

2010

Is Silence Golden in Human-Robot Dialogue?

Robert Ross

Technological University Dublin, robert.ross@tudublin.ie

Follow this and additional works at: <https://arrow.tudublin.ie/scschcomcon>

Recommended Citation

Ross, R. (2010). Is Silence Golden in Human-Robot Dialogue? *Proceedings of Dialog with Robots, 2010 AAAI Fall Symposium Series, Washington, USA*. doi:10.21427/cz90-ek55

This Presentation is brought to you for free and open access by the School of Computer Sciences at ARROW@TU Dublin. It has been accepted for inclusion in Conference papers by an authorized administrator of ARROW@TU Dublin. For more information, please contact arrow.admin@tudublin.ie, aisling.coyne@tudublin.ie.



*This work is licensed under a [Creative Commons Attribution-NonCommercial-Share Alike 4.0 License](https://creativecommons.org/licenses/by-nc-sa/4.0/)
Funder: Irish Research Council for Science Education and Technology (IRCSET)*

Is Silence Golden in Human-Robot Dialogue?

Robert Ross

Artificial Intelligence Group, Dublin Institute of Technology, Ireland.
robert.ross@dit.ie

The physical actions performed by any robot can be used to convey meaning to a user in human-robot interaction. For example, successfully performing an action following a request may be viewed as an acceptance, while performing the wrong action may be construed as a mis-understanding. Even hesitating to perform a requested physical action may be viewed as a signal of non-understanding. Thus, unlike in the more orthodox domain of non-situated dialogue, natural human-robot dialogue must account for physical actions as a natural and effective implicit communication channel.

Though physical actions have not always been explicitly accounted for in dialogue act annotation schemes and models, e.g., vanilla DAMSL lacks a direct mechanism for such implicit communication (Allen and Core 1997), the nature of physical actions as a type of communicative act has been long recognized within the dialogue community (see for example Coulthard & Brazil (1979) for an early account). Indeed, the physical performance of an action can be regarded as a variant on multi-modal interaction (Pfleger, Alexander-son, and Becker 2003). However, while the analysis of physical actions as communicative acts is not new, it is less clear how dialogue planning policies for human-robot interaction should be influenced by the co-occurrence of physical tasks actions. Addressing this issue successfully inevitably depends on knowing whether users consider verbal communication acts alongside physical acts to be superficial or unnatural, and on whether explicit verbal acts can be beneficial given the limitations of imperfect communication.

With these questions in mind, in the following we report on a recently conducted study with an implemented human-robot dialogue system which was designed to assess the importance of compounded physical and verbal communicative acts in human-robot dialogue.

Evaluating Feedback Levels in Human-Robot Dialogue

In this section we briefly summarize the performed study. Before detailing the study and subsequent results, we first introduce the dialogue system and usage scenario for the study.

Copyright © 2010, Association for the Advancement of Artificial Intelligence (www.aaai.org). All rights reserved.

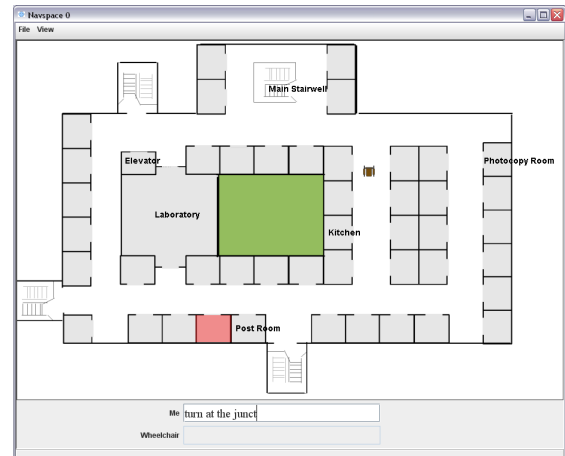


Figure 1: The Navspace application's main screen.

System Design

In this study we made use of an implemented dialogue system which was developed to investigate issues of spatial language and dialogue in human-robot interaction (Ross 2009). While we have previously explored issues of human-robot dialogue with physical robotic wheelchairs (e.g., (Shi, Mandel, and Ross 2007)), in the current work we have focused instead on a simulated environment in order to focus best on language processing concerns. In this simulated robot environment, which we refer to as the *Navspace* application, a user gives verbal instructions to direct an autonomous robot around a partially known office environment. Both the robot and user are assumed to have shared knowledge of their environment, but, while the robot is capable of moving in the environment and following the user's instructions, it is the user who has knowledge of their target destination.

During a typical interaction with the robot, a user views the shared simulated environment from a survey perspective which includes corridors, various rooms, and the robot's position in the environment at any given time. Figure 1 depicts the interface that was used. A user, with a particular known target destination (indicated as a darkened location on the map), is free to direct the robot towards that destination in whatever way that user sees fit. As shown in the bottom

of the figure, interaction with the Navspace application was based on textual rather than spoken channels. This was done to minimize the difficulties introduced by speech recognition and synthesis, and thus allow greater research focus on representation and modelling issues.

