A User Study of Robot Reaction in Human-Robot Dialogue

Heng Ming Su
Group five
University of the West of England
Bristol, United Kingdom
Heng2.Su@live.uwe.ac.uk

Chieh Lee
Group five
University of the West of England
Bristol, United Kingdom
Chieh2.Lee@live.uwe.ac.uk

Shih Kai Lai
Group five
University of the West of England
Bristol, United Kingdom
shih-kai2.lai@live.uwe.ac.uk

Huajun Cai
Group five
University of the West of England
Bristol, United Kingdom
huajun2.cai@live.uwe.ac.uk

Jiamin Liu
Group five
University of the West of England
Bristol, United Kingdom
jiamin2.liu@live.uwe.ac.uk

Yihua Zhu
Group five
University of the West of England
Bristol, United Kingdom
yihua2.zhu@live.uwe.ac.uk

***Abstract*—**

In order to make the robot more natural like humans, especially the daily conversation and so on. In this paper, a method of different rotation angles of the robot head, body and the order between the dialogues is proposed and analyzed to verify the relationship between the movement of the head, body position and the sense of the conversation. According to the recording of the experimental film and the rotation angle data of the head and body position during the man-machine dialogue. The analysis results show that speaking while turning the head and body during the pronunciation of the robot not only expresses natural human behavior, such as expressing a respect for the other party, but also indicating that the degree of sentence comprehension. Discussed the turning angle to the natural sense of human perception, with future studies about other head movements like nodding and gimmick silence, etc., Those factors can also enhance the fluency of voice information in the robot.

Keywords—NAO, Human-Robot Dialogue, natural, friendly, intelligent, polite, head gesture and body motion

I. INTRODUCTION

In order to conduct natural anthropomorphic communication in the human-machine dialogue which is the English conversation content. The body language information such as robot head position movement is non-verbal information for attempting to express intentions, emotional attitudes and understanding, etc. And most importantly, it can increase the anthropomorphic level of the robot.

In fact, movement of the head position is very natural in the conversation between people, and people look at the speaker intentionally or unconsciously. In this case, the action of turning the head usually has a clear meaning in the dialogue. For example, deliberately not looking at the speaker or turning around slowly to see the speaker is often meaning that person who is not interested in the topic or disregards the speech.

However, when most people talk, their head movements are unconsciously turning to ask questions. However, both of those behaviors convey body language such as attitudes to other and not merely the content of the conversation.

The experimental motivation of this project is to find the way that robot in the communication can make people feel more natural and speak as real human. Therefore, we increase the fidelity of robot by imitated people's different turning head speed and angles. We hope in this way the robot's speaking content can be more easily understood. In our experiments, we analyze several conversations between Nao-humans and find a close relationship between turning movements and human perception. Turning around when robot being talked can not only express the fluency of dialogue but also get the affirmation of people. The speaker can think more about the sentence that the robot says and ignore the appearance of the robot.

In the current work, we have tested three kinds of research based on the NAO robot head position steering motion and its speech sequence model to expand the level of anthropomorphic. The follow-up arrangements for this article are as follows. Section 2 describes the related work. In Section 3, a head position steering motion based on method is described. In Section 4, results of the experiment are performed. Section 5 discussion, Section 6 conclusion of survey.

II. RELATED WORK

Ishi, C. T. *et al.* (2010) studied the relationship between dialogue and body language during head movement and highlights that if the robot nod and spoke spontaneously, it would make people feel the special behavior is affirmative. Meanwhile, Ishi, C. T. *et al.* (2010) also points out that when it appears in short sentences, the last syllable is more powerful in the meaning of the phrase. Furthermore, Ishi, C. T. *et al.* (2010) also states that there is some relationship between nodding and cocking head to speech information and anthropomorphism. Moreover, Ishi, C. T. *et al.* (2010) and Lee, C. *et al.* (2004) state that the most natural behavior is to nod and at the same time with express a positive behavior pattern and they mention that while cocking head often represents a symbol of it is thinking or not confident and the dialogue is not fluent. After a few years, Liu, C. *et al.* (2012) states that the head behavior would naturally moves with the speech, with sub-language information as intention, attitude and emotion on the human communication, which implies that the robot head movement behavior is very realistic for human interaction. Meanwhile, Liu, C. *et al.* (2012) indicates that both of head

