Robot Says Hello: An Analysis of Human-Robot Conversation in the Real World

Sadhvi Darisipudi
Department of Computer Science
University of Texas at Austin
Austin, TX 78705
sdarisipudi@utexas.edu

Tracy Zhang
Department of Computer Science
University of Texas at Austin
Austin, TX 78705
tczhang@utexas.edu

***Abstract*— This paper reports the results of human-robot conversation in a real-world setting with the goal of exploring interaction patterns and understanding what tasks human participants want the robot to do for them. We controlled a Building-Wide Intelligence (BWI) Bot using the Wizard of Oz technique in a high-traffic academic building to converse with participants. With the assistance of a data logging system, the dialogues were recorded, transcribed, and analyzed. The results show that participants' responses to the robot can be categorized into two primary categories: questions and requests. These categories can be further categorized into subcategories but regardless, the most common general question posed to the robot was "what is your name?" while the most common request for the robot was prompting for directions around the building.**

I. INTRODUCTION

This project was completed as a part of the Building Wide Intelligence (BWI) Project at the University of Texas at Austin. The aim of the BWI Project is to develop fully autonomous and functional robots that assist students and visitors as an integral part of the Computer Science Gates Dell Complex (GDC) building. In order to reach this goal, it is important to gain an understanding for what kinds of tasks people, especially common visitors of the GDC building, would like for the robots to be able to complete. This project achieved this goal by surveying participants in the GDC for such tasks and studying how the robot most effectively engages people with respect to conversation.

II. BACKGROUND

Before proceeding with this project, a literature survey was taken and exploration was done into other research papers and studies that sought to accomplish a similar goal. While no study was found in which a robot approaches a person with the intent of asking what the robot can do for them or simply more free-flow conversation, many studies about robots approaching humans and the factors that influence this interaction do exist.

For example, Satoru Satake and other researchers at the Advanced Telecommunications Research Institute (ATR) conducted a study on strategies for a robot to initiate interaction with humans [1]. In this particular study, the researchers developed a technique that improved the likelihood of a person engaging with the robot. As a part of this technique, the robot would identify a person who seemed to be idle and predict their walking behavior. If this person could be reached in a frontal direction, the robot would then intercept the predicted walking path thus effectively engaging the person [1]. While there is a movement aspect to our study similar the approach taken in Satake's paper, the focus of our project is still more on analyzing the content of human robot conversations rather than just initiating interactions.

Furthermore, there are many other HRI studies which explore the diverse methods people communicate with robots using natural language [3][4]. For example, Terrance Fong and other researchers developed a collaborative control model in which a human and a robot work as partners through dialogue, collaborating to perform tasks and to achieve common goals [3]. Meanwhile, Cristen Torrey and other researchers proposed that an adaptive robot could estimate the information needs of individuals and change its conversational dialogue to suit these needs [4]. However, both these studies are focused more on processing, understanding, and adapting generated language rather than simply analyzing the underlying content of human robot conversations that take place.

Thus, while these HRI studies regarding robots approaching people and adaptive natural language are somewhat related to this study, our project is still unique and focuses more on the conversational aspect of the interaction, especially what tasks people want to see implemented on the robots.

III. HYPOTHESIS

We hypothesize that the most genuine responses will come from the people approaching the robot themselves, while more non-serious responses will be generated when the robot approaches people. This is because people who are curious and interested will want to talk to the robot since they see it as something novel or noteworthy, while others who do not approach it themselves may be more uninterested or ambivalent.

We also predict that the most common questions and requests asked of the robot will be in regards to directions to a particular location. This is because given the environment of our study, which is the atrium of a large academic building, people will be more likely to ask general questions related to rooms within the space or prompt the robot with similar location-related requests.

IV. METHODOLOGY

We have divided our project into two major components for organization and to ensure efficient implementation. First, there is the logging system used to record each human-robot interaction that takes place. Second, there is the movement behavior of the robot leading up to the interaction as well as the software developed used to remotely control the robot.

