Asian Journal of Information Technology 16 (1): 24-31, 2017
ISSN: 1682-3915

Technogenic Institutional Talk in an Automated Computer-Telephone Interviewing System

Nils Oliver Klowait
Moscow School of Social and Economic Sciences, Moscow, Russia

Abstract: The study aims to conceptualize and evaluate a phone-based, natural-language-employing automated computer-telephone interviewing system. It will be argued that the conversational agent by virtue of its technical limitations is situated squarely within the interactional 'uncanny valley: precisely because it exhibits a rudimentary interactivity and can thereby mimic human agency its inability to be fully humanlike becomes a peculiar interactive feature. The system is shown to take on the role of a highly restrictive interrogator rather than a regular interviewer: it generates 'institutional talk'. This is shown to be the especially the case when users fail to recognize the system as non-human. The findings problematize the overall methodological robustness of state-of-the-art automated surveying agents as such systems may unwittingly introduce response biases to a supposedly impersonal surveying method. Conceptually, the study will be grounded in suchman's 'situated action' paradigm of human-computer interaction as well as Heritage's 'institutional talk' within conversation analysis.

Key words: Computer human interface, situated action, human-computer interaction, institutional talk, analysis

INTRODUCTION

The study aims to evaluate the interactional peculiarity of Automated Computer Telephone Interviewing (ACTI). ACTI allows researchers to record the survey questions once, feed it into the program and subsequently have the ACTI System use that blueprint to call as many respondents as required, using the originally recorded survey questions. Such a system is arguably capable of eliminating interviewer bias as there is no variation in how the survey questions are asked each time. Furthermore, this technology allows for large-scale cost-efficient polling with maximal autonomy. It is therefore likely to be employed with greater frequency in the future. As the respondent is never directly interacting with a human being, this technology may also help when it comes to asking traditionally sensitive questions (Blackstone *et al*., 2009; Corkrey and Parkinson, 2002; Kreuter *et al*., 2008; Pickard *et al*., 2016).

The striking feature of this interactional agent is that it makes it is difficult for the respondent to determine whether they are talking to a computer or a human. As the voice is recorded (i.e., not synthesised via text-to-speech), the illusion of interpersonal conversation can be sustained for a prolonged period. The rudimentary abilities to react to the respondent through context-dependent questions (e.g., based on the previous answers) as well as technical remarks (e.g., "sorry, I could not hear you properly, please repeat your answer") further contribute to the illusion.

In fact, some recordings of these electronic encounters suggest that the respondent does not ultimately realise the non-human nature of the conversation partner. Should this be the case, individually or systematically, an investigation of the consequences of this understanding would help reveal the effects both for the respondent and the survey methodology. For the purposes of this study it shall therefore be assumed as a working axiom that there is the possibility of not understanding that the partner is a talking robot.

The main question is consequently: assuming that the respondent thinks that they are talking to a human being, what kind of interaction order is thereby constituted in the unfolding encounter? In other words, what are the concrete ways in which the technological make-up of the non-human agent can influence, direct or even dictate the occurring interaction?

This study is situated in the tradition of merging the study of Human-Computer Interaction (HCI) with ethnomethodological and Conversation-Analytical (CA) traditions of investigations into the ad hoc interactive and procedural sense-making of concrete human interaction. The choice of utilising ethnomethodology and CA is partly due to the rich ethnomethodological tradition of HCI ever since suchman's plans and situated actions (Suchman, 1987) and more practically due to what is found in the available data: it's talk. Recorded, particular, context-sensitive talk. No matter how peculiar the conversation partners may be, there is no reason to outright deny that interactional rules apply to

them. Whether they do and the effect they may have on this concrete human-computer interaction is the fundamental driving force behind this investigation.

As such, this investigation is further situated in the context of the investigation of anthropomorphic robotic or virtual interactive agents. This research agenda has gained particular traction within the last decade, particularly with the emergence of the 'media equation' paradigm (Nass and Brave, 2005; Nass *et al*., 1994; Reeves and Nass, 1996). Researchers within that paradigm aim to investigate the phenomenon of the human tendency to apply human rules of conduct when engaging certain technologies such as for example, voiced car navigation systems (Nass *et al*., 1997). Nass and Moon (2000) argues that humans, if approached by a non-human agent with a certain number of features (such as having a voice or a face) will subconsciously respond to these cues and act as though they were interacting with a person (Nass and Moon, 2000).