tilt and nodding are more natural and friendlier than just nodding or simply human-faced dialogue through analyzing the experimental model of basic on nod and head tilt motion, then discussed the dialogue of thinking behavior. Furthermore, Oto, K. *et al.* (2017) discusses the delay in giving a response during communication called "silences", which impacts on the important factors of robotic natural dialogues and are classified into semantic silence, syntactic and grammatical silence, interactive silence, and robot silence, and indicates that most people are taking silent as human dialogue not robot silence. Oto, K. *et al.* (2017) also highlight that human beings have special behavior as interactive silence which only close friend can understand and the meaning of silence varies with the content of the robot's dialogue, the age of the experimenter and other factors. In the meanwhile, Morency, L. P. *et al.* (2007) states that when the robot remains silent without moving, participants often considered that as mechanically delayed. Moreover, Morency, L. P. *et al.* (2007) also point out that if the robot makes gestures or head gestures and other actions, this silence is considered to have practical social meaning. Those papers mainly focus on studying how free head movement, nodding and silence impacting on human machine dialogues.

However, these papers are not mentioned that the sequence of the reactions also influence anthropomorphic degree of the robot. Therefore, our study is to find out how the reaction orders influence the users' feeling, based on the information we have obtain from the past studies.

III. METHODS

In order to explore how the reaction sequence of robot influence the appearance, we set up a user study to evaluate the users' feelings while having a conversation with robots. The participants were grouped in pairs and asked to interact with NAO verbally according to the given dialogue. There are three conversation scenarios and the responses of NAO were randomly assigned.

A. *Hypotheses*

When people are intending to start a conversation to a human-like robot, people prefer robots to turn around and reply at the same time, rather than waiting for it to turn around before it replies, since this is a more human-like action.

B. *User Study Design*

In order to create a scenario that the users are going to start the conversation when the NAO is facing somewhere else or focusing on other tasks, we set the experiment to be two participants involved at a time. That is, User A and User B. The performance that the NAO reacts to the user are designed in two modes

Mode 1: Turn the head and the body to face human firstly and then start talking.
Mode 2: Start talking while turning the head and the body to face human.

To make the robot's movement more natural, gestures are added while the robot is talking. Each experiment is separated into three scenarios, in which the lengths of the dialogue are different. Each of the two participants asked one question in each scenario, following the dialogue guiding sheet. The designed dialogue are listed as following. Notice that the response context of the Nao was not given to the participants because we want users to focus on the robot.

During the user study, the NAO was put on the desk that the height of it is close to a normal person, as shown in Figure 3.1. The participants are asked to stand approximately 45 degree to the left and right to the NAO. The NAO is firstly faced to the guider before the conversation start, and it turns to the participant who asks the questions. By doing that, the reactions of the response can be observed during the NAO turning to face the other participant from the previous one. The two modes of the reactions were randomly implemented, participants can see both the reactions to him/herself and the other.

Interaction positions of robot and participants, and dialogues applied in the research is shown in below:

Scenario 1 (short dialogue):
User A: Hi, how are you?
Nao: Great, nice to see you here.
User B: How's your day going?
Nao: Not bad, I enjoy the time singing and dancing last weekend.

Scenario 2 (medium dialogue):
User A: What type of music do you like to listen to?
Nao: I usually listen to pop-music, but I also enjoy classical music recently.
User B: How about your favorite dancing style?
Nao: I like different kinds of dancing, such as hip-pop, Jazz and Ballet.

