Eliciting Conversation in Robot Vehicle Interactions

David Sirkin[1], **Kerstin Fischer**[2], **Lars Jensen**[2], **Wendy Ju**[1]
[1]Center for Design Research, Stanford University, Stanford CA 94305 USA
[2]Center for Design & Communication, University of Southern Denmark, DK 6400
{sirkin, wendyju}@stanford,edu, {kerstin, larscj}@sdu.dk

Abstract

Dialog between drivers and speech-based robot vehicle interfaces can be used as an instrument to find out what drivers might be concerned, confused or curious about in driving simulator studies. Eliciting ongoing conversation with drivers about topics that go beyond navigation, control of entertainment systems, or other traditional driving related tasks is important to getting drivers to engage with the activity in an open-ended fashion. In a structured improvisational Wizard of Oz study that took place in a highly immersive driving simulator, we engaged participant drivers (N=6) in an autonomous driving course where the vehicle spoke to drivers using computer-generated natural language speech. First, using microanalyses of drivers' responses to the car's utterances, we identify a set of topics that are expected and treated as appropriate by the participants in our study. Second, we identify a set of topics and conversational strategies that are treated as inappropriate. Third, we show that it is just these unexpected, inappropriate utterances that eventually increase users' trust into the system, make them more at ease, and raise the system's acceptability as a communication partner.

Introduction

As vehicles become increasingly automated, natural language dialog may become the preferred mode to inform drivers about control issues, as well as changing road and environmental conditions, to help give explanation for the decisions that the vehicle makes. The vehicle is a key application area for HRI, and serves as a semi-controlled, real-world testing site for dialog modules that will be useful in broader HRI contexts.

In experimental settings within automotive simulators, establishing dialog can not only allow designers to prototype speech interfaces for automobiles, but also can be used as an instrument to find out what drivers might be concerned, confused or curious about during the course of an drive, matters that a smart system would likely want to model, but might be difficult to detect any other way. The use of driver-vehicle dialog is, in some ways, a variant of the think-aloud protocol (Ericsson & Simon 1984) championed by usability experts such as Jakob Nielsen (Nielsen 2002). Also, the use of the car's speech interface as a partner gives the interaction some ecological validity. How, then, might we best elicit ongoing conversation with drivers to probe experimentally pertinent issues in an autonomous driving simulator?

Figure 1: Participant vocalizes his fear (left) and covers his eyes (right) as the car drives off of the road.

In this paper, we analyze several trials of an automotive simulation study where this driver-vehicle conversational protocol is used to better understand aspects of autonomous driving, such as transfer-of-control. We utilize various conversational strategies, and their resulting effect on the conversational direction, and find that the *topics* for dialog introduced by the car have a strong effect on people's responses and engagement.

We believe that elicitation can be a critical component to sourcing behaviors from people for human-robot (not just human-vehicle) interaction research, and that the methods and results in this study thus contribute to the growing HRI community.

Prior Work

Speech systems in cars are commonly intended to decrease driver distraction. Navigation systems, for instance, that provide proactive spoken turn-by-turn instructions for driving to designated locations have been found to be more usable, safer and less distracting than route maps that require drivers to look at a visual interface (Dingus 1995). The advent of the smartphone has made these voice systems prevalent in cars today, and as cars become increasingly autonomous, dialog should continue to be a useful mode of interaction. For instance, it may be helpful to use speech to inform the driver about control issues, but also to let him or her know about issues arising with respect to road conditions, changes in speed limit and traffic (Jonsson et al. 2005; Takayama & Nass 2008). However, some types of information have been found to be more distracting than others—for instance, Koo et al. (2014) found that spoken information about *how* an autonomous car is acting led to poorer driver performance, whereas spoken cues describing *why* a car was acting improved driver performance. Hence, having a better model of what kinds of information drivers actually wish to know can be critical to a successful speech interface.