The paradigm's inception has sparked a large number of research projects dealing with seemingly anthropomorphic responses to computers (Brahnam and Angeli, 2012; Eyssel and Hegel, 2012; Kim and Sundar, 2012; Liang *et al*., 2013; Meuse, 1987; Sundar and Nass, 2000; Tay *et al*., 2014; Tay *et al*., 2016; Zhu and Kaber, 2012). Although the present study does not engage the question of why agency may be ascribed to a computer (as it is interested in the mechanics of how this may happen), the media equation may indeed be one of the more fruitful explanatory vectors.

The following study will introduce the reader to the concrete data used as well as the methodological and theoretical approaches employed in the analysis.

MATERIALS AND METHODS

The findings are based on an analysis of 175 audio recordings and corresponding transcripts created by the present author. The available sample is part of a bigger dataset. The survey aimed to collect data on voting preferences in the 2013 Moscow Oblast Gubernatorial Elections. Each human-ACTI encounter occurred in a linear-nonadaptive fashion with the responses lacking branching, follow-up questions or a consistent skip logic. The conversation length ranged from approximately 6-16 min with an average interview time of 8.5 min. The interview questions were for the most part, recorded by a single female interviewer. There was no explicit indication of the robotic nature of the system, the recorded interviewer employed personalized first person language both in the survey questions and in the 'repair responses' (additional probing after undiscernible responses).

The researchers shares the position of ethnomethodology that talk cannot be made sense of through mere reference to ideal language constructions we must go beyond grammatical or semantic structure and the final sentence-meaning thereby inscribed. If therefore there is any hope at treating the present talk recording as data and not just sentence-meaning, the level of conversation analysis is the first acceptable level of depth.

The project of ethnomethodology involved treating social order as a constant, all-pervasive, everyday and routine practical achievement (Garfinkel, 1967). Consequently, the central consideration of sociology was proposed to become the question of how, on a local and interactive level order is created and sustained. This could only be done on a highly localized basis, generating extremely detailed, 'thick' accounts of interaction in order to understand how members, doing routine interaction for the sake of routine interaction, create and maintain a specific local order with stable identities, relations and available actions.

A central concept of this building of order is the concept of accountability: every action within a context-being-ordered is reflexively accountable to the collective that is creating and re-creating local social orderings. This has the important consequence that, if social action is made reflexively accountable as the tool for meaning-making, it cannot be naive: a silence becomes reactive, a 'goodbye' meaningful beyond its dictionary definition.

This general approach was picked up by Sacks and developed into conversational analysis (Schegloff, 1992). In parallel to Garfinkel, sacks looked at talk as a practical action: instead of looking at talk from an abstracted linguistic perspective of meaning and reference, Sacks focused on concrete-for-practical-purposes talk as the means of conversational ordering, spending considerable time studying the minutiae of conversations to reveal their functioning.

To illustrate, the following is one of the first examples of the Sacksian approach to talk. Having obtained recordings of the interaction between a suicide prevention hotline and its callers, sacks chose to investigate when and how the callers manage to provide or not provide their names to the hotline:

- A: this is Mr. Smith may I help you
- Yes, this is Mr. Brown
- A: this is Mr. Smith may I help you
- B: I can't hear you
- A: this is Mr. Smith
- B: Smith (Sacks and Jefferson, 1992)

Using the fragments above, he deconstructed the 'what is going on' from an ordinary understanding of "this is a regular exchange where there seem to be some hearing difficulties on B's side" to the claim that features in contrast to, a sophisticated device, employed by B of avoiding giving one's name in such a way that it isn't notably absent. For example:

- A: this is Mr. Smith may I help you
- B: hello this is Mr. B

The utterance that is left unsaid (illustrated as strikethrough) is perceptively floating in a vacuum. On the other hand, manages to avoid giving the name whilst closing the vacuum or in sacks words, slot, created by the way A furnished their preceding utterance. From there on in, A would have serious difficulty in creating such an easy tit-for-tat nameslot. Straight up asking for the name would have the serious risk of launching a sequence of explicit accounts, started by B asking why A wanted the name something that A's implicit nameslot-creation is protected from. In short, instead of a simple two-line exchange of greeting, a more thorough analysis of the subsemantic reveals a series of strategic and nuanced moves on the battlefield of determining the development of the interaction.