A. The Logging System

The first step we took in completing this project, and perhaps its most core component, was creating a logging system that allowed for data collection. This logging system uses speech-to-text software to create a transcription of the human-robot dialogue that takes place. While data collection can be performed manually by listening to, timing, and transcribing this dialogue, automation of this process allowed for faster information collection as well as more uniform data. Additionally, with this tool, the experiment can be conducted more autonomously and the transcription aspect does not have to be supervised by a researcher. Furthermore, the logging system was developed in such a way that allows for it to function independently of all other components of this project. By doing so, this system can also be used as a data retrieval and recording tool for other future projects, especially further studies on human-robot interaction. The logging system that we developed transcribes dialogue using Google Cloud's Speech-to-Text API. The dialogue is outputted into an XML file and formatted with descriptive tags including a record of the start, end, and duration of the interaction. Prior to use, this logging system was manually tested for accuracy. It is important to note that while the XML file outputted as expected, the transcription accuracy is limited by the capabilities of Google Cloud's Speech-to-Text API as well as the clarity of the participant's speech.

B. The Movement Aspect

The second component of our project involves controlling the movement of the BWI Bot. There are many ways a human-robot interaction may take place but a critical part of that process is first initiating the social encounter. In our study, a proper social encounter between the robot and human is defined as the BWI Bot being positioned at a personal or conversational distance (1.5-4 feet) away from a person in which the robot (or camera) is facing the person, the person is facing it, and both can easily converse and interact with each other.

1) Wizard of Oz Method: As previously mentioned, the control of the robot (speech and movement) utilizes what is known as the Wizard of Oz (WoZ) technique. In this method, a robot is human-controlled, but participants who interact with it are under the impression that it is operating completely autonomously. To do this, we wrote software that enables the controller of the robot to stay at a discreet distance while remotely-controlling the BWI Bot to drive up to people to interact with them. The conversational dialogue that the robot "says" to a participant is also controlled by a researcher. A benefit of this technique is that targets can easily be chosen to approach and initiate an interaction with. All responses collected through this method were recorded using the previously discussed logging system.

2) Implementation of WoZ: In implementing the Wizard of Oz approach, we set up a VPN connection between a BWI Bot, a server desktop, and another client desktop using OpenVPN. This VPN connection allows for a more direct and secure data transfer process. The data transferred between the robot and the client desktop is a live video what the robot sees as well as instructions for movement and also text for how to respond in conversation.

As mentioned, text was typed on the client desktop which was then spoken aloud by the robot. To render this text-to-speech response system functional, we used Festival Speech Synthesis. While the speech synthesis could be understood for the most part, the generated voice had a very robotic and unnatural sound which was at times difficult to understand. Therefore, the experiment was to some extent also hindered by the capabilities of this speech synthesis software as some participants had trouble understanding the robot.

As for remote video streaming, it was a priority to minimize video lag. Thus, we decided to send a compressed video stream, frame by frame, through our established VPN connection. We wrote a program to stream the video feed to a computer window from the robot's Kinect camera to the client desktop. Furthermore, to run the same ROS environment on both machines and to have access to the necessary rostopics, we set the robot as the ROS master on the client desktop.

The final component of this technique is to remotely control the movement of the BWI Bot. Our approach to accomplishing this is similar to that of remote video streaming. Because we were running the same ROS environment over both machines, we simply accessed the teleop functionality already available on the BWI Bot through setting the robot to be the ROS master. The teleop function allowed us to remotely control the speed and direction of the movement of the robot.

V. EXPERIMENTAL SETUP

It is important to note that we did not control many variables within our experimental environment as we want the responses from people to be as natural and reflect as much of the real-world as possible. Additionally, while the human-robot interaction occurs, another researcher observed from a distance and recorded as many qualitative observations as possible (for example, if classes just ended, it is nearing lunchtime, there is a networking event, etc). Due to these characteristics of this study, our project is much more flexible and is more of an ethnographic study rather than a controlled experiment.

1) Location: In designing the experiment, we decided to run the study in the atrium of the computer science building or GDC. This specific location was chosen for a number of reasons. As has been discussed, this project is being completed as a part of the BWI Project at the University of Texas. The goal of this project is for the robot to be used by people at the GDC. Thus, it logically follows that the best people to survey about what they would like the robot be able to do are those at the GDC. Furthermore, we chose the location of the atrium as it has the most foot traffic as well as a relatively diverse range of people. This is because students, professors, staff, as well as university visitors all tend to walk through the atrium. Therefore, the atrium served as the ideal location to run the study as it was the best reflection of the people that would be eventually interacting with the BWI Bot.