Scenario 3 (long dialogue):
User A: I'm looking for Bristol Robotics Lab, do you know how to get there?
Nao: Firstly, go straight one hundred meters, then turn left at the next corner, finally go straight for about two hundred meters, you will see the front door of the Lab.
User B: Oh, I see. Then how can I go back to city center after visiting the Lab?
Nao: You can go back to city center by bus. Either M1 or M3 goes to city center, but don't forget to double check with the driver before you buy the ticket.

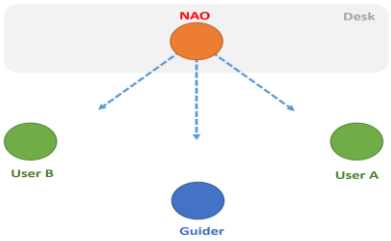


Figure 3.1. Robot and Participants Positions During the Interaction.

C. *User Study Procedure*

The participants were led to the laboratory, after a short introduction, they were given the participant information sheet and the consent form. After they agreed to participate, they were asked to answer several questions for their personal information and background. Then, we led them to the assigned position in front of the NAO, and the user study began. We told them when to begin the conversation and let

them know when the scenario ended. After all the three scenarios were finished, they were asked to complete a questionnaire to rate their feelings of NAO's reaction during the conversations. Finally, there was a small interview asking participants extra questions about their feelings or experience of the NAO robot

D. Dependent Measures

The first questionnaire we collected the basic personal information of the participants, including the age, gender, job, education level and so on. These questions were only related to the factors that we considered could influence the perception of the robot. We also asked about the personalities of the participants and the previous experience of robots. In order to explore the participants' rating to the interaction, we designed the second questionnaire given to participants after the conversation with NAO. In this questionnaire, the participants were asked to rate the feelings of the robot's reaction in each scenario, in terms of polite, intelligent, human-friendly, and how natural the robot behaved. The questions were presented as statements and the participants were supposed to rate how much they agree to the statements from one (totally disagree) to ten (totally agree).

E. Participants

Eight volunteers participated in this experiment (5 males and 3 females), half of them are friends of our experiment professor and half of them are students who have this course. Their age ranged from 21 to 40 years, and four of them are from 21 to 25, three of them are from 31 to 40. Apart from the basic situations, their education levels are all the postgraduate master or higher than it, and the situation about their job is that six of them are full time student, and two of them are employed full time. Finally, the country of them also have differences, four of them are from China, two of them are from UK, and the last two are from Algeria and Lithuania.

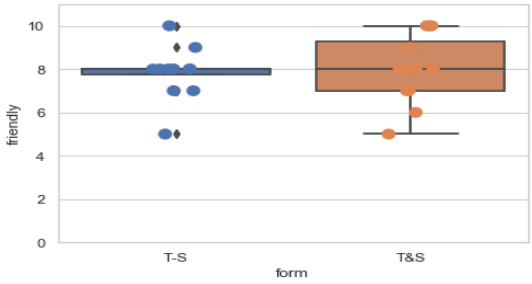
IV. RESULTS

This user study applies four measured factors: the feelings of friendly, intelligent, polite and natural. We use these four factors to determine the better actions order of turning and speaking. Meanwhile, this research quantify the four measured factors conditions into ten integer marks between one to ten. The ten mark states that participants agree with the statement of the robot's actions while one means totally disagree. Furthermore, the study also tries to figure out the impact on four measured factors from the length of dialogue between robot and human and personal experience.

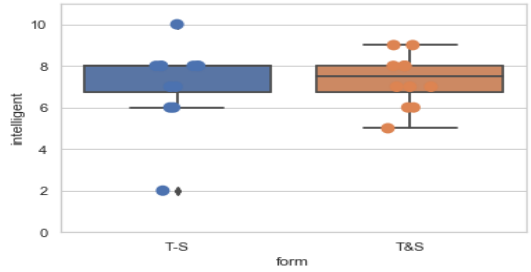
A. Action Order

Action order in the study are divided into T-S and T &S two order groups. T-S order group implies that the robot will turn its body to the participant first and then speak. Meanwhile, T & S order indicates that the robot will speak and turn at the same time.