For the purposes of the current investigation, we shall restrict ourselves to the following conversation analytic insights about talk: talk goes beyond linguistic one-purpose ideal-language constructs. It is a toolset for furnishing reflexive order. It is sequentially organised with only one speaker at a time taking a turn to talk. Each turn stands in a relationship to both the preceding and successive turn. The particular situated sequence of turns has fundamental bearings on the way the interactants relate to one another.

The wider theoretical foundations of the study are found in a common conversation analytic synthesis of goffmans observation that interaction "represents an institutional order sui generis in which interactional rights and obligations are linked not only to personal face and identity but also to macro-social institutions" with Garfinkel's stressing that interactants "by their actions [...] exhibit an analysis or an understanding of the event in which they are engaged but by acting they also make an interactional contribution that moves the event itself forward on the basis of that analysis".

An encounter between individuals within a particular structural environment is principally seen as indeterminate before it occurs. When it occurs, all of the interactants contribute to the formation of a consensus.

If interactants want to affect what is going on in a particular moment, they need to have a certain degree of autonomy. Encounter norms are therefore at any given point malleable to interactant influence; norms of interaction can be taken up, declined, ignored and upheld (Goffman, 1967).

This goes against the idea that there are prefabricated interactive norms that can be simply imposed on a particular context, while not denying the influence of certain interactive archetypes, especially in heavily ritualised or institutional contexts: there is such a thing as a "conference study presentation" but there is no guarantee that it will remain one unless a particular selective attention, norm of behaviour and organisation of space is maintained by the participants.

The capability of the conversation partners to actively stabilise the encounter is of paramount importance to the success of an interaction.

Interactant autonomy: For the purposes of this investigation, the autonomy of an interactant may thus be defined as its ability to contribute to the form of interaction, firstly through the particular mode of entering the conversation, secondly through being able to contribute to the flow of a conversation as a competent participant; to enter into a conversation as X have this X sustained and using X to influence the direction, type, topic, mode and purpose of the interaction. Everyday conversation encounters, for instance, in the absence of constraining institution or ritual, allow a maximum of expression of said dynamic (Schegloff, 1992).

Restrictive environments are consequently those where the dynamic is somehow limited, either by pre-existing norms that impose a limit on which interactions are appropriate (e.g., the more "formal" court proceedings, lectures, etc.) or by denying interactant autonomy gradually (e.g., in a situation where one interactant vastly outranks the other and can deny certain directions such as a switch to an overly jovial interaction) or completely (e.g., a total boycott, the denial of interactant status) (Atkinson, 1982).

Formal environments: Formal environments, in that sense, are environments which govern the interaction within them in a restrictive fashion. In order to understand the toolset of a formal environment, the initial core assumptions should be extended by the observation that "first that at least one and not more than one speaker speaks at a time and second that speaker change occurs" (Atkinson, 1982), i.e. that conversations are structured in distinct turns (Sacks and Jefferson, 1992).

Atkinson (1982) argues that managing turn allocation is integral to "achieving and sustaining the shared attentiveness of co-present parties to a single sequence of actions", especially in environments with greater audiences, where turn-taking cannot be easily managed by ad hoc conventions: without turn taking conventions, a mass gathering of people will struggle to maintain shared focus of attention and will be at risk to fall back to more manageable groups. As an example, take a non-assertive teacher who loses control of his or her class; the latter breaks down into sub-units with different conversations, where it is possible to oversee the individual interactants without significant turn-taking rules.

While the breakdown of a spontaneous mass gathering is relatively unproblematic, institutional environments in particular require careful management of turn-taking and mode of communication, for these environments are primarily goal-oriented. For example, an interaction between a bank clerk and a customer is regulated based on the goal of the visit (some form of transaction or other bank business), the institutional roles of clerk/customer and the constraints of action based on the latter. While small-talk may be commonly employed in semi-formal institutional environments, care is taken to prevent a full breakdown of the purpose of the visit: it would be inappropriate for the bank clerk to start a discussion on the merits of post-modern literature with a customer.

Based on the success of maintaining the focus and roles of an institution, an inferential framework is referenced: a doctor that fails to sustain an authoritatively knowledgeable position will risk appearing as medically incompetent; a trial proceeding in a chaotic manner might be seen as an illegitimate farce.