Three variants on the baseline dialogue system were used in the current study. These three variants differ in the amount of feedback given to the user but are identical in every other way. Version 1 provided a straightforward dialogue strategy where clarification questions were posed to the user where necessary and confirmations were conveyed through short acknowledgment statements (i.e., “ok”) even when a physical action was being performed in response to a user instruction. Version 2 provides explicit clarification or confirmation utterances to the user after every recognized user contribution. Thus in addition to clarification questions, explicit acknowledgment statements were formed in response to each successfully interpreted user command. It should be noted that these clarification questions and confirmation statements did not simply repeat or rephrase the surface form of user contributions, but were instead formed from a deep contextualization of user utterances. Finally, Version 3 provided verbal feedback only in the case where explicit clarification questions were necessary; thus no verbal feedback was used where a physical action could be interpreted as providing the necessary illocutionary response.

Procedure & Participants

This study was web-based and involved users interacting with the application through a Java applet which connected to a server that performed all computational modeling and processing. Each participant was randomly assigned to one of the three dialogue system variants. Once assigned, participants were then required to perform 5 separate tasks involving directing the simulated robotic wheelchair to a target destination. Destinations and initial locations were randomly ordered, and initial instructions given to the participants were explicitly designed not to prime any specific language usage or discourse strategies. Following the completion of all five trials, users were required to fill out an on-line questionnaire which collected both demographic information and a number of Likert-scale (five-point) ranked questions for system analysis. The data for 34 participants were retained for analysis.

Results

Table 1 summarizes initial results of the study in terms of mean questionnaire results and a number of objective metrics. Two things are immediately notable. There is a direct correlation between the amount of feedback provided by the system and the mean user utterance length and number of utterances. More interestingly, while the minimal feedback system (V3) was not perceived as particularly robust (high system rejection and clarification rate as well as a low value for user perception of robustness), users perceived this *silent* version as better than the verbose and intermediate systems on a number of subjective measures (note that difference in preference for this system over a joystick based system was statistically significant ($p < 0.05$)).

Table 1: Study Results. V1 = Dialogue system with short verbal acknowledgments. V2 = Dialogue system with complete verbal acknowledgments. V3 = Dialogue system with no verbal acknowledgments.

	V1	V2	V3
Rather use this system (verbal) than a joystick based system	2.17	2.33	3.00
System was robust	3.00	3.50	2.90
System was efficient	3.42	3.25	3.50
System was friendly	3.67	3.92	4.00
Enjoyed working with system	3.58	3.58	3.90
Pace of interaction acceptable	3.08	3.75	4.10
Number of User Utts.	39.67	43.58	35.80
Average Time Taken	91.94	112.33	93.73
Number of System Rejections	2.25	4.17	4.00
Number of System Clarifications	5.92	5.67	7.50
Average User Utt. Length	2.04	2.27	3.35
Number of Participants	12	12	10

Conclusions & Future Work

In this paper we reported on a study which analyzed the relative importance of omitting verbal feedback in situated human-robot dialogue. We found that while a lack of explicit feedback can and does lead to more errors in dialogue, overall task performance times are improved, while users perceive the resultant system as being better on a number of subjective measures. We recognize that these results are likely specific to the language used in our dialogue system, and indeed the particulars of the interaction scenario we investigated. Nevertheless, we believe that the question of appropriate coordination of physical behavior with discourse policy is an important facet of human-robot dialogue research, and we thus hope to continue investigating these issues in future research.

Acknowledgments The author gratefully acknowledges the support of the Irish Research Council for Science, Engineering and Technology (IRCSET).

References

- Allen, J., and Core, M. 1997. Draft of DAMSL: Dialog act markup in several layers. Unpublished manuscript.
- Coulthard, R., and Brazil, D. 1979. *Exchange Structure*. Number 5 in Discourse Analysis Monographs. Birmingham: University of Birmingham English Language Research.
- Pfleger, N.; Alexandersson, J.; and Becker, T. 2003. A robust and generic discourse model for multimodal dialogue. In *IJCAI-03 Workshop on Knowledge and Reasoning in Practical Dialogue Systems*.
- Ross, R. J. 2009. *Situated Dialogue Systems: Agency & Spatial Meaning in Task-Oriented Dialogue*. Ph.D. Dissertation, Universitat Bremen.
- Shi, H.; Mandel, C.; and Ross, R. J. 2007. Interpreting route instructions as qualitative spatial actions. In *Spatial Cognition V*, Lecture Notes in Artificial Intelligence. Springer Verlag; 14197 Berlin; <http://www.springeronline.com>.