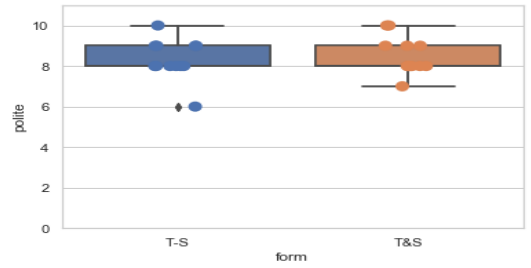
According to Figure 4.1 and Figure 4.2, through the distribution of marks on the measured factors, we can learn that action order has a large impact on the friendly feeling but less impact on intelligent, polite and natural feelings. Meanwhile, in terms of data from Table 4.1 and Table 4.2, the means and variances of two action orders have not such a big difference and the order of speaking and turning body at the same time have higher marks, which implies the users prefer T&S order.



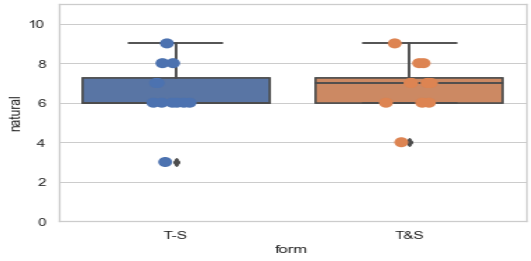
(a) Box plot of action order and friendly



(b) Box plot of action order and intelligent

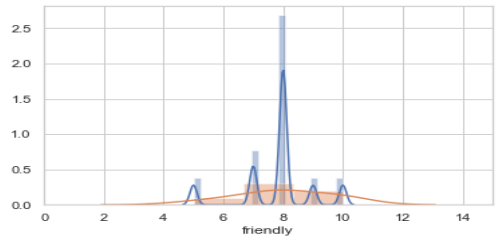


(c) Box plot of action order and polite

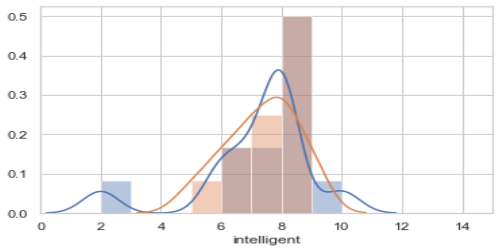


(d) Box plot of action order and natural

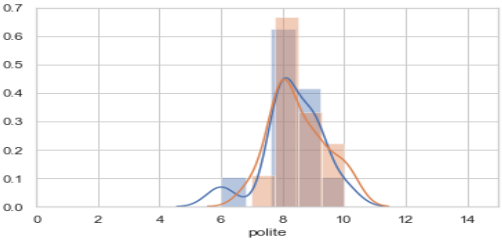
Figure 4.1. Box plots of friendly, intelligent, polite and nature feeling with human four measured factors and action order of turning body and talking with human



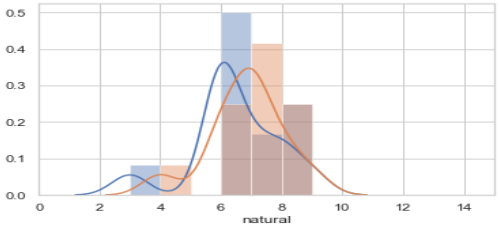
(a) Graph of parametric tests between action order and friendly



(b) Graph of parametric tests between action order and intelligent



(c) Graph of parametric tests between action order and polite



(d) Graph of parametric tests between action order and natural

Figure 4.2. Graphs of parametric tests between action order and friendly, intelligent, polite, nature with human four measured factors, where the blue color and orange color represent the T-S and T & S action order respectively and x axis and y axis are shown the possibility and mark respectively.