Consequently, formal interactions are in need to be carefully managed. This need arises out of their goal-orientation and is realised by the adherence of the respective interactants to norms of interaction specific to the institution: the "formality" of an institution is thereby not given by their structure but constantly maintained and confirmed by each successive turn. In other words, "formal" and "informal" is necessitated by the norms necessary to sustain a situation. A mediation session could thereby be described as pleasantly informal because it is typically dealing with a smaller audience and therefore does not need pre-formed rules. The conversation can be managed by the co-present interactants. Once the co-presence loses its immediacy, i.e., people are not able to assess the entirety of the group at once and react organically, formalisation takes place.

Institution is also maintained in the structure of the turns themselves. Different settings prescribe answer/question formats to particular institutional actors. Speaking out of turn or changing the restrictive format of the interaction is directly enforced, e.g., in courts: it would be highly uncalled for the defendant to start asking the judge questions instead of the other way around. Who speaks next and for how long is vigilantly monitored.

Having introduced the basic characteristics of informal interactions, formal interactions and interactant autonomy, it is now possible to apply these categories to the ACTI system in question.

Characteristics of the non-human interactant: From the technical parameters described in the introduction, a future encounter can be said to have pre-determined characteristics.

First of all, a telephone interaction is necessarily more formal than a face-to-face encounter. This is because a telephone conversation is initiated without the necessary knowledge of who is picking up and why. This makes it necessary to have some form of introduction to set the desired tone to which the respective other interactant can react. Callers and the called are asymmetrical in the sense that the former knows who they want to call while the called cannot (except in the case of caller ID) determine who called them. No matter the level of authority normally commanded, some form of introduction is needed from the caller while the called does not have to reciprocate accordingly. This observation is symptomatic of the telephone conversation 'genre'. In general, telephone conversations require a higher degree of conventions because of the absence of a whole body co-presence of interactants that would otherwise allow for a more tacit management of the encounter.

A telephone interview itself is another level of formalisation, for it is goal specific: the caller wants information from the called. In the context of the institution of the telephone interview, therefore the caller proposes, through an introduction, a set of turn norms and institutional roles. If accepted, the caller becomes the interviewer and the called becomes the interviewee. This firstly means a particular set of proposed actions (the caller asks questions, the interviewee is expected to respond) and secondly that the failure to adhere to the institutional roles can threaten the interaction.

The non-human interactant has a few peculiarities that go beyond the general restrictions and peculiarities of the telephone interview.

Turn imposition: Due to the technical limitations, the ACTI System is unable to be undogmatic about turns; a recording has to be played in its entirety. After that, silence has to be maintained for a certain amount of

time. This gap between recordings is the proposed turn sequence. A respondent may interrupt a question with a response but even if the program is equipped with the ability to count interruptions as proper answers, this just restarts the turn sequence originally imposed: another recording is played, a turn is proposed with a request and subsequent silence.

Turn type imposition: Likewise, this technology is not as yet so advanced as to be able to interpret answers in their everyday complexity. A set of possible answers can be anticipated with slight variations but in general what is expected is an audible enunciation of a numeric value or a datum from a finite set of pre-determined categories (months, a set of names, countries, etc). In some cases, the acceptable answers are enumerated in the question itself.

This is highly problematic in everyday interaction as a respondent might respond to the question "How many children do you have?" with "A beautiful pair of blonde-locked twins". A human interviewer can extract the numeric value from said statement. What might be even more crucial is that a competent interviewer might be able to somehow respond to the emotional quality of the response and engage in brief smalltalk before moving on. Whether it is desirable that such smalltalk occurs is a different question, however, what cannot be avoided is reacting to the precedent turn as "by producing their next actions, participants show an understanding of a prior action and do so at a multiplicity of levels [...]. These understandings are (tacitly) confirmed or can become the objects of repair at any third turn in an ongoing sequence" (Schegloff, 1992). Even though an interviewer can choose to disregard the emotional addition, the subsequent turn will still be a response to the prior; in this case it will likely be seen as rejecting the addition.

So, if any response is a contextual one and the computer responses are recorded in advance one for each question each turn by the computer is a harsh statement on the preceding turn, in which any proposed deviance from a programmed turn type is met with complete disattendance.

A non-standard response is either rejected (i.e., not actively acknowledged) and another question initiated or a repeat answer is requested. Both options communicate primarily that there is a rigid institutional norm at hand that does not allow for any transgressions.