2) Participants: We also employed a loose strategy for choosing who to approach while conducting the study. Since the control and movement of the robot was not autonomous, we were able to analyze social cues to figure who was most likely to interact with the robot and approach them. For example, people wearing headphones were oftentimes not approached as they demonstrated an unspoken desire to not be bothered. Furthermore, people who appeared to be in a rush were also avoided as they would be unlikely to stop and take time to converse with the robot. Essentially, the strategy employed in seeking out people to talk to involved targeting more idle people who seemed to be more open and likely to engage with the BWI Bot. Overall, 517 students and GDC visitors took part in the study. They were not paid for participation in any way.

3) Robot: The BWI Bot used in the study was designed to eventually become a GDC building assistance of sorts. In this study, we attached a microphone and speaker to the robot for the speech-to-text API to pick up recordings and generate speech. The desktop on the robot displayed the terminal running teleop and the RViz localization map of the robot.

4) Dialogue: The final component involved in the design of this experiment was deciding on the dialogue and what the robot should say. It was important to us to minimize prompting as much as possible to allow for the most organic flow of conversation and to leave questions and task requests as open ended as possible. Therefore, we initially designed the dialogue to start with the simple greeting of ”Hello”. This particular phrase was used to create a more natural dialogue and allow genuine interest to develop, and as a result, hopefully more serious requests for task implementations to come into the conversation. We soon came to realize that this greeting was not enough for people to stop and converse with the robot as many people seemed to be unsure of how to respond. Due to this observation, we used the question ”What can I do for you?” in several cases as the participant interacting with the robot better understood what to say to the robot. As for the robot’s conversational responses, we aimed to keep them as neutral as possible for the same reasons.

Thus, this is the procedure we used to remotely control the robot to drive up to people and engage them in conversation. We also ran our developed speech-to-text transcription software while conversations took place.

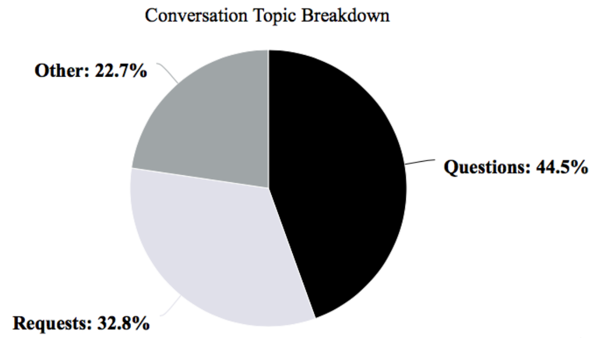

Fig. 1. Pie chart of human and BWI Bot conversation topic breakdown.

TABLE I
CONVERSATIONAL ENGAGEMENT

Category	Value
Total Number of People Approached	517
Total Number of People Engaged	61
Conversational Engagement Percentage	11.79%

VI. RESULTS

Upon collecting data, we organized it into several categories of verbal responses and nonverbal reactions. The three overarching categories of responses are ”Questions”, ”Requests”, and ”Other”. A response was categorized as a question if the person had asked the robot about something, requests consisted of a person asking the robot to do something, and lastly the ”Other” category encompassed all reactions we received that do not fall into either of the two preceding categories. We found that 44.5% of participants asked questions to the robot, while 32.8% gave it a task to do. Finally 22.7% of participants spoke the the robot in other ways (see Fig. 1).

Another analysis performed on the data collected was classifying them into proactive and reactive responses. Proactive responses were those in which a person approached and spoke to the robot themselves without any sort of prompting. On the other hand, reactive responses were those in which the robot was controlled to approach a person and then successfully spoke to a person; these responses were simply a reaction of sorts to the robot. Overall, we had 517 participants in this study. Of those, only 61 participants or 11.79% actually talked to the robot (see Table I). Of all these verbal responses, 19.67% were proactive while 80.33% were reactive (see Table II). In general, the conversations that did take place had an average duration length of 71.39 seconds.

