communication +1

Volume 5
Issue 1 Machine Communication

Article 7

Fall 9-1-2016

Computational Interpersonal Communication: Communication Studies and Spoken Dialogue Systems

David J. Gunkel
Northern Illinois University, dgunkel@niu.edu

Abstract

With the advent of spoken dialogue systems (SDS), communication can no longer be considered a human-to-human transaction. It now involves machines. These mechanisms are not just a medium through which human messages pass, but now occupy the position of the other in social interactions. But the development of robust and efficient conversational agents is not just an engineering challenge. It also depends on research in human conversational behavior. It is the thesis of this paper that communication studies is best situated to respond to this need. The paper argues: 1) that research in communication can supply the information necessary to respond to and resolve many of the open problems in SDS engineering, and 2) that the development of SDS applications can provide the discipline of communication with unique opportunities to test extant theory and verify experimental results. We call this new area of interdisciplinary collaboration "computational interpersonal communication" (CIC).

Keywords

Spoken Dialogue Systems, Communication, Natural Language Processing, Artificial Intelligence

Computational Interpersonal Communication: Communication Studies and Spoken Dialogue Systems

Cover Page Footnote
This essay is the product of a class taught a Northern Illinois University (USA) in the spring semester of 2016. The course, COMS 493: AI, Robots and Communication, was concerned with examining the importance of communication in AI and robotics and the impact of recent developments in these fields on the discipline of communication studies. My thanks to the COMS 493 students for their unbridled enthusiasm for this subject matter and their insightful questions and relentless curiosity.

Whether we recognize it as such or not, we are in the midst of a robot invasion. The machines are everywhere and doing virtually everything. They may have begun by displacing workers on the factory floor, but they now participate in many aspects of our lives. We work with them online. We play with them in digital games. We rely on their recommendations to make informed decisions about everything from the films we watch to potential romantic partners. And we have even begun conversing with them in situations that are beginning to resemble interpersonal communication. Right now these spoken dialogue systems (SDS) tends to be limited to what Gabriel Skantze calls "command-based" interactions[1], which can be seen with a number of recently introduced commercial implementations, like Apple's Siri for the iOS, Microsoft's Cortana, Amazon's Echo/Alexa, and the social robots Jibo and Pepper[2].

The command-based approach to SDS design works reasonably well, because it predetermines much of the semantic context, communicative structure, and social variables, keeping conversational interactions within manageable boundaries. Although interpersonal communication also includes this form of interaction, human-to-human exchanges involve a much wider array of communicative opportunities, contexts, and variables. For this reason, the development of more robust SDS—what Skantze calls "conversational systems"[3]—will rely not only on advancements in engineering but will also require better understanding and modeling of the actual mechanisms and operations of human-to-human communicative behaviors. Unfortunately, the two disciplines that deal with these subjects—engineering and interpersonal communication—have not recognized and/or exploited this interdisciplinary opportunity and challenge. Engineers, for their part, either have tried to reinvent the wheel themselves or have sought advice from research and researchers in other disciplines, like social linguistics or psychology. Communication scholars, who have in fact spent decades studying human-to-human interpersonal relationships and producing the kinds of theories that would be necessary for developing more robust conversational SDS, have not done much better. They have often limited their research efforts and findings to human communication, and when they have dealt with computers or bots, have typically considered the mechanism as a *medium* of human

---

[1] Gabriel Skantze, *Error Handling in Spoken Dialogue Systems: Managing Uncertainty, Grounding and Miscommunication* (Gothenburg, Sweden: Graduate School of Language and Technology. University of Gothenburg, 2007), 11. http://www2.gslt.hum.gu.se/dissertations/ skantze.pdf

[2] Siri and Jibo are two species of "sociable robots." According to Cynthia Breazeal (*Designing Sociable Robots*. Cambridge, MA: MIT Press 2002, 1), "a sociable robot is socially intelligent in a human-like way, and interacting with it is like interacting with another person." Whereas Siri, which is a software application integrated into Apple's iOS, is a Spoken Dialogue System (SDS) or what Andrea Guzman ("Making AI Safe for Humans: A Conversation with Siri," in *Socialbots: Digital Media and the Automation of Sociality*, ed. M. Bakardjieva and R. W. Gehl. New York: Routledge, 2016) also calls a Vocal Social Agent (VSA), Jibo, which is a small table top device with an expressive movable head, is more properly described as an Embodied Conversational Agent (ECA) or "social robot." See Justine Cassell, Joseph Sullivan, Scott Prevost and Elizabeth F. Churchill, *Embodied Conversational Agents* (Cambridge, MA: MIT Press, 2000) and Matthias Rehm and Elisabeth André, "From Annotated Multimodal Corpora to Simulated Human-Like Behaviors," in *Modeling Communication with Robots and Virtual Humans*, ed. Ipke Wachsmuth and Guenther Knoblich (Berlin: Springer, 2012), 1-17.

[3] Skantze, 11.

communicative exchange—what is called “computer mediated communication” or CMC[4]. This essay attempts to remediate this missed opportunity by articulating what research in interpersonal communication might be able to contribute to the design and development of SDS and what efforts in SDS engineering, in turn, are able to contribute to the field of communication studies[5].

1. Talking with Machines

Developing technologies that can work with and produce human language content is the purview of a sub-field of computer science called Natural Language Processing (NLP). Despite the recent proliferation of commercially available “digital assistants” or “intelligent personal assistants” (and the names for these products vary), like Apple’s Siri or Microsoft’s Cortana, NLP is not a new area of research and development. It has been central to both the theory and practice of artificial intelligence (AI).

1.1 AI and NLP

From the beginning, it is communication—and specifically, a tightly constrained form of conversational interpersonal dialogue—that provides AI with its definitive characterization and test case. This is immediately evident in the agenda-setting paper that is credited with defining machine intelligence, Alan Turing's "Computing Machinery and Intelligence," which was first published in the journal *Mind* in 1950. Although the term "artificial intelligence" is a product of the Dartmouth Conference of 1956, it is Turing's seminal paper and the "game of imitation" that it describes—what is now routinely called "the Turing Test"—that defines and characterizes the field. “The idea of the test,” Turing explained in a BBC interview from 1952, “is that the machine has to try and pretend to be a man, by answering questions put to it, and it will only pass

[4] David J. Gunkel, “Communication and Artificial Intelligence: Opportunities and Challenges for the 21st Century.” *Communication +1* 1, no. 1 (2012): 1-25. http://scholarworks.umass.edu/cpo/ vol1/iss1/1/. Steve Jones, “People, Things, Memory and Human-Machine Communication.” *International Journal of Media & Cultural Politics* 10, no. 3 (2014): 245–258