	Friendly		Intelligent		Polite		Natural	
	X	U	X	U	X	U	X	U
T&S	8.00	1.60	7.30	1.23	8.50	0.90	6.83	1.27
T-S	7.83	1.19	7.16	1.95	8.30	0.98	6.50	1.51

Table 4.1 Means (X) and variances (U) of parametric tests between action order and friendly, intelligent, polite, nature with human four measured factors.

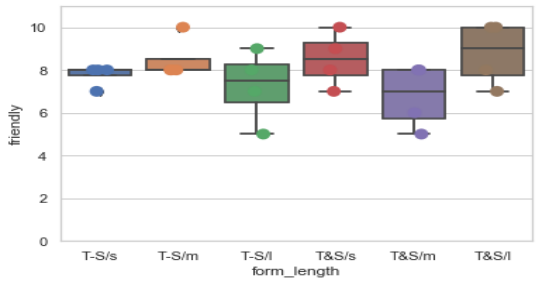
	t	p	d
Friendly	-0.29	0.77	0.42
Intelligent	-0.25	0.80	0.20
Polite	-0.43	0.67	2.49
Natural	-0.59	0.56	1.37

Table 4.2 Interoperation of Cohen's of parametric (d) tests and statistic (t), probability (p) between action order and friendly, intelligent, polite, nature feeling with human four measured factors.

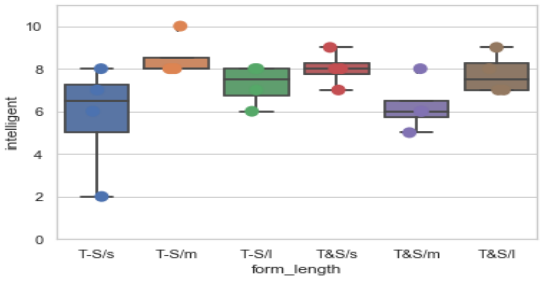
B. *Length of Dialogue*

Length of dialogue in this study consists of short (s), median (m) and long (l) length three length of dialogue. In order to figure out the impact on the measured factors from length of dialogue, the experiment combine the length of dialogues and action orders to produce T-S/s, T-S/m, T-S/l, T&S/s, T&S/m, T&S/l six dialogue and action pairs. For example, T-S/s means that the robot will do action order in T-S order group with short length of dialogue.

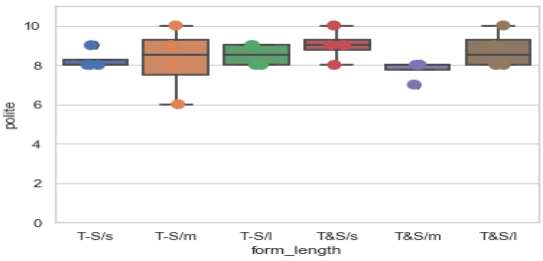
In terms of Figure 4.3 and 4.4, the length of dialogue will have large impact on the friendly feeling and will have some effect on the intelligent and polite feelings. Meanwhile, the length of dialogue have a few effect on the natural feeling by summarizing distribution of mark on the measured factors and length of dialogue. But in all, according to Table 4.3, the length of dialogue doesn't have too much effect on human feelings.



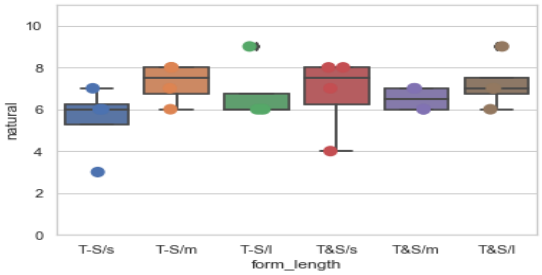
(a) Box plot of action order and friendly