Next, we split the ”Questions”, ”Requests” and ”Other” sections into more subcategories to better analyze interaction patterns and understand what people generally want the robot to do. See Table IV for a percentage breakdown of the conversational topics that occurred.

A. Questions

We split the ”Questions” section into more subcategories as it is still rather broad and ambiguous. The questions asked of the robot can be grouped as: Anthropomorphic,

TABLE II

CONVERSATIONAL RESPONSE TYPES

Category	Value
Total Number of Proactive Responses	12
Total Number of Reactive Responses	49
Total Number of People Engaged	61
Proactive Response Percentage	19.67%
Reactive Response Percentage	80.33%

TABLE III

MOST COMMON QUESTIONS AND REQUESTS

Questions	Requests
What is your name?	Take me to the closest [location].
What can you do?	Go to [location].
Where is [location]?	Drive to that [object].
Where is the [object]?	Spin in a circle.
Are you able to [verb]?	Go forwards/backwards/around me.

Seeking Information, Pleasantries, Trivia, Seeking Info about the Robot, and Non-serious Questions.

1) Anthropomorphic: Anthropomorphic questions consisted for those that treated the robot like a person with real feelings, opinions, and emotions. An example of such a question we received from a person is "What is your favorite color?" as this question implies the robot has opinions or thoughts on colors.

2) Seeking Information: An informational or information seeking question is one that asks for factual information. An example of such a question is "Where is Peter Stone's office?"

3) Pleasantries: Another type of question is commonplace sayings or pleasantries. For the most part, this category consisted of the questions "How are you?" or "How are you doing?" which was asked to the robot several times.

4) Trivia: Another category of questions is trivia. These questions were those that asked about general facts or something that would usually be found through the use of a search engine such as "How tall is the Empire State Building?".

5) Seeking Info about Robot: The next category is questions regarding the robot itself which included questions such as "Who is your creator?" or "What is your name?".

6) Non-serious: The last category consisted of non-serious questions which were asked more as rhetorical jokes and were not meant to be real questions. For example, a person asked the robot about issues in their code for an homework assignment, saying "Do I have pointer errors?" when they were well aware the robot did not have access to their code.

B. Requests

Similar to "Questions", this section was also split into several subsections: Location-related, Dummy, Non-serious, and other Miscellaneous Requests.

1) Location-related: The first category of requests was Location-related tasks This category encompasses all requests asking the robot to take them to a desired location or the find the closest or a certain amenity and them take them there. An example of such a request is "Take me to the closest elevator".

2) Dummy: Dummy Requests were requests that served no real purpose and were mainly asked of the robot only to test its functionality and/or to see if the robot could actually understand them. There were several examples of such request such as "Spin in a circle". The person asking the robot to complete such a task gained little to nothing from the robot performing the task but rather only tested to see if the robot was functional. Dummy requests can be distinguished from the next category (Non-serious requests) because they are very simple things such as spinning or moving from one point to another.

3) Non-serious: The next category of requests, as previously mentioned, is Non-serious requests. These were more complicated tasked that were jokingly asked to the robot that were not truly expected of the robot due to how outrageous they were. A good example of such a request was "You should juggle" or "Cook me a three-course meal".

4) Miscellaneous: The final category of requests is simply Miscellaneous tasks. This is the largest subcategory that groups everything that does not belong in the other three subcategories. The requests in this subcategory are all quite random and cannot necessarily be grouped further among themselves. Some examples of requests that would belong in this category are "Play music", "Get me a drink", "Race me", etc.

C. Other

This section encompasses everything that falls neither in the category of "Questions" nor in the category of "Requests". The "Other" category are things that were said and can be split into these sections: Compliments, Insults, Clarifications, and Nonsensical Statements.

1) Compliments: Compliments are kind things said to the robot such as "You are the smartest robot I have ever met".

2) Insults: Insults were disparaging remarks made such as "You are creepy".

3) Clarifications: The next subcategory, and the largest subcategory is Clarifications. This encompasses people asking the robot to repeat itself or speak louder as they could not understand what was being said. This number was surprising large and was mostly due to weaknesses in Festival speech-to-text technology as well as the hardware limitation of the volume on the speaker.