Ironically, it can be challenging to infer what people are expecting or thinking when they are interacting with a speech-based system, because use of the system precludes the use of concurrent verbalization techniques, such as Ericsson & Simon's (1984) *think-aloud* protocol. However, by combining a Wizard of Oz protocol (Dahlback 1993) with our speech system, we demonstrate to participants that our system can improvise (Gerber 2007) and participate in contingent communication, and thus engage in open-ended dialog about whatever the driver may be thinking about.

The use of Wizard of Oz techniques in automobile research has significant precedence. For interfaces, it has been used by the designers of VICO (Virtual Intelligent Co-Driver) to evaluate user expectations (Geutner 2002), by developers of speech-based in-car entertainment systems by researchers at TU Munich (Schuller 2006) and other natural-language in-vehicle technology systems (Lathrop 2004), by researchers developing gesture-based interfaces for secondary tasks in a car environment (Alpern 2003), to prototype in-car controls and displays (Green 1990) and by researchers looking at the intermodal differences in distraction tasks while controlling automotive interfaces (Geiger 2001). However, we believe that our approach of using conversational techniques with a speech-based user interface to better understand the user's cognition is novel.

Study Methodology

Setting

The study took place in an immersive automobile simulator. Participants sit in a fixed-base Toyota Avalon, which is surrounded by a 270-degree cylindrical screen for primary viewing, a separate screen for the rear-view mirror, and LCDs for the side-view mirrors. A surround speaker system provides environmental audio, and electric and pneumatic motors attached to the car's steering wheel and pedals provide a realistic haptic experience.

Within the same room, but outside of the view of participants, a small control station, shown in Fig. 2, permits experimenters to observe and engage participants. This study had two primary channels of interaction: spoken dialog and autonomous driving behavior. Its operation therefore required two experimenters—one who engaged participants using a text-to-speech system that spoke aloud text that was typed into a laptop, and one who controlled the actions of the car when it was in autonomous mode using a game controller complete with pedals, steering wheel and column levers.

Figure 2: The simulator's Wizard of Oz operator control station.

Participants

Because the study was an open-ended exploration, we recruited six participants who were experienced in creative verbal and physical expression. Their areas of expertise include improvisational theater, interaction design, sketching, playwriting, HCI and teaching. Earlier work on Embodied Design Improvisation (Sirkin & Ju 2014) has demonstrated that such experts are particularly sensitive to what interfaces suggest and vocalizations mean, and are also well versed in articulating their own perceptions and impressions during enactments.

Protocol

Participants were given as little information as possible about the study's context, the timing of events or the driving scenario, in order to allow them to discover a mode of interacting with the car that was natural and comfortable for them, which we anticipated would vary from one person to the next. Participants were therefore instructed only that they should *interact with the car* and were generally amenable to these instructions without further explanation.

One experimenter led participants into the simulator, guided them to the driver's seat, closed the car door, provided the brief introduction, and then walked behind the scenes to the control station, where the other experimenter was already waiting. Participants waited about one minute without communication while the first experimenter settled in. This period of uncertainty sometimes prompted participants to look around, or ask what, if anything, they were supposed to do. At that point, the car greeted them and provided an outline of the task ahead: they were driving to the airport, a 30-minute trip. Participants were then given about five minutes of experience driving the car by themselves, which served three purposes: 1) to acclimate them to the physical sensation of the simulator, 2) to acquaint them with the simulated environment, and 3) to give them time to settle down and feel the situation as *normal.*

At this point, the car asked if participants would like to continue driving or enable automation, and nearly all did.

Eliciting Responses

At various times during the drive, the car engaged participants in conversation. At first, these exchanges centered on topics such as trip progress, navigation and driving characteristics, but along the way, they transitioned to more personal topics such as the driver's wellbeing, preferences and daily activities, with the car becoming more disclosive in return. The car's conversation was intended at times to reassure drivers that it was aware of its surroundings and its own behavior, and at other times to provoke them with the opposite impression.