General features of ACTI as a conversational partner: As a general feature, therefore, the computerised conversation partner is rigid. A conversational norm that is normally sustained and affirmed by the input of its participants is set in stone, nonreactive. The institutional roles are enforced without possible compromise. Should the respondent have an unanticipated (i.e., not programmed) request (e.g., "Could you please speak up!"), the computer is incapable of either reacting to (e.g., raising the volume of the next recording) or acknowledging it (e.g., "Okay, I will speak up"). Conversely, when the respondent is answering, the answer is not necessarily followed by an acknowledgement in the form of "okay". Instead, the next question is played.

As a result, a computer is quite dictatorial as a conversation partner: it chooses the direction of the conversation, the type of acceptable response, the turn length, all the while completely disattending any personal remarks the respondent may utter or even in the case of reiterating the question disciplining the respondent regarding their place in the exchange.

How this behaviour is interpreted depends highly on the person in question as there are many possible explanations for it. It could be a very dispassionate or bored interviewer, just reading a script, it could be interpreted as a form of authority, leading an interrogation; conversely, the interviewer could just be socially incompetent and so on. What is important, however, that each interpretation will reference an inferential framework that will have to be consistently sustained: if the dictatorial imposition of the form of engagement is based on a supposed authority, said authority has to be maintained and made sense of with each turn.

However, if the reiteration of a question can be construed as being the voice of displeasure with the preceding answer, it can be made coherent with the status of an authoritative agent. This way, robot peculiarities (i.e., re-asking a question repeatedly until the expected answer occurs) can be made sense of as coherent features of the institutional persona. Thus, an authoritative interpretation is arguably more reasonable to assume for a respondent as it is one of the few options where ACTI's actions are legitimate. As it stands, Occam's razor cuts in favour of 'interrogator' over 'non-human' which may further contribute to the prevalence of the 'authoritative interviewer' interpretation.

RESULTS AND DISCUSSION

Illustration: The following is a series of selective transcripts from the dataset. The original phone recording was Russian; transcripts are provided in the form of a translation made by the present author. It was attempted to more or less mimic the overall sentence structure of the

original. Brackets are used to delineate moments where talk is coinciding. As one person is talking to a computer, a classic interruption does not occur: the sentence "interrupted" is allowed to run its course with one negligible exception (Algorithm 1 and 2).

Note that this data is merely illustrative of a theoretical concept, not an attempt to ultimately interpret with absolute certainty what was going on. This study is therefore meant to suggest a future course of research, where such transcripts could be looked at en masse, follow-up interviews be conducted and paradata collected.

Legend:

- IE = Interviewee
- IR = Interviewer
- [= Interruptions/Coinciding speech

Algorithm 1:

IE: Hello? Hello...
IR: This is the All-Russian Center for the Study of Public Opinion. Good evening.
We are conduct[ing a public opinion survey of residents of the Moscow Oblast about...
IE: [Good evening!
IR: ...the socio-political situation in the region.
would like to emphasise that there are no correct o[r incorrect ways to answer them.
IE: [Could you speak up a bit!
IR: It is impor[tant to us to know your personal opinion about the situation of our Oblast.
IE: [Hello...?
IR: You can be absolutely certain in the confidentiality of your responses they will only be used in a summarised fashion
Did you participate in public opinion polls during the last three months?
IE: ... no
IR: Are you a citizen of the Russian Federation? [Please answer yes or no.
IE: [yes, yes, yes. Born, raised, grew old.
IR: Please ans[wer yes or no.
IE: [Hello...?
Yes..yes
IR: Please answer yes or no.
IE: Yes! Yes!

Algorithm 2:

IR Do you think that you will be able to come on that day and participate in the elections? Possible [answers:
IE [Of cour...
IR It's: likely, somewhat unlikely, don't know.
IE Oh...ummmm... of course I'll come... will go... vote. (laughter)
IR Excuse me, I did not understand. Will you personally be able to take part in the elections on the 8th. Possible answers: It's likely, somewhat unlikely, don't know.
IE It's likely!

In this short 'conversation', lines 7-9 and 15-17 illustrate an attempt by the IE to assert themselves as conversation partners. In both cases as per the technical limitations, the request in 7 and the back-story in line 15 is entirely ignored. In both cases, the IE reacts with an inquisitive/puzzled "Hello...?". There is a marked contrast in tone between the cheerful line 15 and the relatively muted line 18.