[5] Although the argument presented here is ultimately concerned with making a case for accommodating work in communication studies to developmental efforts in SDS engineering and vice versa, the essay itself is philosophical. That is, it aims to demonstrate the need for and the opportunities and challenges of this interdisciplinary effort, but it does not yet engage in the work that it calls for and characterizes. This is both a practical limitation and a methodological necessity. Practically there is simply not sufficient room in one essay to introduce the research effort and to undertake the many tasks that would be required to execute it. For this reason, the essay seeks the rather modest goal of opening up new avenues of research and cooperation between two fields that are commonly separated and considered somewhat incompatible. Methodologically this effort is consistent with philosophical research, which does not seek to articulate answers to an available problem or set of questions but endeavors to describe the contours of the problem in the first place and formulate the questions that should be asked, but are often not being asked. See Slavoj Žižek, “Philosophy, The ‘Unknown Knowns,’ And the Public Use of Reason.” *Topoi* 25, no. 1 (2006): 137–142. This essay, therefore, isolates and describes what we should be asking about in the face of increasingly social, interactive, and communicative machines in an effort to organize and to direct future work in this area.

if the pretense is reasonably convincing. A considerable proportion of a jury, who should not be experts about machines, must be taken in by the pretense. They aren't allowed to see the machine itself—that would make it too easy. So the machine is kept in a faraway room and the jury are allowed to ask it questions, which are transmitted through to it"[6]. According to Turing's stipulations, if a computer is capable of successfully simulating a human being in communicative exchanges (albeit exchanges that are constrained to the rather artificial situation of typewritten questions and answers) to such an extent that human interlocutors (or "a jury" as Turing calls them in the 1952 interview) cannot tell whether they are talking with a machine or another human being, then that device would need to be considered intelligent[7].

---

[6] Alan Turing, "Can Automatic Calculating Machines Be Said to Think?" in *The Essential Turing*, 487-505 (Oxford: Oxford University Press, 2004), 495.

[7] It is important to recognize the way the concepts of *communication* and *intelligence* have been operationalized in this context. In Turing's game of imitation, communication is understood to be of instrumental value; it is considered to be a sign of intelligence. This instrumentalist formulation is based on the assumption that communicative behavior is made possible by and is an expression of cognitive activity. This way of thinking is ultimately rooted in a classic epistemological limitation that philosophers call "the other minds problems." As Paul Churchland famously characterized it: "How does one determine whether something other than oneself—an alien creature, a sophisticated robot, a socially active computer, or even another human—is really a thinking, feeling, conscious being; rather than, for example, an unconscious automaton whose behavior arises from something other than genuine mental states?" (Paul Churchland, *Matter and Consciousness*. Cambridge, MA: MIT Press, 1999, 67). In other words, when faced with and confronting another entity, I can never be entirely certain that it is another thinking, feeling, intelligent thing like I assume myself to be. The best I can do—the best anyone can do—is to talk to it and read the signs of that interaction as indicative of intellectual activity or not. Turing's game of imitation is predicated on this epistemological exigency. If a machine gives reasonably understandable answers to questions in a text-based conversation, that activity is taken to be evidence that there must be some kind of intelligence behind the communicative interaction. But the equally famous counter example to Turing's Test, John Searle's "Chinese Room," complicates the picture, demonstrating that there can be significant communicative behavior without intelligence. This thought experiment, first introduced in 1980 with the essay "Minds, Brains, and Programs" (*Behavioral and Brain Sciences* 3, no. 3: 417–457) and elaborated in subsequent publications, was initially offered as an argument against the claims of strong AI. "Imagine a native English speaker who knows no Chinese locked in a room full of boxes of Chinese symbols (a data base) together with a book of instructions for manipulating the symbols (the program). Imagine that people outside the room send in other Chinese symbols which, unknown to the person in the room, are questions in Chinese (the input). And imagine that by following the instructions in the program the man in the room is able to pass out Chinese symbols which are correct answers to the questions (the output). The program enables the person in the room to pass the Turing Test for understanding Chinese but he does not understand a word of Chinese." (John Searle, "The Chinese Room," in *The MIT Encyclopedia of the Cognitive Sciences*, ed. R. A. Wilson and F. Keil, 115–116. Cambridge, MA: MIT Press, 1999, 115). The point of Searle's imaginative albeit somewhat ethnocentric illustration is quite simple—simulation is not the real thing. Merely shifting symbols around in a way that looks like linguistic understanding is not really an understanding of the language. Consequently, what the Chinese Room thought experiment demonstrates is that seemingly "intelligent" communicative behavior—the simulation of conversational interaction—is not necessarily an indication of intelligence and cognitive ability. Simply moving linguistic tokens around in such a way

At the time that Turing published the paper proposing this test-case, he estimated that the tipping point—the point at which a machine would be able to successfully play the game of imitation—was at least half-a-century in the future. "I believe that in about fifty years' time it will be possible to programme computers, with a storage capacity of about 10^9, to make them play the imitation game so well that an average interrogator will not have more than 70 per cent chance of making the right identification after five minutes of questioning"[8]. It did not take that long. Already in 1966 Joseph Weizenbaum demonstrated a simple natural language processing application that was able to converse with human interrogators in such a way as to appear to be another person. ELIZA, as the application was called, was what we now recognize as a "chatterbot," although Weizenbaum did not use this terminology[9]. This proto-chatterbot was actually a rather simple piece of programming, "consisting mainly of general methods for analyzing sentences and sentence fragments, locating so-called key words in texts, assembling sentence from fragments, and so on. It had, in other words, no built-in contextual framework of universe of discourse. This was supplied to it by a 'script.' In a sense ELIZA was an actress who commanded a set of techniques but who had nothing of her own to say"[10]. Despite this, Weizenbaum's program demonstrated what Turing had initially predicted:

> ELIZA created the most remarkable illusion of having understood in the minds of many people who conversed with it. People who know very well that they were conversing with a machine soon forgot that fact, just as theatergoers, in the grip of suspended disbelief, soon forget that the action they are witnessing is not 'real.' This illusion was especially strong and most tenaciously clung to among people who know little or nothing about computers. They would often demand to be permitted to converse with the system in private, and would, after conversing with it for a time, insist, in spite of my explanations, that the machine really understood them[11].

1.2 From Chatterbots to SDS

Since the debut of ELIZA, there have been numerous advancements in chatterbot design, and each year many of them are assembled to compete for the Loebner Prize, a competition that has been described as "the first formal instantiation of a Turing Test"[12]. Even though there have been

as to supply answers to questions is not intelligence, even if doing so gives the impression of carrying on an "intelligent dialogue."

[8] Alan Turing, "Computing Machinery and Intelligence," in *Computer Media and Communication: A Reader*, ed. Paul A. Meyer, 37-58 (Oxford: Oxford University Press, 1999), 44.

[9] Gunkel, 5.

[10] Joseph Weizenbaum, *Computer Power and Human Reason: From Judgment to Calculation* (San Francisco, CA: W. H. Freeman, 1976), 188.

[11] Ibid, 189.