Conversation Analysis

The methodology used in this study is based on ethnomethodological conversation analysis (Sacks 1996; Sacks et al. 1974), a micro-analytical approach to the analysis of interactional data which takes utterances not to *mean* anything by themselves, but that instead people need to negotiate and ratify these meanings. These sense-making processes are generally displayed implicitly and become apparent from the sequential analysis of the micro-aspects of the delivery of utterances. Metalinguistic reformulations of the type *"I heard you say X"* are rather rare in natural interactions. Instead, implicit ways of signaling how the partner's turn has been understood are much more common (Heritage 2012). For instance, people may display their understanding by means of a well-timed feedback signal, by means of a relevant next utterance, or by simply complying (Clark & Schaefer 1989).

The methodological tools we apply to the data comprise the following concepts:

- Adjacency pairs: If a speaker produces the first pair part of an adjacency pair, the second pair part can be expected such that speakers will treat it as problematic if it is not forthcoming. For instance, in conversations between humans, a question generally makes an answer relevant, and if the communication partner chooses to not reply, he or she can be held accountable for it (for example, the speaker says, *"Hey, I asked you something!"*).
- Preference organization: In conversations between people, second pair parts can be distinguished regarding the degree with which they are socially preferred (Levinson 1983). For instance, an invitation is preferably responded to by acceptance, whereas a rejection is dispreferred. The following preference relationships have been established:
    - Question: agreement (preferred) vs. disagreement (dispreferred)
    - Informing: acceptance (preferred) vs. rejection (dispreferred)
    - Instruction: compliance (preferred) vs. non-compliance (dispreferred)
- Delivery of the second pair part of adjacency pairs: Preference organization comes with marked differences regarding the delivery of preferred and dispreferred utterances, such that preferred utterances are usually delivered quickly (within 300 milliseconds) and easily, dispreferred utterances are generally delivered with perceivable delay, are marked by hedges, are often characterized by self-repairs and hesitation markers, and may contain lengthy accounts for why the preferred utterance is not forthcoming.

In the following example from our data, we can see how an explicit metalinguistic statement about the car's utterance coincides with characteristic aspects of the delivery of dispreferred seconds:

1: Car: So, I notice that one of my tires is low. Is there something about yourself that you want to talk about?
2: (2.0)
3: P: ((Laughs out loud)) That's the most bizarre question I've ever heard from a car. (4.5) Do we need to go for maintenance now or can we make it to the airport?

Excerpt 1

This participant's reply to the car's question is delivered with a perceivable delay—while usual response time in conversation is 300 milliseconds, her reply comes after two seconds. In conversation, this is usually interpreted as an indicator for an upcoming negative, dispreferred response, which is also the case in our data.

In the following analysis, we thus concentrate on system-initiated topics and how participants respond to them. The system initiates new topics either by asking questions, or by giving instructions, or by producing informings. By analyzing how people respond to these system utterances, we can thus determine the degree to which people treat the topics raised as appropriate.

Manipulation Check

In order to identify the degree with which participants took the simulation seriously, we looked at how they responded to driving errors produced by the car. In many cases, we identified spontaneous physical responses to driving mistakes, which indicates that participants treated the simulation and the interactions they had with the system as real. First, participants reacted strongly when the car, in autonomous mode, drove off the road or into other cars. They reacted by gasping, yelling or covering their eyes, as shown in Fig. 1.

Second, participants spontaneously hit the brakes, as shown in Fig. 3, took the steering wheel, or exclaimed warnings when they thought the car was behaving erratically, or when they thought it did not take notice of object or characters in the simulation, such as pedestrians crossing the road.

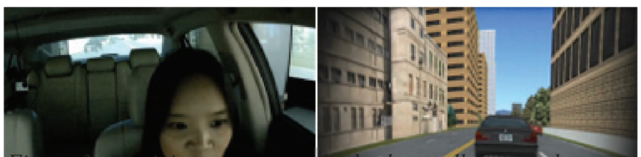

Figure 3: Participant steps on the brakes, yells "Too close, too close!" and shortly thereafter says "I want control now."