In the entire sequence, the IR speaks the entire question without hitches, pauses or interruptions, despite attempts of IE to prematurely answer or voice requests. It is not immediately clear how to interpret the utterances in lines 9 and 17, respectively. A possible interpretation is that it's a "is anybody there?" utterance akin to line 1. In relation to the core assumption that the person does not seem to realise that they are talking to a computer, it could be an attempt to confirm that they have been, in fact, utterly ignored, either by will (disattendance) or by technical difficulties (phone interference). The more favourable explanation of phone interference is, however, immediately disputed by line 19 which restates the rigid response.

This is a transcript from the beginning of the second third of the conversation. An interruption is still occurring in line 2, the answer (likely "of course!"). The IE is audibly struggling to find a formulation that will be accepted, fails and is reminded of the proper format of answering. The laughter could, again, be interpreted in many different ways but it seems likely that it is in this case a form of dealing with being embarrassed about the self-perceived incoherence of the precedent sentence. In parallel to the first transcript, the "correct" answer is finally uttered with a high degree of assertion.

Again, the IE attempts to prematurely answer but stops talking, letting the IR resume with the enumeration of options. Consequently, the IE unsuccessfully tries to reproduce one of the standard answers (line 4). In parallel to Algorithm 3, this is followed by laughter. This transcript documents an exchange that is well into the "conversation". It appears that in this case the IE did indeed not manage to determine that this person was a computer.

Algorithm 3:

IR How certain are you that you will vote exactly like this during the mayoral election? Possible ans[wers:
IE [One hundred...
IR I'm quite sure that it will be so, if nothing extraordinary occurs; not quite sure of my choice, everything can change; absolutely unsure, my vote will likely change.
IE Quite...sur....likeeel....hey...[what was it again? Geeee...fo[r him. I'll vote for him is what I mean. (laughter)
IR [Exc... [ip...

What can be gained from these examples? Firstly, that it is not completely out of the question that some people do not realise that they are talking to computers. The entirety of this particular recording is filled with requests and informal talk-things that one might extend to a person but likely not a computer.

It would be principally relevant to look at the conversation conventions in a case where there is little doubt about the robotic nature of the IR, for example, through a clearly robotic voice or through a clearly audible announcement that the IEs are not dealing with a real person. It could very well be that anything supposed to be directed at humans is also present in obviously robotic contexts.

It is reasonable to assume that there are groups that are more or less susceptible to the working assumption. This way, one could hypothesize that some groups will approach the interview with varying degrees of formality. A person realising that they are talking to a computer might relax and even have a parallel conversation while a person thinking him or herself in a highly rigid and authoritative context might be more occupied but it does not principally invalidate the data thereby obtained. It is not at all clear how exactly the uncompromising interviewer appears to the individual interviewees as there are several options: as rude, incompetent, authoritative, professional or dispassionate. It would be beneficial to compare how the various options relate to accepting the continuation of the telephone interview: some people might agree to bear the formality of a criminal investigator whilst not the authority of an unfamiliar religious organisation.

So, while there is definite technically necessitated rigidity in the status quo of computerised interviews, it is ambiguous how they appear to the group of people who were tricked by the human voice.

Problematical, however is the unintended generation of a highly authoritative institution when it comes to revealing voting patterns: thinking that there is a wrong or right form of answer might lead the respondent to assume that there is a right or wrong answer in general. In that context, it would be relevant to single out those respondents that were likely to be more familiar with authoritative contexts whilst also being less familiar with cutting-edge telephone interview technologies such as the older generation in Russia. If that group does indeed exist and is prevalent (e.g., due to the disappearance of landlines, the times called, etc.), it would have to be investigated whether their responses are significantly affected by the impression of there being a correct answer.

CONCLUSION

In summary, this study aimed to point out that technical characteristics of interviews will become increasingly relevant to the interactive context thereby forced upon the respondents. With the development of intelligent agents, it may become harder to distinguish humans from computers. At the same time, it will likely take a much longer time for machines to be able to mimic human interactive capabilities. The result may contribute to the emergence of human-like automatons that constantly force a particular mode of conversation upon the human interactants, thereby diminishing the regular basic expressive autonomy an interactant possesses. As such, the case of this surveying system may be a cautionary tale with increasing relevance to HCI practitioners as the increasing pervasiveness of interactive systems is likely to also perpetuate the side-effects of them being not interactive enough. It may not be outright undesirable that a surveybot behaves like a 'heartless interrogator', yet having the knowledge that this is occurring is crucial for having control over survey proceedings and consumer interactions. This knowledge can only be gained with an increased attention to the building-blocks of talk which go considerably deeper than the sentence-level.