[12] Loebner Prize. 2016. http://www.loebner.net/Prizef/loebner-prize.html

impressive performances year after year, no chatterbot has succeeded in "fooling all of the people all of the time," and for this reason, the gold medal has yet to be bestowed. Furthermore, all chatterbots, irrespective of design, inherit two important practical limitations following from Turing's original proposal. First the mode of interaction is restricted to a very narrow range of interpersonal behaviors. Beginning with Turing's game of imitation, chatterbots have been designed as question answering systems. That is, their social involvement is intentionally limited to situations where human interrogators asks questions and the machine is designed to provide responses. This mode of interaction is prescribed by the initial set-up of the game of imitation, and its influence and legacy is evident in published transcripts from the annual Loebner Prize competition. Restricting communicative interaction to Q&A exchanges, although clearly expedient for the purposes of conducting the experiment, is a deliberate and rather artificial restriction that severely limits the range of conversational activity.

Second, these Q&A interactions are restricted to typewritten text. Although one might think that this was done in order to accommodate technical limitations with speech recognition and synthesis, which were all but non-existent during Turing's time, this is not the sole or principal reason. For Turing, and the chatterbots that follow his lead, the use of textual interaction is not a technical exigency; it is a necessary and deliberate element of the imitation game's design. Turing's test was initially formulated in terms of gender: "The new form of the problem," as Turing explains, "can be described in terms of a game which we call the 'imitation game.' It is played with three people, a man (A), a woman (B), and an interrogator (C) who may be of either sex. The interrogator stays in a room apart from the other two. The object of the game for the interrogator is to determine which of the other two is the man and which is the woman"[13]. Consequently, the main reason for limiting the interrogation to text form is to level the playing field: "In order that tones of voice may not help the interrogator the answers should be written, or better still, typewritten. The ideal arrangement is to have a teleprinter communicating between the two rooms"[14]. Obviously restricting interaction to textual exchanges is technically expedient once a computer takes the place of the man or the woman in this elaborate game of gender performance, but that was not, at least according to Turing's formulation, the principal reason for limiting the game to typewritten questions and answers.

Recent developments in SDS have begun to push beyond and to work outside these initial restrictions by focusing on spoken dialogue exchanges where the mode of communication is not necessarily restricted to answering questions. These SDS implementations, especially commercially available products like Siri, Echo/Alexa, and Cortana, are not one technology but consist of an ensemble of several different but related technological innovations: "automatic speech recognition (ASR), to identify what a human says; dialogue management (DM), to determine what that human wants; actions to obtain the information or perform the activity requested; and text-to-speech (TTS) synthesis, to convey that information back to the human in spoken form"[15] (figure 1). Despite their apparent complexity and technical advancement beyond text-based chatterbots like ELIZA, SDSs are still designed for and operate mainly with text data.

---

[13] Turing, "Computing Machinery and Intelligence," 37.

[14] Ibid.

[15] Julia Hirschberg and Christopher D. Manning, "Advancements in Natural Language Processing," *Science* 349, no. 6245 (July 2015): 262. http://science.sciencemag.org/content/349 /6245/261

The principal task of the ASR, a technology that was originally developed in the late 1950's to assist people with physiological disabilities operate computers, is to transform spoken discourse into "a textual hypothesis of the utterance"[16] that can then be processed by the DM. The DM "parses the hypothesis and generates a semantic representation of the utterance" in order to fabricate "a response on a semantic level"[17]. This processing can be accomplished through the application of different NLP methodologies, extending from modified versions of ELIZA's simple keyword search and sentence assembly to more sophisticated systems like Stanford CoreNLP, which is able to work out the syntactic structure of individual sentence by using a dependency grammar analysis[18], and machine learning algorithms trained on various corpora of human conversational interactions. The task of the TTS is to convert the output of the DM, which is typically a text string[19], into intelligible speech sounds. The TTS therefore takes the output generated by the DM, and transforms this data into an audible form that simulates spoken discourse either though concatenation synthesis, which uses a library of prerecorded samples (either whole words or individual phonetic elements), or formant synthesis, which algorithmically produces the audio waveform.

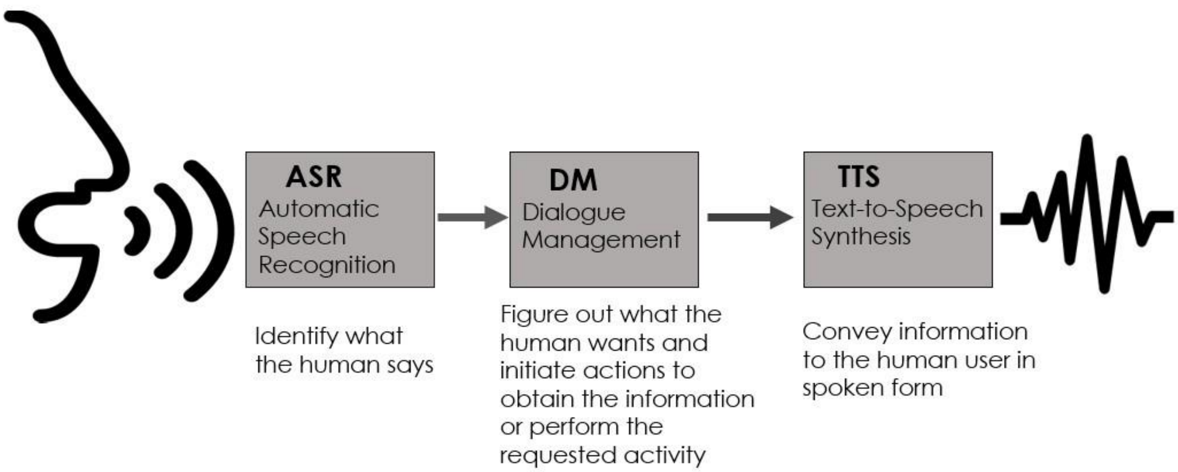

Figure 1 – Simplified block diagram of the technological components of SDS[20]

[16] Skantze, 19.

[17] Ibid.

[18] Hirschberg and Manning, 262.

[19] Skantze, 19.

[20] This simplified block diagram, derived from Hirchberg and Manning's review article in *Science*, is an analytic construct useful for the purpose of representing standard SDS architecture. It is, however, not the only or even the best schematic available. Gabriel Skantze, for instance, provide a more detailed formulation that involves 5 elements: Automatic Speech Recognition (ASR), Natural Language Understanding (NLU), Dialogue Management (DM), Natural Language Generation (NLG), and Text-to-Speech synthesis (TTS). Although Skantze recognizes that some elements in his model could be combined into one integrated module (i.e. "the ASR and NLU may be done in the same processing step"), he argues that a more precise distinction between interacting elements is useful for system

Because of numerous technical improvements, i.e. increased efficiency in data throughput, better algorithms for processing linguistic data like the Hidden Markov Model (HMM), and the application of machine learning capabilities, current SDS application appear to perform rather well, at least in pre-defined circumstances and tightly controlled domains. “Although SDSs now work fairly well in limited domains, where the topics of the interaction are known in advance and where the words people are likely to use can be predetermined, they are not yet very successful in open domain interaction, where users may talk about anything at all”[21]. In other words, commercially available SDS products like Siri, Echo/Alexa, Cortana, and Jibo work well as long as our interactions with them are limited to a narrow range of predefined possibilities: answering questions, providing recommendations, or performing basic actions by delegating requests to cloud-based services. Although representing an impressive development in practical NLP technology, these systems are nowhere near to achieving anything close to what one might considered “normal” interpersonal conversational capabilities.