(b) Box plot of action order and intelligent

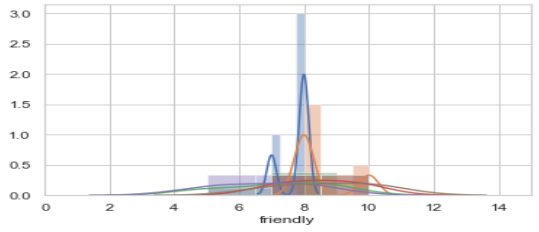


(c) Box plot of action order and polite

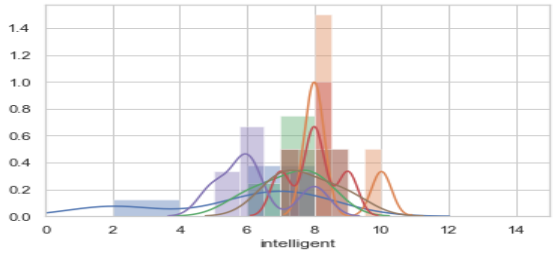


(d) Box plot of action order and natural

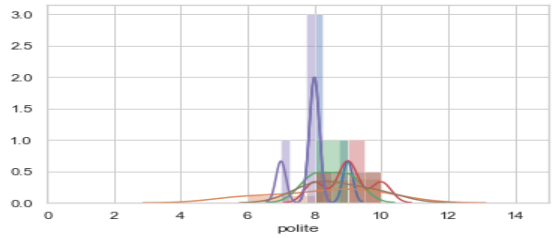
Figure 4.3 Box plots of friendly, intelligent, polite and nature four measured factors and length of dialogue



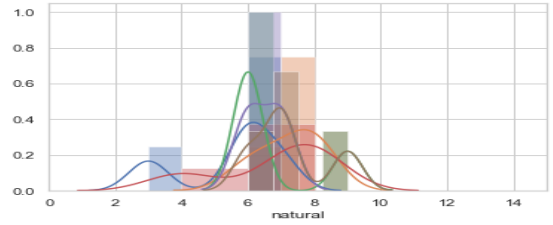
(a) Graph of parametric tests between six dialogue and action combinations and friendly



(b) Graph of parametric tests between six dialogue and action combinations and intelligent



(c) Graph of parametric tests between six dialogue and action combinations and polite



(d) Graph of parametric tests between six dialogue and action combinations and natural

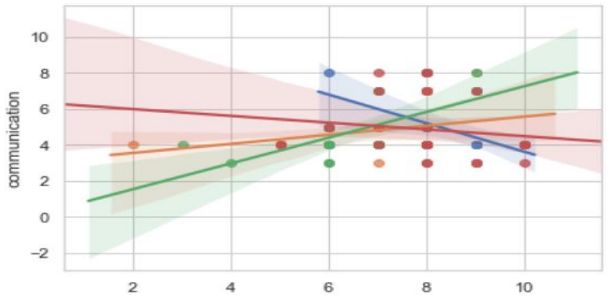
Figure 4.4. Graphs of parametric tests on six dialogue and action combinations and friendly, intelligent, polite, nature four measured factors, where the blue color, orange color, green color, red color, purple color, brown color represent T-S/s, T-S/m, T-S/l, T&S/s, T&S/m, T&S/l six dialogue and action combinations respectively and x axis and y axis are shown the possibility and mark respectively.

	Friendly		Intelligent		Polite		Natural	
	X	U	X	U	X	U	X	U
T&S/l	8.75	1.50	7.75	0.96	8.75	0.96	7.75	1.26
T&S/m	6.75	1.50	6.25	1.26	7.75	0.50	6.50	0.58
T&S/s	8.50	1.29	8.00	0.82	9.00	0.82	6.75	1.89
T-S/l	7.25	1.71	7.25	0.96	8.50	0.58	6.75	1.50
T-S/m	8.50	1.00	8.50	1.00	8.25	1.71	7.25	0.96
T-S/s	7.75	0.50	5.75	2.63	8.25	0.50	5.50	1.73