4) Nonsensical: The final subcategory is Nonsensical Utterances that did not mean much or simply could not be understood. An example of a nonsensical utterance is "I need a man".

D. Other Reactions

While many people chose to speak to the robot, many other people communicated with or engaged with the robot in a nonverbal fashion. Of the 517 participants, 69 or 13.35% engaged with the robot in nonverbal ways. This could be split into four main categories; waving, filming

TABLE IV
CONVERSATION TOPIC PERCENTAGES

Topic	Number	Percentage
Questions		
Anthropomorphic	6	2.78%
Seeking Information	47	21.80%
Pleasantries	10	4.63%
Trivia	7	3.24%
Seeking Info About the Robot	22	10.19%
Non-serious	4	1.85%
Requests		
Location-related	4	1.85%
Dummy	13	6.02%
Non-serious	17	7.87%
Miscellaneous	37	17.10%
Other		
Compliments	4	1.85%
Insults	2	0.93%
Seeking Clarification	42	19.44%
Nonsensical	1	0.46%
Total	216	100.00%

TABLE V
REACTION TYPE PERCENTAGES

Reactions	Number	Percentage
Waving	33	47.83%
Taking Pictures/Filming	21	30.43%
Looking at Screen	11	15.94%
Touching Screen/Keyboard	4	5.79%
Total	69	100.00%

and/or taking pictures, looking at the screen, and touching the screen or keyboard (see Table V for more information). Waving oftentimes took place with the robot approached a person with the greeting ”Hello” and the person usually had no intent of engaging in an actual conversation so they simply waved and then left. Filming and taking pictures also happened many times as many people recognized the robot as something new in the atrium, something out of place, or simply something that was noteworthy. Next, people who did not watch or see the direction in which the robot moved mistook the screen to be the front of the robot and thus tried to read the screen to understand what was taking place. Finally, a few people even tried to touch the touch-screen monitor on the robot. These people often tried to read and understand the open terminal windows and some even tried to open a browser window to access the internet through the robot.

VII. DISCUSSION

Overall, we have found that the most common question asked of the robot is ”What is your name” while the most common requests were location-related oftentimes involving directions. See Table III for a breakdown of the top questions and requests asked of the robot. Furthermore, there were more non-serious questions asked when the response was reactive (71.23%) than proactive (28.77%), thus both our hypotheses were correct.

A. Factors Affecting Engagement

As is reflected in table I, there was a relatively low engagement percentage (11.79%) which may have been due to a number of factors. While running this experiment, we observed several factors that affected the likelihood of a person engaging in conversation with the robot. Some of these factors could be controlled while others could not but those that could be controlled were used to our advantage and to maximize engagement with the robot.

1) Phrasing: While running our study, we noticed that the phrasing used to get the attention of a person significantly affected how they reacted. In the beginning of the study, we began with using the simple greeting of ”Hello”. This did not work well in engaging most people in conversation as a majority of participants ignored the robots or waved and walked away. Changing this opening line to the phrase ”Hello. What can I do for you?” increased the prompting but improved engagement as people seemed to have more direction in how to respond to the robot. Even when using this phrase, however, many people still ignored the robot. The most useful technique in grabbing the attention of a passerby proved to be specifying the exact person the speech was directed at by pointing out a distinguishing characteristic of that particular person. The characteristics we chose for the robot to ”point out” and ”say” tended to be articles of clothing or objects the person was holding. An example of such a phrase spoken by the robot is: ”Excuse me. Man in the black cap, what can I do for you?”. By specifying the intended recipient of the question and eventual conversation, the likelihood of them holding dialogue with the robot increased greatly. This technique was used many times during the study to better engage people in conversation and collect data.

2) Individuals vs. Groups: Another observed variable that affected whether or not a person interacted with the robot was how many other people they were with. A group or pair of people tended to be more likely to speak to the robots while individuals were less likely to do so. In multiple cases, encouragement from an individual for another individual to speak to the robot was observed. Furthermore, individuals seemed to be more uncomfortable with speaking to the robot while the conversation was more natural and even jovial in a group setting.