1.3 Open Problems in SDS Development

So what is still needed? According to Julia Hirschberg and Christopher D. Manning, building more robust SDS systems will require not only improvements in the design and operation of the various technical components that make up these systems but also better knowledge concerning human conversational behavior:

> There are many challenges in building SDSs, in addition to the primary challenge of improving the accuracy of the basic ASR, DM, and TTS building blocks and extending their use into less restricted domains. These include basic problems of recognizing and producing normal human conversational behaviors, such as turn-taking and coordination. Humans interpret subtle cues in speakers’ voices and facial and body gestures (where available) to determine when the speaker is ready to give up the turn versus simply pausing. These cues, such as a filled pause (e.g., “um” or “uh”), are also used to establish when some feedback from the listener is desirable, to indicate that he or she is listening or working on a request, as well as to provide “grounding” (i.e., information about the current state of the conversation)[22].

There are ongoing technological challenges with the design and operation of the various components that make up SDS, like error correction with ASR in noisy environments or latency

---

design and fabrication: “The advantage of the division into components–especially for research systems–is that the components can be developed individually by different developers working with different approaches (and possibly different programming languages), as long as the interfaces between the components are well defined.” (Skantze, 19).

[21] Hirschberg and Manning, 262.

[22] Hirschberg and Manning, 263.

in the processing of data by the DM system. These are technical issues properly addressed by engineers and device manufacturers. But what is also necessary, and what for now remains an "open problem" in effective SDS design and development, is "recognizing and producing normal human conversational behaviors, such as turn-taking and coordination" and making sense of the nonverbal cues and non-linguistic verbal expressions that are commonly used to fill pauses, signal the need for feedback, or provide information on the status and state of the interpersonal relationship.

In addition to these problems, Hirschberg and Manning also mention disambiguation and conversational entrainment. The former refers to the ability that humans have to sort out important differences between words "such as 'yeah' and 'okay,' which may have diverse meanings—including agreement, topic shift, and even disagreement—when spoken in different ways"[23] The latter concerns the way that human interlocutors are able to accommodate their communicative interactions to each other in order to negotiate differences. "In successful and cooperative conversations humans also tend to entrain to their conversational partners, becoming more similar to each other in pronunciation, word choice, acoustic and prosodic features, facial expressions, and gestures"[24] Consequently, there is a wide range of social/interactional issues that need to be properly identified, modeled, and eventually made computable. These are not engineering problems, at least not yet. They are first and foremost a matter of research—observation, data collection, and theory generation—in interpersonal communication.

For the most part, SDS developers have attempted to address these aspects of human conversational behavior by drawing on research from the field of linguistics. Svetlana Stoyanchev, Alex Liu, and Julia Hirschberg, for instance, utilize work in theoretical and applied linguistics to develop a computational model capable of producing more natural clarification questions in dialogue systems[25]. And Gabriel Skantze bases his extensive work with SDS on linguistic research: "Before discussing the components and implementation of dialogue systems we will briefly describe some fundamental properties of human-human conversation from a linguistic perspective"[26]. There is undoubtedly a lot to be obtained from linguistic analysis, and the coupling of linguistics with mathematical modeling and computation—what is called "computational linguistics"—already has a proven record of success. But linguistics generally focuses attention on the elements and operations of language, and in the case of SDS, spoken as opposed to written language. "A useful unit for analysis of written text," Skantze explains, "is the sentence. Sentences are delimited by punctuation marks, where each sentence commonly express one or more propositions. For spoken dialogue, on the other hand, such units are much less adequate for analysis…A unit that is commonly used for segmenting spoken dialogue is instead the *utterance*. In dialogue, speakers exchange utterances, with the intent of affecting the

---

[23] Ibid.

[24] Ibid.

[25] Svetlana Stoyanchev, Alex Liu and Julia Hirschberg, "Towards Natural Clarification Questions in Dialogue Systems," in *AISB Symposium on Questions, Discourse and Dialogue* 20 (2014), http://doc.gold.ac.uk/aisb50/AISB50-S21/AISB50-S21-Stoyanchev-paper.pdf

[26] Skantze, 13.

other speaker in some way"[27] In addition to considering the communicative function of the utterance—what is actually said—linguistic analysis also seeks to parse and process *disfluencies*, like filled pauses, repetitions, repairs, and false starts.[28] Despite this rather broad consideration of conversational activity, however, linguistics is not typically concerned with other, equally important, aspects of the communicative encounter, e.g. social context, tone of voice, spatial proximity, nonverbal behaviors, etc. For this reason, the discipline of linguistics, for all its usefulness in SDS development, does not provide a complete picture of the full range of interpersonal behaviors.

2. SDS and Communication Studies

Many of the widely recognized "open problems" in SDS research and development are precisely what is targeted and studied in communication research, especially efforts in interpersonal communication. Conversely many SDS implementations, like Siri or Jibo, offer unique opportunities to test communication theory, verify the results of communication research, and develop new forms of generating experimental evidence. In other words, it appears that SDS could learn a thing or two from communication research and communication research could, in turn, benefit from many of the opportunities made available by SDS development (figure 2).

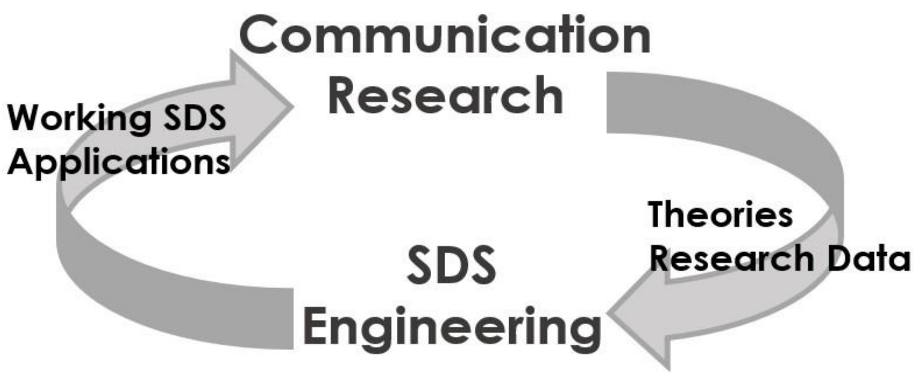

Figure 2 - SDS and Communication Studies

2.1 What Communication Research Can Contribute to SDS

Research in the discipline of communication can supply theory and data to help address and work to resolve open problems in SDS. According to Hirschberg and Manning's review article, one of the fundamental problems in SDS design is "recognizing and producing normal human

[27] Ibid.

[28] Ibid., 17.

conversational behaviors"[29], especially as regards *turn-taking* and the use of both *verbal and non-verbal cues* to coordinate conversational interaction.