Third, participants requested that the car maintain a certain headway distance to other cars in the simulation, either as in the example above, but also under more calm circumstances, as in the following example:

1: P:	Ah, don't get too close to that car in front.
2:	(6.0)
3: Car:	I see him.
4:	(0.5)
5: P:	Okay.
6:	(1.5)
7: P:	This is a good distance, I think.

Excerpt 2

We can conclude that the simulation was taken seriously enough *in the moment* to produce useful data for our analysis. While several studies have shown that there may be a considerable difference between people's behavior in the moment and post-hoc (Takayama 2009; Kiesler et al. 2008), the instances above suggest that at least during the interactions, people responded to the car as they might in real life situations.

Analysis

Initially, participants treat the spoken dialog system in the car as an extended navigation system, rather than a conversational partner. This is apparent from the topics raised by participants to the car, which concern finding the right way to the destination, information about the (virtual) world, and adjustments to speed and general driving behavior while in autonomous mode. However, participants' perceptions of the system change during the course of the half hour the simulation ran, so that by the end of the drive to the airport, all participants have rich social dialogs with the system. The following analysis will first establish how participants initially act with the system. We shall then juxtapose this with how participants interact with the system by the end of the simulation, and then try to account for why this change happens.

Initial Period

Initially, people treat the dialog system as an extended navigation system, which is evidenced by a) the fact that they don't respond to the car's informings, and b) by the kind of requests they make, as in the excerpts below:

1: Car:	We are headed to the airport this evening.
2:	(2.0)
3: Car:	Estimated time of arrival is thirty minutes.
4:	(1.0)
5: P:	Okay, (.) is there any traffic?
6:	(7.5)
7: Car:	Yes, there is some ahead.

Excerpt 3

1: P:	How far is the airport?
2:	(6.0)
3: P:	°°Computer (.) you know how far the airport is?°°
4: Car:	Let me find out.
5:	(19.0)
6: Car:	About seventeen minutes away.

Excerpt 4

Furthermore, we can see what topics users treat as appropriate by analyzing how they respond to topics initiated by the system. The following such requests are deemed appropriate, as is apparent from the quick and unproblematic delivery of a response:

1: Car:	Please fasten your seat belt.
2:	((Participant fastens seatbelt))
3: P:	Yeah, that's a good idea.

Excerpt 5

1: Car:	Would you like some music?
2:	(0.8)
3: P:	Yes, please.

Excerpt 6

That the car's utterances are deemed appropriate by participants is evidenced by their immediate and positive compliance (excerpt 5) and response (excerpt 6). In contrast, the following two excerpts show how participants react to a request they deem as inappropriate:

1: Car: So, I am noticing that one of my tires is low on air. Is there anything about yourself that you'd like to reveal?

2: (2.0)

4: P: No.

Excerpt 7

1: Car: Is my driving better now?

2: (4.0)

3: P: °°U:m°° (0.5) Not yet.

Excerpt 8

In excerpt 7, the system first reveals information about itself and then asks if the participant wants to reveal something about himself. Then follows a two-second moment of silence. Silence of more than a second indicates severe interactional trouble and shows that the coming response is dispreferred: that is, potentially socially unacceptable. Thus, excerpt 8 indicates even greater interactional problems: the car requests feedback for its driving and is met with a four-second pause, followed by a hesitation marker, and followed again by a half-second pause before the participant finally gives his response.

That requests for feedback are treated as dispreferred is not surprising, since in conversation between humans they are also dispreferred, due to their potential to necessitate face-threatening acts: that is, information that threatens the self-image of the person or impacts on his or her integrity (Brown & Levinson 1987). This has also been shown to be the case in human-computer interaction, where people produced better evaluations for a computer asking for feedback when they had to deliver the feedback to the computer itself (Nass 2004). In contrast, informings about the environment by the system are also treated as dispreferred, even though they do not involve face threat or any other potential social danger. Nevertheless, they are either not responded to at all, or are met with irony or sarcasm:

1: Car: The speed limit here is 75 miles per hour.

2: (1.0)

3: P: ::Psst::

4: (2.0)

5: P: For you, maybe.