3) Walking vs. Sitting: Next, the movement or state of the person also influenced whether or not they spoke to the robot. Through the remote control of the robot, we approached people walking by and people sitting in the atrium. To our surprise, people who were sitting in the atrium interacted with the robot significantly less. In fact, one group of student sitting in the atrium even told the robot to ”go away”, indicating that they were busy. This was probably due to the fact that people who were sitting were engaged in some other activity such as doing homework and were unlikely to interrupt this to talk to the robot. In contrast, people walking through the atrium were not usually actively engaged in any other task and thus were more likely to talk to the robot.

B. Success of Project

Given that this study is centered around interaction and conversation between a robot and a person, we expect the data to lack a certain degree of objectivity. Due to this, measuring the success of the experiment is difficult. However, we did this by looking at the size of the sample space of people we are able to collect data from. The more people we were able to involve in the experiment, the greater understanding we could gain for what kinds of tasks people want the BWI Bots to complete and the better we can understand how different variable affect the human-robot interaction that takes place. We believe that this study involved a decently sized sample size as we were able to approach over five hundred people. Furthermore, we believe this sample to be an accurate representation of the larger population of the GDC. However, it is difficult to decide whether or not enough people have participated in the study. In the case of this study, the limiting factor was time and the study could not be run for longer but it is fair to assume that running the study longer would have produced greater breadth in results.

Another metric for measuring the success of this project is to analyze the effectiveness of the logging system. The more accurate the data collection from the logging system is, the more successful this project can be deemed to be. While the logging system performed just as expected, it had several weaknesses. For example, the accuracy of the text generated in reflection of the speech was limited by the capabilities of the Google Speech-to-Text API. Several inaccuracies between what was said in real life and what was recorded in the log were observed. In addition, the logging system was designed to split conversations by observing silences and creating a new conversation in the log every time a pause of approximately 30 seconds was observed. However, in many cases conversations between the robot and separate people took place back-to-back which did not allow for the conversations to be separated in the log.

Overall, the experimental design and the software developed allowed for the experiment to succeed in achieving its goal of understanding what visitors of the GDC would like the robot be able to do. While there were weaknesses in the software and way study was conducted, we believe it is fair to say that the project was successful overall,

C. Limitations

As discussed, while there were weaknesses in the technology and experimental design such as difficult to understand speech synthesis, there are also a number of limitations regarding the applicability of the data received from the study. This is due to a number of factors surrounding the experiment performed. We noticed from the data received that many of the requests and questions posed to the robot were circumstantial and/or relevant to something that was happening around the time the study was run. For example, one non-serious request asked by an individual was "Go make Mike Scott miserable". This study was performed around the time of finals at the university and therefore a presumably frustrated student joking expressed ill will towards Mike Scott, a computer science lecturer. While this was not a serious request of the robot, this example serves to show that the questions and requests people had for the robot were shaped in part by other things happening during the time this study was run. If the study had been run earlier this year or later this year, it logically stands that different data regarding questions and request may have been collected. For example, had the study been run during football season at the university, students may have a asked the robot about football scores or where to watch an upcoming game. Therefore, a limitation of the study is that some of the data is specifically reflective of the time of year the study was run and may not be an accurate depiction of what would have been said during another time of the year.

Another limitation of the applicability of the data collected is that this experiment was very much tailored to the larger project it is a part of, the BWI Project. For example, the experiment was only run in the university's computer science building as the goal of the BWI project is to have fully autonomous robots in this building. While the data collected is useful to the purpose of the BWI Project, it may not be as applicable to any other studies or a different situation. This is because the data is not reflective of a larger population or people in general, but rather only university students, faculty, and visitors that specifically go to the computer science building. Therefore, the data collected is useful for the purposes of this project but the it may not be as applicable to other situations or studies which is another limitation.

VIII. CONCLUSION

Ultimately, the aim of our project was to contribute to the overarching goal of the BWI Project which is to develop fully autonomous and functional robots that assist visitors in the GDC building. Using the Wizard of Oz technique, we controlled the BWI Bots to roam the GDC atrium collecting and logging data from all human-robot conversations that took place. We found that people responded to the robots differently with conversational topics including questions, requests, and others as well as different nonverbal reactions. Ultimately, people most commonly want the robot to give them directions and in general they do not know what sorts of capabilities already exist on the robots.