Turn-Taking. According to John Wiesmann and Mark Knapp, "the phenomenon by which one interactant stops talking and another starts, in a smooth synchronized manner is considered to be the most salient feature of face-to-face conversation"[30] and research focusing on experimental studies of face-to-face human interaction begins to develop traction in the 1970s. In 1972 Starkey Duncan published a widely cited study in the *Journal of Personality and Social Psychology*, which sought to identify the mechanisms of turn-taking through experimental observation of face-to-face human interactions[31]. And in 1974, Harvey Sacks, Emanuel Schegloff and Gail Jefferson published what is considered to be the seminal research paper in conversational turn-taking, establishing what many consider to be the benchmark for speech-exchange systems[32]. Since that time, there have been numerous empirical investigations of turn-taking in both face-to-face interpersonal interactions and group decision making. More recent studies have focused on the way turn-taking is effected and modified by technological mediation though telecommunications, video conferencing, and computer mediated communication[33]. Although a good deal of the published research in conversational turn-taking involves modeling linguistic performance[34], there is a significant aspect of it, as Wiesmann and Knapp point out, that depends on nonverbal contributions[35]. As a result, "turn-taking" is and remains one of the main topics covered in standard textbooks addressing interpersonal communication. A more complete inventory and understanding of all the variables and factors involved in regulating

---

29 Hirschberg and Manning, 263.

30 John M. Wiemann and Mark L. Knapp, "Turn-Taking in Conversations," *Journal of Communication* 25, no. 2 (1975), 75.

31 Starkey Duncan, "Some Signals and Rules for Taking Speaking Turns in Conversations," *Journal of Personality and Social Psychology* 23, no. 2 (1972): 283-292.

32 Harvey Sacks, Emanuel Schegloff and Gail Jefferson, "A Simplest Systematics for the Organization of Turn-Taking for Conversation," *Language* 50, no. 4 (1974): 696-735.

33 Starr Roxanna Hiltz, Kenneth Johnson and Murray Turoff, "Experiments in Group Decision Making Communication Process and Outcome in Face-to-Face Versus Computerized Conferences," *Human Communication Research* 13, no. 2 (1986): 225-252. Sara Kiesler and Lee Sproull, "Group Decision Making and Communication Technology," *Organizational Behavior and Human Decision Processes* 52, no. 1 (1992): 96-123. Robert Hopper, *Telephone Conversation* (Bloomington, IN: Indiana University Press, 1992). Mia Lobel, Michael Neubauer, and Randy Swedburg, "Comparing How Students Collaborate to Learn About the Self and Relationships in a Real-Time Non-Turn-Taking Online and Turn-Taking Face-to-Face Environment," *Journal of Computer-Mediated Communication* 10, no. 4 (2005). http://onlinelibrary.wiley.com/doi/10.1111/j.1083-6101.2005.tb00281.x/full

34 Duncan, "Some Signals and Rules for Taking Speaking Turns in Conversations." Sacks, Schegloff, and Jefferson, "A Simplest Systematics for the Organization of Turn-Taking for Conversation." J. C. P. Auer, "Review of B. Oreström's Turn-taking in English Conversation," *Linguistics* 21 (1983): 742–748. Bengt Oreström. *Turn-taking in English conversation* (Lund, Sweden: Liber 1983). D. C. O'Connell, S. Kowal and E. Kaltenbacher, "Turn-Taking: A Critical Analysis of the Research Tradition," *Journal of Psycholinguistic Research* 19, no. 6 (1990), 345-373.

35 Wiemann and Knapp, 75.

human-to-human conversational turn-turning may help with the design of human-machine SDS that are more natural, engaging, effective, and personal.

Verbal and non-verbal Cues. A good deal of conversational interaction is negotiated through nonverbal elements, which can include, visual cues, or "body language," vocal intonation or paralanguage, chronemics, and oculesics like eye contact and gaze direction. Turn-taking, for instance, is often indicated by either a change in vocal tone, a pause in the temporal sequence of the verbal delivery, or a visual cue, such as a nod or glance. Although research in these aspects of human communication have been pursued and published in the fields of semiotics[36] and the social sciences of anthropology and psychology[37], it is communication studies that has staked a claim to this particular area of research since the mid-1970s[38]. Since this time, attending to the importance of both verbal and non-verbal cues has become not just a legitimate area of communication research, but is considered one of the central concerns of the discipline. And recent publications in the field, like that of Jones and LeBaron have sought to correlate the study of the nonverbal and verbal components, which have been historically distinguished, in order to formulate "more integrated approaches to the study of verbal and nonverbal communication so that more holistic understandings of social interaction may emerge"[39]. Right now commercially available SDS applications, like Siri and Echo/Alexa, are only attending to what is said. How it is said and in what particular fashion it is articulated is not necessarily part of the current implementations. Siri in other words, can parse and process the words that users speak but the system is currently unable to make sense of the pauses, the social context, the vocal tone, and the gestures made by users in the process of speaking. But it is possible to imagine a more sophisticated implementation of SDS that would be capable of processing these other elements to assist the effectiveness and efficiency of communicative interaction between human users and the SDS application/appliance.

2.2. What SDS Can Do for Communication Research

The manufacture of working SDS implementations can provide research in communication studies with unique opportunities to verify results, test conclusions, and even improve methodology. Currently, theory produced in communication research is typically tested and verified in experiments with human subjects. SDS provides researchers with some other options and research opportunities. First, the design of working SDS systems will require that various

[36] Paul Ekman and Wallace Friesen, "The Repertoire of Nonverbal Behavior: Categories, Origins, Usage, and Coding," *Semiotica* 1, no. 1 (1969): 49-98. Jacqueline Lindenfeld, "Verbal and Non-verbal Elements in Discourse," *Semiotica* 3, no. 3 (1971): 223-233.

[37] R. G. Harper, A. N. Wiens and J. D. Matarazzo. *Nonverbal Communication: The State of the Art*. (New York: John Wiley & Sons, 1978). Marvin A. Hecht and Nalini Ambady, "Nonverbal Communication and Psychology: Past and Future," *Atlantic Journal of Communication* 7, no. 2 (1999): 156-170.

[38] Judee K. Burgoon, "Nonverbal Communication Research in the 1970s: An Overview," *Communication Yearbook* 4 (1980): 179-197.

[39] Stanley E. Jones and Curtis D. LeBaron, "Research on the Relationship Between Verbal and Nonverbal Communication: Emerging Integrations," *Journal of Communication* 52, no. 3 (2002): 499.

concepts and theories of human communication be made computable and employed to control the behaviors of various types of SDS operations. The development of SDS, therefore, offers opportunities for the testing of theory through the construction of mechanisms that employ or embody a particular theoretical model. In the process, results obtained from experimenting with different instantiations of theory can be reflected back into the discipline of communication for improving the accuracy and formulation of the theoretical models. This approach has proven to be extremely useful in other fields, like computational linguistics. As Hirschberg and Manning explain, "the development of probabilistic approaches to language [necessary for SDS development] is not simply about solving engineering problems: Probabilistic models of language have also been reflected back into linguistic science," and, as a result, "many areas of linguistics are themselves becoming more empirical and more quantitative in their approaches"[40]. It is likely that similar outcomes would be obtained from developing probabilistic approaches to interpersonal communication.