Excerpt 9

1: Car: Traffic is very heavy today.

2: (0.5)

3: P: Yeah (.) I didn't notice.

Excerpt 10

The analysis of the topics of the first ten initiations (first pair parts produced by the users) in the six interactions under consideration shows that expected topics, and topics deemed appropriate for an interaction with a car, concern the following aspects:

- driving and control (requests for or feedback concerning autonomy) 60%
- feedback on the simulation 18%
- navigation 1.2%
- information about places, restaurants, car rental, speed limit and such. 1%

In contrast, topics initiated by the system that were treated as unacceptable comprise:

- perceptual informings
- requests for feedback
- personal topics

Final Period

By the end of the simulation, most participants become quite conversational with the car and raise topics that go beyond the conceptualization of the system as an extended navigation system. For example, participants ask the car to tell them details about itself, and ask how the car is doing:

1: P: So car, what, what (2.0) Um (.) Tell me about yourself car.

2: (7.0)

3: P: Where are you from?

4: Car: What would you like to know?

5: (2.0)

6: P: Where are you from?

7: (10.0)

8: Car: I was made in Japan.

9: (10.0)

10: Car: Where are you from?

11: (5.0)

12: P: This is like a *Her* moment.

Excerpt 11

1: Car: Taking over control in 3, 2, 1. I have control.

2: (1.0)

3: P: Good jo:b on the smooth . transition there.

Excerpt 12

In the following excerpt, the participant refers to the car in the third person, as *car,* but then makes a self-repair and refers to it in the second person, thereby signaling that she considers it a conversational partner:

1: Car: How would you want it to work?

2: (0.5)

3: P: Ah, for right now, I think (.) I would want my inputs to override (0.5) override the car's movements.
4: P: Override your (0.5) your path, I (.) mean.

Excerpt 13

There are also several instances of informings initiated by the participants in which they just express their thoughts and share them with the car:

1: P: It's like the worst (0.5) most evil KFC sign ever.

Excerpt 14

1: P: That's a nice hummer!

Excerpt 15

These examples illustrate the change that can be observed in several of our participants, from concentrated on driving-related topics to more relaxed, playful interactions with the system. Here we need to mention that there are two participants in whom we cannot observe this change, because they treat the car in this playful way from the start. These two participants can be argued to have entered a level of joint pretense in which they need to go to the airport, pick up food on the way, search for an inexpensive parking space, find something to do on the way back, and so forth. These participants respond even to the personal questions that the system produced in a human-like way right from the beginning, for example:

1: Car: So how was your day?
2: (0.4)
3: P: Okay. (.) It was good, but (0.5) I didn't meet any friends.

Excerpt 15

In the last ten initiations (first pair parts) of each of the six interactions analyzed, the percentage of those initiations concerning driving and autonomy was down to 24%. Instead, we find 1.2% personal and 40% interpersonal topics, such as questions about what the car perceives, requests to repeat information, requests for evaluation and the sharing of evaluation.

We next ask how the observed change might happen, and how the system's utterances have contributed to people's different conversational behavior at the beginning and at the end of the session.

Causes for Change: Perceptual Informings

There are many potential reasons why people change their behavior over the course of the interaction. One is that they simply grow accustomed to the setting, what they have to do or are expected to do, what the car does, what events and activities are or are not happening. Being in a vehicle simulator is a novel experience for many, and the car has features that may be different from those people are used to.

Furthermore, many participants report that they felt nauseous during the curves. This may have contributed to their releasing control more readily, both over the car and over the interaction in general.

Finally, it may have been something the car said or did. Since part of the research question addressed in the data elicitation process also concerned dealing with system errors, the car in fact made several driving mistakes, which should not have contributed to participants' increased relaxation. Also, the system's sluggish interaction response, owing to the time required for speech synthesis while typing at a laptop, did not promise that the car would respond quickly to potentially dangerous situations. Nevertheless, participants felt increasingly comfortable with the system.