IX. FUTURE WORK

There are many ways in which this project can be extended in the future. For example, the experimental setup and the way the interaction between human and robot was setup may have influenced the responses received. Because the robot was not autonomous but actually human controlled, there was a certain degree of fluidity in the conversation. In fact on one occasion, a person speaking to the robot suspected the robot was not actually performing autonomously, saying "I'm pretty sure someone is typing responses". Therefore, it logically follows that the responses would have been different if the robot had actually been completely autonomous in both movement and speech. A useful extension of the project would then be to develop software for the robot to able to autonomously seek out an idle person and converse with them. This would also be useful to do as it ties in better with the BWI Project. The goal of the BWI Project is to have the BWI Bots be completely autonomous and become an integral part of the GDC ecosystem. Therefore, the best reflection of the interaction between a person and the robot would be to have the robot perform autonomously as this is the eventual goal of what will happen.

Another way to expand this project in the future is to develop an autonomous patrol for the robot. As hypothesized and previously discussed, the most genuine responses were the proactive responses in which the person approached the robot and not vice versa. Therefore, devising a method in which all interactions that take place are proactive responses would produce valuable data. A way to do this is to have the robot autonomously patrol a given area stopping only when a person interacts with it and recording the interaction and data collected. While this would be a useful extension of the project, it would likely require a large period of time to run the study as only twelve responses from this study were proactive responses.

ACKNOWLEDGMENTS

The researchers would like to thank Dr. Justin Hart for his guidance throughout this project. They would also like to acknowledge the lab mentors for their support and advice. Lastly, they would like to give a special thank you to Mayuri Raja for her mentorship.

REFERENCES

[1] Satoru Sakate, Takayuki Kanda, Dylan F. Glas, Michita Imai, Hiroshi Ishiguro, and Norihiro Hagita. 2009. How to Approach Humans? - Strategies for Social Robots to Initiate Interaction. ACM/IEEE (March 2009), 109-116. DOI: https://doi.org/10.1145/1514095.1514117

[2] P.A.M. Ruijten and R.H. Cuijpers. 2017. Stopping distance for a robot approaching two conversating persons. IEEE/ROMAN (August 2017), 224 - 229. DOI: https://doi.org/10.1109/ROMAN.2017.8172306

[3] Fong, T., Thorpe, C., and Baur, C. (2001). Collaboration, dialogues, and Human-Robot Interaction. Proc. Int. symposium of Robotics Research.

[4] Torrey, C., Powers, A., Marge, M., Fussell, S., Kiesler, S. (2006). Effects of adaptive robot dialogue on information exchange and social relations. In Proc. of HRI06, 126-133.

APPENDIX

Below are some raw sentences taken from human side of the conversations recorded with the logging system. As can be observed, there are some inaccuracies in what was said and what was recorded due to a lack of clarity of speech. The API used also omits punctuation. Here are just six examples out of all the transcriptions we collected:

Example 1: ”Do you have a name Roberto nice to meet you yeah yeah the pleasure is mine how smart are you Roberto how smart are you Roberto okay you are probably the smartest robot I've ever seen no problem do you have any beliefs yeah no problem what do you believe what what does that mean okay fair enough pretty cool I got to go do homework but you should go talk to more people okay bye”

Example 2: ”who are your owners isn't there a second one what's your name can you go backwards Roberto okay Roberto can you go backwards yes what's your favorite color I'm going to class bye goodbye”

Example 3: ”I give you a command yourself to do something to spin around okay can you spin around wow that's really impressive I mean ideally I would ask you to like do my homework for me but I feel like your particular model probably can't thank you is nice meeting you Roberto”

Example 4: ”hi do a 360 cool what's the weather out it is nice who is your creator that's cool”

Example 5: ”hello wake me up at 5:15”

Example 6: ”hello there go in circles I told you the answer all my God that's actually really cool do you want to fight I told you I'm pretty sure we're pretty sure someone's like typing responses on this is turning circles I did that what's your name it said his name is Roberto go to the trash can over there Roberto”