Second, the effectiveness of operational implementations of SDS will need to be tested and evaluated in actual interactions with human users. Although the immediate goal of this effort might be to "stress test" the design in actual social circumstances, these situations will also provide researchers working in communication with a unique opportunity to investigate how human subjects interact with other kinds of communicative agents. In other words, because the manufacturing and marketing of more effective SDS will require countless hours of controlled testing with human users, scholars of communication will have a unique opportunity to study new forms of social interaction and to use this data for both SDS improvement and theory development. This will be crucial not only for the discipline of communication studies but for the social sciences in general. As Norbert Wiener, the progenitor of the science of cybernetics, accurately predicted over a half-century ago: "It is the thesis of this book [*The Human Use of Human Beings*] that society can only be understood through a study of the messages and the communication facilities which belong to it; and that in the future development of these messages and communication facilities, messages between man and machines, between machines and man, and between machine and machine, are destined to play an ever-increasing part"[41]. In the social relationships of the 21st century, machines will no longer comprise mere instruments or media through which human users communicate and socialize with each other. Instead they will increasingly come to occupy the position of another social actor with whom one communicates and interacts, and communication studies is in a unique position to investigate and develop knowledge about this new social arrangement and its consequences.

2.3 By Way of Example

These suggestions are just that, suggestions. Demonstration of the usefulness of interpersonal communication research for the design and development of SDS and the expediency of SDS implementations for communication research will only be achieve in practice. This kind of interdisciplinary work—the very work this essay seeks to identify and initiate—remains to be

[40] Hirschberg and Manning, 266.

[41] Norbert Wiener, *The Human Use of Human Beings: Cybernetics and Society* (Boston, MA: Da Capo Press, 1988), 16.

undertaken. At this point, the best we can do is project the opportunities and challenges of this effort by considering a concrete example: the effect of eye-gaze, non-verbal communication on conversational turn taking.[42] This is the subject of a recent article by Gabriel Skantze, Anna Hjalmarsson, and Catharine Oertel. Their study applies findings from research with human-to-human interactions to robots in order to model and study the effects of non-verbal eye gaze behaviors on human/robot interactions[43]. The investigation sounds promising insofar as researchers are engaged with available studies in human communication and are applying results of this research to the design and operation of social robots. A quick look through the article's bibliography, however, reveals an obvious lack of research from the field of interpersonal communication, even though communication scholars have been involved with the study of this subject matter and have made considerable contributions to the research literature[44].

This absence is not necessarily a problem. Identifying the lack of something is never sufficient evidence that it should have been included in the first place. What is needed, therefore, is not just an indication that something is missing but a demonstration that shows how the inclusion of this absent material would be able to add a crucial dimension currently unavailable in the works that have been cited. In other words, we would need to demonstrate that research in communication studies has some significant insight to contribute to this area of investigation such that its exclusion represents a missed opportunity for developing better and/or more robust forms of research. This is, in fact, the case with research in nonverbal communication. Typically scholars have divided verbal and non-verbal communication into separate channels of message transmission such that information supplied by one channel can be reinforced, re-contextualized, or even subverted by the information supplied by the other[45]. This differentiation persists in many of the studies conducted with human-robot interaction, mainly because the division is already operationalized by the literature researchers have called upon and due to the fact that this kind of separation assists computational modeling, i.e. each channel (verbal and nonverbal) can be isolated and independently modeled and controlled. Research efforts in interpersonal communication, however, have begun to challenge this conceptual duality[46] and have sought to devise more integrated approaches that consider the verbal and nonverbal elements as interacting components of a holistic system[47]. This alternate way of framing and modeling conversational

---

[42] Although gaze behavior is not a pressing concern for the design of speech-only SDSs, like Siri, it is increasingly important for more sophisticated implementations, like Jibo, which has a movable head with a round faceplate screen that can be repositioned in such a way that the device appears to be "looking at" users.

[43] Gabriel Skantze, Anna Hjalmarsson and Catharine Oertel, "Turn-Taking, Feedback and Joint Attention in Situated Human–Robot Interaction," *Speech Communication* 65, no. 1 (2014): 50–66.

[44] For an overview, see Charles R. Berger, "Interpersonal Communication: Theoretical Perspectives, Future Prospects," *Journal of Communication* 55, no. 3 (2005): 415-447.

[45] Jones and LeBaron, "Research on the Relationship Between Verbal and Nonverbal Communication."

[46] Jürgen Streeck and Mark L. Knapp, "The Interaction of Visual and Verbal Features in Human Communication," in *Advances in Nonverbal Communication*, ed. F. Poyatos (Amsterdam: Benjamins, 1992), 3-23.

[47] Jones and LeBaron, "Research on the Relationship Between Verbal and Nonverbal Communication," 500.

behavior could supply improvements for SDS design and, as a result, produce better or "more natural" human-robot conversational interactions. Obviously, however, this is still just a hypothesis. Its definitive test and demonstration will require practical implementations. But such implementations need to begin by recognizing the potential that is already available in this marginalized body of literature.

At the same time, the design of SDS systems can be used to model interpersonal communication scenarios in such a way that researchers are able to have precise control over experimental variables. This is precisely what Skantze, Hjalmarsson and Oertel pursue in their study: "We have systematically manipulated the way the robot produces turn-taking cues. We have also compared the face-to-face setting … with a setting where the robot employs a random gaze behaviour, as well as a voice-only setting where the robot is hidden behind a paper board. This way, we can explore what the contributions of a face-to-face setting really are, and whether they can be explained by the robot's gaze behaviour or the presence of a face at all"[48]. This kind of systemic manipulation of experimental variables could be very useful to communication researchers. For example, Judee K. Burgoon and colleagues have published a series of influential studies concerning the effects of eye gaze on social perceptions and outcomes[49]. These studies have, following a standard practice in interpersonal communication research, utilized "confederates," who were trained by the researchers to produce three different levels of eye gaze response in simulated interviews with test subjects: "The confederate interviewees were six undergraduate students, three males and three females, who were trained to keep their verbal replies and all other nonverbal behavior consistent across interviews. The eye contact manipulation consisted of one of three levels: high, medium or low"[50]. Although expedient for the purposes of conducting the study, utilizing confederates introduces significant limitations and unwanted variability. "While the numerous significant differences due to confederates are an experimenter's nightmare, they do underscore the need to conduct interpersonal communication research using multiple confederates. The current findings strongly demonstrate the idiosyncratic differences in individual communication styles that may mediate communication outcomes. Had only one or two of the confederates in this experiment actually been used, very different results might have appeared. One has to wonder how many interpersonal experiments have been subject to this kind of undetected confound"[51]. This confounding problem, as the researchers clearly point out, is not unique to this one study but constitutes a persistent and often unidentified difficulty across interpersonal communication research in general. Using tightly controlled and programmable SDS implementations for this kind of research might provide better experimental controls than can be obtained with any number of trained confederates, mainly because

---

[48] Skantze, Hjalmarsson and Oertel, 51.