REFERENCES

Atkinson, J.M., 1982. Understanding formality: The categorization and production offormal interaction. Br. J. Soc., 33: 86-117.

Blackstone, M.M., D.J. Wiebe, C.J. Mollen, A. Kalra and J.A. Fein, 2009. Feasibility of an interactive voice response tool for adolescent assault victims. Acad. Emergency Med., 16: 956-962.

Brahnam, S. and D.A. Angeli, 2012. Gender affordances of conversational agents. Interact. Comput., 24: 139-153.

Corkrey, R. and L. Parkinson, 2002. A comparison of four computer-based telephone interviewing methods: Getting answers to sensitive questions. Behav. Res. Methods Instrum. Comput., 34: 354-363.

Eyssel, F. and F. Hegel, 2012. She's got the look: Gender stereotyping of robots1. J. Appl. Soc. Psychol., 42: 2213-2230.

Garfinkel, H., 1967. Studies in Ethnomethodology. Prentice Hall, Englewood Cliffs, New Jersey,.

Goffman, E., 1967. Interaction Ritual: Essays in Face-to-Face Behavior. Aldine Publishing Company, Chicago, Illinois,.

Kim, Y. and S.S. Sundar, 2012. Anthropomorphism of computers: Is it mindful or mindless?. Comput. Hum. Behav., 28: 241-250.

Kreuter, F., S. Presser and R. Tourangeau, 2008. Social desirability bias in CATI, IVR and Web surveys the effects of mode and question sensitivity. Public Opin. Q., 72: 847-865.

Liang, Y.J., S.A. Lee and J.W. Jang, 2013. Mindlessness and gaining compliance in computer-human interaction. Comput. Hum. Behav., 29: 1572-1579.

Meuse, K.P., 1987. A review of the effects of non-verbal cues on the performance appraisal process. J. Occup. Psychol., 60: 207-226.

Nass, C. and Y. Moon, 2000. Machines and mindlessness: social responses to computers. J. Soc. Issues, 56: 81-103.

Nass, C., J. Steuer and E.R. Tauber, 1994. Computers are social actors. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems CHI-94, April 24-28, 1994, ACM, Boston, Massachusetts isBN:0-89791-650-6, pp: 72-78.

Nass, C., Y. Moon and N. Green, 1997. Are machines gender neutral? Gender-stereotypic responses to computers with voices. J. Appl. Soc. Psychol., 27: 864-876.

Nass, C.I. and S. Brave, 2005. Wired for Speech: How Voice Activates and Advances the Human-Computer Relationship. MIT Press, Cambridge, Massachusetts isBN:9780262140928, Pages: 296.

Pickard, M.D., C.A. Roster and Y. Chen, 2016. Revealing sensitive information in personal interviews: Is self-disclosure easier with humans or avatars and under what conditions?. Comput. Hum. Behav., 65: 23-30.

Reeves, B. and C. Nass, 1996. The Media Equation: How People Treat Computers, Television and New Media Like Real People and Places. Cambridge, Cambridge University Press, UK.

Sacks, H. and G. Jefferson, 1992. Lectures on Conversation. Wiley-Blackwell, Hoboken, New Jersey isBN:9781557863584, Pages:

Schegloff, E.A., 1992. Repair after next turn: The last structurally provided defense of intersubjectivity in conversation. Am. J. Sociology, 97: 1295-1345.

Suchman, L.A., 1987. Plans and Situated Actions: The Problem of Human-Machine Communication. Cambridge University Press, Cambridge, England,ISBN:0-521-33739-9, Pages: 207.

Sundar, S.S. and C. Nass, 2000. Source orientation in human-computer interaction Programmer networker or independent social actor?. Com. Res., 27: 683-703.

Tay, B., Y. Jung and T. Park, 2014. When stereotypes meet robots: The double-edge sword of robot gender and personality in human-robot interaction. Comput. Hum. Behav., 38: 75-84.

Tay, B.T., S.C. Low, K.H. Ko and T. Park, 2016. Types of humor that robots can play. Comput. Hum. Behav., 60: 19-28.

Zhu, B. and D. Kaber, 2012. Effects of etiquette strategy on human-robot interaction in a simulated medicine delivery task. Intell. Serv. Rob., 5: 199-210.