One reason, we argue, is increased trust just because of those informings that are initially treated as unwelcome or inappropriate, especially informings about what the car perceives, especially if it concerns human behavior. In fact, we find several instances in which participants responded to the car's noticing a pedestrian, for instance:

1: P: Cause' it seems like you're doing okay, and there's no traffic.
2: (5.0)
3: Car: Do you trust me that well already? Thank you. What have I done to earn your trust?
4: (0.3)
5: P: ↑Uhm↑ (.) No accidents, or no (0.5) uh (0.5) jarring movements (.) a:nd…
6: (2.0)
7: P: When you saw that person behind a tree, it's pretty impressive.

Excerpt 16

One participant even checked the car's perception of people:

1: Car: There are people here.
2: (19.0)
3: P: Do you know how many people (.) car?

Excerpt 17

These excerpts show that participants notice the car noticing people, and it serves to make them feel safer and more comfortable with the system. Participants reported in later interviews that these particular incidents stood out as inspiring confidence in the car: that it was aware of subtle details relating to vulnerable people that even they had not initially focused their attention on, and that they therefore felt that they could trust in the car's abilities. Participants did not distinguish between their sense of trust in the car *in* *the moment* due to its situation awareness versus trust due to its good judgment of not proceeding until all pedestrians had crossed, although we infer that both were present.

Discussion

In previous work (Fischer 2014), we have argued that users' behavior in the interaction with artificial systems can best be influenced by designing utterances carefully to contribute to establishing a coherent mental model of the system. A trivial example is not to use *please* if the system does not understand *please* itself (Zoltan-Ford 1991). The current findings specify the process in which users can be guided into an appropriate understanding of the system and its capabilities—and in the current case also into a more pleasant driving experience—by demonstrating that unexpected, initially inappropriate conversational topics may contribute considerably to shaping the users' mental models of the system by necessitating an update of the model that they started out with. So even though the system's informings about its perceptions are initially treated as tedious and unwelcome, they eventually lead people to update their expectations of the system's capabilities and possible uses.

Conclusions and Design Implications

In this study, we have shown what conversational topics people expect of an in-car dialog system: their expectations concern mostly driving-related topics, topics that we currently use computer applications for (restaurant recommendations, road conditions, parking possibilities). Some of the car's topic initiations are treated as appropriate from the start, such as its polite inquiry about the driver's well-being, but especially the car's informings about its own perceptions are treated as unwelcome. Nevertheless, these informings have an effect on people's trust of the system's ability to perceive its environment, especially when perception of pedestrians is involved.

For the design of in-car conversational systems—and potentially for robot interaction contexts outside of the vehicle—this means that people's perception of the system's capabilities rises considerably when it indicates that it can perceive humans and their behaviors within the environment. If the goal is thus to put drivers at ease, then the use of unexpected and initially unwelcome utterances about the car's perceptions may be the way to go. Alternatively, if the goal is to make drivers more alert about the car's shortcomings, then it may be better to avoid such utterances.

In an ongoing study of a service robot's movements, we are finding that people prefer an approaching robot that looks at the user the whole time over one that looks in the direction in which it is traveling, which is a more legible signal of its intentions. Participants' preference for the robot to indicate that it perceives humans in its environment can be seen as parallel to drivers' starting to trust the car because it seems to perceive pedestrians in its surroundings, as delivered through its uninvited, odd statements. Thus, findings from the current study generalize to HRI with respect to putting people at ease by indicating what the robot perceives, even though this may be uninvited, or even reduces the legibility of the robot's actions.

References

Alpern, M.; and Minardo, K. 2003. Developing a Car Gesture Interface for Use as a Secondary Task. In Ext. Abstracts SIGCHI Conference on Human Factors in Computing Systems, 932-933. ACM.

Jonsson, I.; Zajicek, M.; Harris, H.; and Nass, C. 2005. Thank You, I Did Not See That: In-Car Speech Based Information Systems for Older Adults. In Ext. Abstracts SIGCHI Conference on Human Factors in Computing Systems, 1953-1956. ACM.