[49] Judee Burgoon, Valerie Manusov, Paul Mineo and Jerold U. Hale, "Effects of Gaze on Hiring, Credibility, Attraction and Relational Message Interpretation," *Journal of Nonverbal Behavior* 9, no. 3 (1985): 133-146. Judee Burgoon, Deborah A. Coker and Ray A. Coker, "Communicative Effects of Gaze Behavior: A Test of Two Contrasting Explanations," *Human Communication Research* 12, no. 4 (1986): 495-524. A special note of thanks to my colleague Mary Lynn Henningsen for helping me navigate through the literature of interpersonal communication research.

[50] Burgoon, Manusov, Mineo and Hale, 137.

[51] Ibid., 143.

researchers can directly and precisely control each variable in the experiment and ensure consistent behavior across multiple trials.

3. Limitations and Directions for Future Research

At the end, this investigation arguably produces more questions than definitive answers, and that is by design. The purpose of the essay is to open up new avenues of research that can contribute to both SDS development and research in interpersonal communication. We end, therefore, not with a set of conclusive outcomes that put an end to investigation, but with an indication of future research possibilities and the identification of one important caveat.

3.1 New Opportunities and Reciprocal Benefits.

Although it has not been widely recognized, research in communication can supply the data and theories necessary to develop more robust SDS. Extant communication theories—theories that have been produced and tested in countless hours of experimental observations of human-to-human social interactions in a variety of situations and contexts—can help in the design of SDS applications by supplying generalizable models that can be used to develop computable program instructions. Likewise raw data—typically video and audio recordings and anonymized written transcripts from communication research projects—can furnish the material for training learning algorithms and neural networks on the standard patterns of human communicative behavior. The best design strategies will, following the recent success of Google DeepMind's AlphaGo, probably draw on both methodologies—the predictions available in theory and the experiences of learning in practice—to develop future SDS applications. At the same time, work in SDS will produce operationalized applications of communication research that can be tested in actual encounters with human subjects. This will not only provide communication researchers the unique opportunity to validate available theories and evidence but will also, we predict, lead to new opportunities for communication researchers. In much the same way that computational linguistics made the discipline of linguistics more empirical and data driven, we can expect that SDS in particular and human-machine communication in general will transform communication studies from a "soft" social science to something more empirical and quantitative.

3.2 Interdisciplinary Approaches

This effort requires the development of an interface between the fields of engineering and communication studies. We will, therefore, need the equivalent of an academic API (application program interface) for these two disciplines. This is going to necessitate, on the one hand, mining the literature of communication studies and porting its finding in such a way that they can be utilized outside the discipline in which they were initially cultivated and developed. Doing so will involve making theory computable so that the insights that have been generated by decades of communication research are not just human readable but are also rendered machine executable. At the same time, and on the other hand, engineers will need to learn to recognize and to appreciate how scholarship in this so-called "soft science" can speak to and contribute the data necessary to address many of the open problems in SDS development. As Hirschberg and Manning recognize, the current crop of open problems in SDS design concern not only engineering challenges but also better understanding of human conversational behavior. Although there is some small movement in this direction as is evident in the work of Hirschberg and Manning and Skantze[52], this is still a wide open and largely untapped intellectual resource. Finally, mobilizing these interdisciplinary connections is not something that is or should be limited to the specific case of SDS design. Similar opportunities and challenges are available with embodied conversational agents (ECA) and social robots. In fact, it is in these other areas that research in communication—especially as concerns nonverbal forms of communicative interaction—would be most needed but has been, for now at least, conspicuously absent.

3.2 Underlying Assumptions

One fundamental assumption behind this effort is that SDS *should* emulate or simulate human-level communicative behaviors. But this is an assumption, as Skantze accurately recognizes:

> An argument for moving towards conversational dialogue, as opposed to a command-based, is that human-like conversation generally is considered to be a natural, intuitive, robust and efficient means for interaction. Thus, the advantage of command-based speech interfaces over traditional graphical user interfaces is often restricted to the fact that users may use the hands and eyes for other tasks, and their usefulness may thus be limited to special contexts of use, such as when driving a car. Conversational dialogue systems hold the promise of offering a more intuitive and efficient interaction. Whether this promise will be met remains to be seen[53].

This insight has been verified by James R. Glass: “While it is clear that the study of human-human conversations can provide valuable insights into the nature of dialogue, it is still a matter of debate how human-like spoken dialogue systems should be. The ability to handle phenomena commonly used in human conversations could ultimately make systems more natural and easy to use by humans, but they also have the potential to make things more complex and confusing”[54].

We began by following the suggestion of Hirschberg and Manning that SDS can and should be more conversational. This was not questioned, because it seems, at least to begin with, to be rather intuitive. But modeling human-to-machine (h2m) communication on human-to-human (h2h) communication might be the wrong place to begin, just as modeling “machine intelligence” on human cognition turned out to be a significant impediment to progress in

[52] Gabriel Skantze, “Exploring Human Error Recovery Strategies: Implications for Spoken Dialogue Systems,” *Speech Communication* 45 (2005): 325–341.

[53] Skantze, *Error Handling in Spoken Dialogue Systems*, 12.

[54] James R. Glass, “Challenges for Spoken Dialogue Systems,” in *Proceedings of the 1999 IEEE ASRU Workshop* (December 1999): 3. http://www.cs.cmu.edu/afs/cs.cmu.edu/Web/People/dod/papers/glass99.pdf

artificial intelligence[55]. Identifying this assumption, however, does not mitigate against the argument for including communication research in SDS development. In fact, it actually makes such research more important and valuable. If fabricating human level conversational agents is in fact a worthwhile objective (i.e. the assumption is true), then research in interpersonal communication will provide the necessary data and theories to inform this effort. If, however, the opposite is the case, and human-level communicative behavior for h2m is not in fact more effective and efficient, then it is research in interpersonal communication that will help prove this point by assembling the data necessary to disprove the initial assumption. Either way, research in communication will be crucial to successful SDS design, development, and implementation.

[55] Rodney A. Brooks, "Intelligence Without Representation," *Artificial Intelligence* 47 (1991): 139–159.

Bibliography

Auer, J. C. P. "Review of B. Oreström's Turn-taking in English Conversation." *Linguistics*. 21 (1983): 742–748.

Berger, Charles R. "Interpersonal Communication: Theoretical Perspectives, Future Prospects." *Journal of Communication* 55.3 (2005): 415-447.

Breazeal, Cynthia. *Designing Sociable Robots*. Cambridge, MA: MIT Press, 2002.

Brooks, Rodney. A. "Intelligence Without Representation. *Artificial Intelligence* 47 (1991): 139–159.

Burgoon, Judee K. "Nonverbal Communication Research in the 1970s: An Overview." *Communication Yearbook* 4 (1980): 179-197.