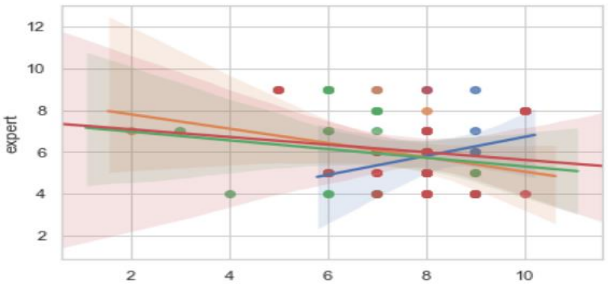
Table 4.3 Means (X) and variances (U) of parametric tests on six dialogue and action combinations and friendly, intelligent, polite, nature four measured factors

C. Personal Experiences

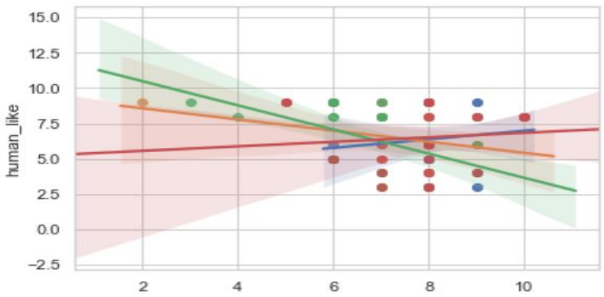
Personal experiences in the study are divided into four evaluated standards with levels in each standard: the levels of being familiar with communication with people you do not know (communication), the levels of being an expert in the robotics (expert), the levels of being familiar with human-like robot (human like) and the levels of how the people like robot (like robot). Furthermore, these standards can be quantified into ten integer mark between one to ten. The ten and one marks state that participants think that he or she are perfectly suitable, totally unsuitable to describe this personal experiences respectively. According to the Figure 4.5 and Table 4.4, the personal experience of participants will have a slight impact on the human feelings.



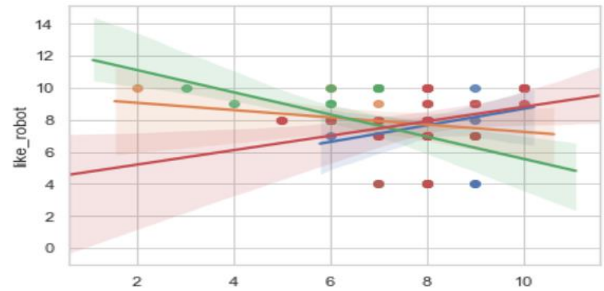
(a) Graph of person correlation test communication standard and measured factors



(b) Graph of person correlation test expert standard and measured factors



(c) Graph of person correlation test robot human-like standard and measured factors



(d) Graph of person correlation test like robot standard and measured factors

Figure 4.5. Graphs of person correlation test between personal experiences of four evaluated standards and four measured factors, where the red color, orange color, blue color, green color, represent friendly, intelligent, polite, nature four measured factors and both axes are shown the marks.

	Communication		Expert		Human-like		Like robot	
	R	P	R	P	R	P	R	P
Friendly	-0.16	0.46	-0.14	0.51	0.09	0.65	0.33	0.11
Intelligent	0.24	0.25	-0.31	0.14	-0.28	0.18	-0.19	0.37
Polite	-0.45	0.03	0.24	0.26	0.13	0.56	0.26	0.23
Natural	0.60	0.00	-0.16	0.45	-0.53	0.01	-0.51	0.01

Table 4.4 linear correlation rate (R) and possibility (P) of person correlation test between personal experiences of four evaluated standards and four measured factors

V. DISCUSSION

First of all, the results of the survey indicate that the sequence of actions do have a great impact on the friend's friendliness. The score of friendliness about speaking first and then turning head is more consistent. However, when the speaker with speaking at the same time, the level of friendliness float is greater. Besides, it also has upper confidence limit of friendliness. Surprisingly, whether you turning the head and body first then talking or speaking with turning head and body, the intelligence is the same. Moreover, the polite of the human-robot interaction experiment is not affected by the action sequence. As for the Graph of parametric tests charts, only the figure that action order and natural distribution lines are different. The blue colour and orange colour represent the T-S and T&S action can clearly find that speaking while turning head and body around is significantly higher than turning the head and body first then speaking. In other words, robot is more natural when speaking with turning head and body.