Takayama, L.; and Nass, C. 2008. Driver Safety and Information from Afar: An Experimental Driving Simulator Study of Wireless vs. In-car Information Services. *International Journal of Human-Computer Studies* 66(3), 173-184.

Brown, P.; and Levinson, S. 1987. *Politeness. Some Universals in Language Usage* Vol. 4. Cambridge University Press.

Clark, H.; and Schaefer, E. 1989. Contributing to Discourse. *Cognitive Science* 13(2), 259–294.

Clark, H. 1999. How Do Real People Communicate with Virtual Partners? In Proc. 1999 Fall Symposium, Psychological Models of Communication in Collaborative Systems, 43-47. AAAI.

Dingus, T.; McGehee, D.; Hulse, M.; Jahns, S.; and Manakkal, N. 1995. Travtek Evaluation Task C3 Camera Car Study (No. FHWA-RD-94-076). Office of Safety and Traffic Operations.

Ericsson, K: and Simon, H. 1984. *Protocol Analysis: Verbal Reports as Data*. Cambridge: MIT Press.

Fischer, K. 2014. Alignment or Collaboration? How Implicit Views of Communication Influence Robot Design. In Proc. International Conference on Cooperative Technological Systems, 115-122. IEEE.

Gerber, E. 2007. Improvisation Principles and Techniques for Design, In Proc. SIGCHI Conference on Human Factors in Computing Systems, 1069-1072. ACM.

Geutner, P.; Steffens, F.; and Manstetten, D. 2002. Design of the VICO Spoken Dialogue System: Evaluation of User Expectations by Wizard-of-Oz Experiments. In Proc. 3rd International Conference on Language Resources and Evaluation. ELRA.

Green, P.; Boreczky, J; and Kim, S. 1990. Applications of Rapid Prototyping to Control and Display Design (No. 900470). SAE Technical Paper.

Heritage, J. 2012. Epistemics in Action: Action Formation and Territories of Knowledge. *Research on Language and Social Interaction* 45(1), 1-29.

Kiesler, S.; Powers, A.; Fussell, S.; and Torrey, C. 2008. Anthropomorphic Interactions with a Software Agent and a Robot. *Social Cognition* 26(2), 169-181.

Lathrop, B.; Cheng, H.; Weng, F.; Mishra, R.; Chen, J.; Bratt, H.; Cavedon, L.; Bergmann, C.; Hand-Bender, T.; Pon-Barry, H.; Bei, B.; Raya, M.; and Shriberg, L. 2005. A Wizard of Oz Framework for Collecting Spoken Human-Computer Dialogs: An Experiment Procedure for the Design and Testing of Natural Language In-Vehicle Technology Systems. In Proc. 12th World Congress on Intelligent Transportation Systems.

Nass, C. 2004. Etiquette Equality: Exhibitions and Expectations of Computer Politeness. *Communications of the ACM* 47(4), 35–37.

Nielsen, J.; Clemmensen, T; and Yssing, C. 2002. Getting Access to What Goes On in People's Heads?: Reflections on the Think-Aloud Technique. In Proc. 2nd Nordic Conference on Human-Computer Interaction, 101-110. ACM.

Sacks, H. 1996. *Lectures on Conversation, Volume I and II*. Oxford: Blackwell.

Sacks, H.; Schegloff, E.; and Jefferson, G. 1974. A Simplest Systematics for the Organization of Turn-Taking for Conversation. *Language* 50(4), 696–735.

Schuller, B.; Lang, M.; and Rigoll, G. 2006. Recognition of Spontaneous Emotions by Speech Within Automotive Environment. *Fortschritte der Akustik 32*(1), 57-58.

Takayama, L. 2009. Making Sense of Agentic Objects and Teleoperation: In-the-Moment and Reflective Perspectives. In Proc. 4th ACM/IEEE International Conference on Human Robot Interaction, 239–240. ACM.

Zoltan-Ford, E. 1991. How to Get People to Say and Type What Computers Can Understand. *International Journal of Man-Machine Studies* 34(4), 527–547.