Burgoon, Judee K., Valerie Manusov, Paul Mineo and Jerold U. Hale. "Effects of Gaze on Hiring, Credibiliity, Attraction and Relational Message Interpretation." *Journal of Nonverbal Behavior* 9.3 (1985): 133-146.

Burgoon, Judee K., Deborah A. Coker and Ray A. Coker. "Communicative Effects of Gaze Behavior: A Test of Two Contrasting Explanations." *Human Communication Research* 12.4 (1986): 495-524.

Cassell, Justin., Joseph Sullivan, Scott Prevost and Elizabeth Churchill. *Embodied Conversational Agents*. Cambridge, MA: MIT Press, 2000.

Churchland, Paul. *Matter and Consciousness*. Cambridge, MA: MIT Press, 1999.

Duncan, Starkey. "Some Signals and Rules for Taking Speaking Turns in Conversations." *Journal of Personality and Social Psychology* 23.2 (1972): 283-292.

Ekman, Paul and Wallace V. Friesen. "The Repertoire of Nonverbal Behavior: Categories, Origins, Usage, and Coding." *Semiotica*, 1.1 (1969): 49-98.

Glass, James. R. "Challenges for Spoken Dialogue Systems." *Proceedings of the 1999 IEEE ASRU Workshop* (December 1999). http://www.cs.cmu.edu/afs/cs.cmu.edu/Web/People/dod/ papers/glass99.pdf

Gunkel, David. J. "Communication and Artificial Intelligence: Opportunities and Challenges for the 21st Century." *Communication +1* 1.1 (2012): 1-25. http://scholarworks.umass.edu/cpo/ vol1/iss1/1/

Guzman, Andrea. "Making AI Safe for Humans: A Conversation with Siri." In *Socialbots: Digital Media and the Automation of Sociality*, ed. M. Bakardjieva and R. W. Gehl. New York: Routledge, 2016.

Harper, R. G., A. N. Wiens and J. D. Matarazzo, *Nonverbal Communication: The State of the Art*. John Wiley & Sons, 1978.

Hecht, Marvin. A., & Nalini Ambady. "Nonverbal Communication and Psychology: Past and Future." *Atlantic Journal of Communication 7*.2 (1999): 156-170.

Hiltz, Starr Roxanne, Kenneth Johnson and Murray Turoff. "Experiments in Group Decision Making Communication Process and Outcome in Face-to-Face Versus Computerized Conferences." *Human Communication Research* 13.2 (1986): 225-252.

Hirschberg, Julia and Christopher D. Manning. "Advancements in Natural Language Processing." *Science*, 349.6245 (July 2015): 261-266. http://science.sciencemag.org/content/349/6245/261

Robert Hopper, *Telephone Conversation*. Bloomington, IN: Indiana University Press, 1992.

Jones, Stanley E. and Curtis. D. LeBaron. "Research on the Relationship Between Verbal and Nonverbal Communication: Emerging Integrations." *Journal of Communication* 52.3 (2002): 499-521.

Jones, Steve. "People, Things, Memory and Human-Machine Communication." *International Journal of Media & Cultural Politics* 10.3 (2014): 245–258.

Kiesler, Sara and Lee Sproull. "Group Decision Making and Communication Technology." *Organizational Behavior and Human Decision Processes* 52.1 (1992): 96-123.

Lindenfeld, Jacqueline. "Verbal and Non-verbal Elements in Discourse." *Semiotica* 3.3 (1971): 223-233.

Lobel, Mia, Michael Neubauer and Randy Swedburg. "Comparing How Students Collaborate to Learn About the Self and Relationships in a Real-Time Non-Turn-Taking Online and Turn-Taking Face-to-Face Environment." *Journal of Computer‐Mediated Communication* 10.4 (2005). http://onlinelibrary.wiley.com/doi/10.1111/j.1083-6101.2005.tb00281.x/full

Loebner Prize. 2016. http://www.loebner.net/Prizef/loebner-prize.html

O'Connell, D. C., S. Kowal and E. Kaltenbacher. "Turn-Taking: A Critical Analysis of the Research Tradition." *Journal of Psycholinguistic Research* 19.6 (1990), 345-373.

Oreström, Bingt. *Turn-taking in English conversation*. Lund, Sweden: Liber, 1983.

Rehm, Matthias and Elisabeth André. "From Annotated Multimodal Corpora to Simulated Human-Like Behaviors." In *Modeling Communication with Robots and Virtual Humans*, ed. Ipke Wachsmuth and Guenther Knoblich, 1-17. Berlin: Springer, 2008.

Sacks, Harvey, Emanuel Schegloff and Gail Jefferson. "A Simplest Systematics for the Organization of Turn-Taking for Conversation." *Language* 50.4 (1974): 696-735.

Searle, John. "Minds, Brains, and Programs." *Behavioral and Brain Sciences* 3.3 (1980): 417–457.

Searle, John. "The Chinese Room." In *The MIT Encyclopedia of the Cognitive Sciences*, ed. R. A. Wilson and F. Keil, 115–116. Cambridge, MA: MIT Press, 1999.

Skantze, Gabriel. "Exploring Human Error Recovery Strategies: Implications for Spoken Dialogue Systems." *Speech Communication* 45 (2005): 325–341.

Skantze, Gabriel. *Error Handling in Spoken Dialogue Systems: Managing Uncertainty, Grounding and Miscommunication.* Graduate School of Language and Technology. University of Gothenburg, Sweden, 2007. http://www2.gslt.hum.gu.se/dissertations/skantze.pdf

Skantze, Gabriel, Anna Hjalmarsson and Catharine Oertel. "Turn-Taking, Feedback and Joint Attention in Situated Human–Robot Interaction." *Speech Communication* 65. 1 (2014): 50–66.

Stoyanchev, Svetlana, Alex Liu and Julia Hirschberg. "Towards Natural Clarification Questions in Dialogue Systems." In *AISB Symposium on Questions, Discourse and Dialogue*, vol. 20, 2014. http://doc.gold.ac.uk/aisb50/AISB50-S21/AISB50-S21-Stoyanchev-paper.pdf

Streeck, Jürgen and Mark L. Knapp. "The Interaction of Visual and Verbal Features in Human Communication." In *Advances in Nonverbal Communication*, ed. F. Poyatos, 3-23. Amsterdam: Benjamins, 1992.

Turing, Alan. "Computing Machinery and Intelligence." In *Computer Media and Communication: A Reader*, ed. Paul A. Meyer. Oxford: Oxford University Press, 1999.

Turing, Alan. "Can Automatic Calculating Machines Be Said to Think?" *The Essential Turing*, 487-505. Oxford: Oxford University Press, 2004.

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

Wiemann, John M., & Mark L. Knapp. "Turn-Taking in Conversations." *Journal of Communication* 25.2 (1975): 75-92.

Wiener, Norbert. *The Human Use of Human Beings: Cybernetics and Society*, Boston, MA: Da Capo Press, 1988.

Žižek, Slavoj. "Philosophy, The 'Unknown Knowns,' And the Public Use of Reason." *Topoi* 25.1 (2006): 137–142.