As for the background knowledge of the volunteers. Because we only have eight volunteers participated in the experiment (5 males and 3 females).The four factors include being familiar with communication with people you do not know (communication), the levels of being an expert in the Robotics (expert), the level of being familiar with human-like robot (human like) and the level of how the people like robot (like robot) have merely slight influence on the graph of person correlation test communication standard and measured factors. It is possible because our sample number do not enough, if there are as many as 20 people or more may be the result will have significant change. Besides, we also hope to balance number of men and women may be able to see the impact of gender on it.

Our experiment is very hard to find similar research in google scholar. Our research is hoped that through robot increases the anthropomorphic degree of the robot through different head and body speeds and positional silence. Perhaps most scientists believe speaking while turning head and body is the most friendly. Even though scientists are not studying depth in this way, they still offer many good ideas in other details that can improve this field. For example, the robot increases the anthropomorphic degree of the robot through different head speeds and positional silence.

Our study can improve by adding free head movement as nodding, blinking eyes and head tilt motion. Those head movement would make people feel it is special behaviour and more nature. Besides, the delay in giving a response during communication as silence can sometime view as social silence.

The reason why our research is limit as follow

1. There is a huge problem with the built-in sensory system of the robot. The speaker must respond to his eyes(cameras), which means that speaker need to be the range of cameras (180 degrees).

2. The robot can't turn around very quickly and with the rotor as the human body or easily to turn the head and body in different angles like the person.

3. Because our sample number is not enough, there may be a significant change if the experimental volunteers are more than 20 people.

4. Because the experiment is for two people at a time, when one of the experimenters talks to attract the robot. The other experimenter must look at the back of the robot and cannot recognize whether there is natural courtesy and so on.

VI. CONCLUSION

Our research is to find out which type of head movement is most polite and natural . The experimental data clearly shows that robot is more polite and natural when it turn around and speak at the same time. I believe the main contribution of our study is that we have confirmed that the majority of people are talking about turning the head by default is the most human nature is the fact.

we also make an objective investigation of human-machine dialogue, and our future research will based on this situation to investigate more details. This study results are very vital to the people who need to talk with robots face to face. Especially when the elderly face medical robots, they urgent need more realistic robots. The future directions and studies of this field as mentioned previously are mainly about free head movement as nodding, blinking eyes, head tilt motion and silence. Those four factors will be the key points to make robots be more natural, friendly, intelligent and polite.

REFERENCES

[1] Ishi, C. T., Liu, C. R., Ishiguro, H. and Hagita, N. (2010). Head motion during dialogue speech and nod timing control in humanoid robots. *Human-Robot Interaction (HRI) 2010 IEEE International Conference on*. pp. 293-300.

[2] Lee, C., Lesh, N., Sidner, C. L., Morency, L. P., Kapoor, A. and Darrell, T. (2004). Nodding in Conversations with a Robot. *Conference on Human Factors in Computing Systems: CHI04 extended abstracts on Human factors in computing systems*. pp. 785-786.

[3] Liu, C., Ishi, C. T., Ishiguro, H. and Hegita, N. (2012). Generation of nodding, head tilting and eye gazing for human-robot dialogue interaction. *Human-Robot Interaction (HRI)* 2012 7th ACM/IEEE *International Conference*. pp. 285 – 292.

[4] Oto, K., Feng, J. M. and Imai, M. (2017). Investigating how people deal with silence in a human-robot conversation. *Robot and Human Interactive Communication (RO-MAN) 2017 26th IEEE International Symposium.* pp.195 – 200.

[5] Morency, L. P., Sidner, M. C., Lee, C. and Darrell, T. (2007). Contextual Recognition of Head Gestures. *ICMI 05 Proceedings of the 7th international conference on Multimodal Interfaces* pp. 18-24.