...

**SACK:** I was reluctant to create a publishable form of this interview because I am not finished with my computer program. We have been talking, in theory, about the construction of an artificial participant, but, in practice I have not yet finished designing and implementing the software for my artificial participant; i.e., the supposed focus of this interview. Nevertheless, you have managed, using a variety of interviewing techniques, to draw out a very long series of statements from me.

**DUMIT:** Okay.

**SACK:** Hopefully, when I am done with my artificial participant and when I release it onto the Internet it will work as Garfinkel's experiments do. Hopefully, it will engender a series of deautomatizations.

**DUMIT:** Hmm. Is the common sense of a person going to change once there are devices like this out there? Even now in electronic spaces like MUDs, you have to take into account the fact that quite a few of the "people" you run into are bots. One always has to at least have more than one quick interaction to make sure that it isn't a bot.

**SACK:** Right.

**DUMIT:** The only bot that I wrote said, "Just a minute, I'm busy, I'll be right back."

**SACK:** No! Did you watch it?

**DUMIT:** A few times. It would come into a room, a couple of people would talk to it. They would say, "Hi, welcome". It would say, "Just a minute, I'm busy, I'll be right back." And then a little while later it would say, "I'd really like to talk with you, but something came up." And it would go.

**SACK:** (laughing) Great!

**DUMIT:**

**SACK:**

**DUMIT:**

**SACK:** Hello?

---

<sup>64</sup> Charles L. Briggs. *Learning to Ask: A sociolinguistic appraisal of the role of the interview in social science research* (New York: Cambridge University Press, 1986, page 2).