2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)
October 25-29, 2020, Las Vegas, NV, USA (Virtual)

This or That: The Effect of Robot's Deictic Expression on User's Perception *

Dahyun Kang, Sonya S. Kwak, Hanbyeol Lee, Eun Ho Kim, and JongSuk Choi

Abstract— **The purpose of this study is to investigate a robot's impression perceived by users as well as the accuracy of perception of location information, which the robot provided according to the modality type of the robot. To explore this, we designed two 2 (verbal types: deictic vs. descriptive) x 2 (nose pointing: with nose vs. without nose) x 2 (eye pointing: with eyes vs. without eyes) mixed-participant studies. In the first study, we investigated the impacts of the robot's modality type in the imperative pointing situation. As a result, participants identified the robot's pointing gesture with nose as more effective, social, and positive, than the robot's pointing gesture without nose. Moreover, the descriptive speech robot was evaluated as more positive than the deictic speech robot. In terms of the accuracy of perception of location information, which the robot provided, participants identified the robot-designated chair more accurately when the robot delivered a deictic speech than when the robot delivered a descriptive speech. For the second study, we explored the effects of the robot's modality type in the declarative pointing situation. As a result, the robot's descriptive speech was rated as effective, social, natural, competent, trustworthy, and more positive than deictic speech. In the case of the robot's pointing gestures, pointing gesture with nose was evaluated as more effective, social, natural, competent, trustworthy, and positive than that without nose. In terms of the accuracy of location information perception, participants perceived the location of the object designated by the robot more accurately when the robot used descriptive speech, pointed with nose and without eyes.**

I. Introduction

"Look at that! (pointing to the object)"

In conversation with a person in the same space, the objects, places, and people around us frequently become the main topic, and we exchange locational information. In addition, when we share the location information, we naturally use deixis and gesture. Social robots share spaces and communicate with human users. Moreover, robots have been used to manage smart appliances. Therefore, location-specific information can be exchanged between the robots and their users. Unlike other media, such as computers and smartphones, the robot has a physical embodiment, which enables not only verbal communication, but also nonverbal communication [1]. Similar to human–human interaction, location information could be exchanged between a robot and a human through deixis speech and gesture. It is necessary to explore how the robot's modalities should be used and how robots can effectively communicate with humans by using its modalities. Therefore, this study aims to explore the impact of the robot's various modalities to express locational information on user's perception.

* This work was supported by the KIST Institutional Program (2E30280 & 2E28970). This work was also supported by the Technology Innovation Program (10077553, Development of Social Robot Intelligence for Social Human-Robot Interaction of Service Robots) funded by the Ministry of Trade, Industry & Energy (MOTIE, Korea). D. Kang is with the Center for Intelligent and Interactive Robotics, Korea Institute of Science and Technology, Seoul, 02792 Korea (e-mail: dahyun@kist.re.kr). S. S. Kwak is with the Center for Intelligent and Interactive Robotics, Korea Institute of Science and Technology, Seoul, 02792 Korea (corresponding author to provide phone: +82-2-958-6814; email: sonakwak@kist.re.kr). H. Lee is with the Center for Intelligent and Interactive Robotics, Korea Institute of Science and Technology, Seoul, 02792 Korea (e-mail: yihaanstar@kist.re.kr). E. H. Kim is with Human and Culture Convergence Technology R&BD Group, Korea Institute of Industrial Technology, Ansan-si, Gyeongi-do, Korea (e-mail: kimeunho@kitech.re.kr). J. Choi is with the Center for Intelligent and Interactive Robotics, Korea Institute of Science and Technology, Seoul, 02792 Korea (e-mail: cjs@kist.re.kr).

II. Theoretical Background

A. Locative Deixis

The conversation is a process of converting a speaker's intended idea into a linguistic form that a listener uses. The speaker uses the knowledge structure of the listener [2]. In addition, the listener seeks nouns (e.g., a demonstrative determiner, like "this" and "that") and understands what they stand for. Also, a reference can be treated as an activity of using the linguistic form to allow the listener or reader to understand.

Deixis is a frequently used item in conversation. We find personal pronouns, demonstratives, and adverbs of time and place virtually in every text [3]. Locative deixis (e.g., this, that, these, and those), which point to features of the surrounding context help to find referents and understand a speaker's utterance, and to smooth communication. In addition to using demonstratives, a pointing gesture expresses locative deixis.

B. Pointing Gesture

Pointing is one of the communication skills that an infant learns [4]. This behavior appears a few weeks before the first word is spoken. An infant communicates with both pointing gestures and words once he/she begins to say things [4]. We use deictic gestures to reduce cognitive load. It removes complicated verbal descriptors, by using the gesture to the referent [5]. In this way, pointing gestures have been used as fundamental, effective, and intuitive communication tools. This gesture is often considered a pan-culturally agreed, prototypical indexical gesture, used as a universal language [4], [6]. The prototypical pointing gesture is a body language that represents a direction from the body. This direction shows a specific location or an object [4]. The pointing gesture is even necessary during the conversation about locating referents [4].

Pointing is a novel human behavior [4]. In other words, only people communicate with each other by pointing. However, with the development of technology, a new entity,

978-1-7281-6211-9/20/$31.00 ©2020 IEEE

which is a robot, has emerged that can communicate with humans using pointing gestures. The robot is capable of nonverbal communication because it has a physical embodiment, as compared to other devices such as a computer, smart phone, and tablet PC [1]. Former studies on robot deictics reveal that describing an area of space, which is usually difficult to describe, returns only a marginally worse accuracy rate than describing an object [7]. Robots use a visual cue for spatial difference and the deictics together to confirm listeners recognize the right area of space [8].

Although human-robot interaction (HRI) studies have applied human pointing behaviors in various applications, the study of the robot's pointing gestures has been primarily limited to pointing with the finger of the robot. However, humans, also use various body parts to point, such as arms, head, eyes, lips, and nose [9], [10], [11], [12]. In addition, it is necessary to understand which robot modality, or which combination of the robot modalities, could give users the effective evidence to identify the pointed object.

C. Head Pointing

In conversations with people, real-world objects (except speakers and listeners) often become the main topic. When people refer to objects, they point with their head and hand gestures [13], [14]. This common behavior is also essential for the natural interaction between humans and robots [15].

In human-human interaction, head pointing to a speaker is considered to be a polite attitude, of listening closely, and a crucial expression, along with gestures [16], [17], [18]. HRI studies have also demonstrated that a robot's head pointing had a similar effect to a human's head pointing [19]. There have been many researchers studying the influence of the general appearance of a robot, revealing that the face of the robot plays a crucial role in its appearance [20], [21], [22], [23], [24], [25], [26]. Various current commercialized and research-purposed robots have their faces visualized on a screen because it gives vast versatility to the possible faces of the robot and creates a new design opportunity that can build a personality and perceived characteristics of the robot [27]. In this way, the face of the robot can easily be animated, such as blinking, eye gaze, and other facial expressions [27]. In addition to those benefits of the display, it is unassured that a face with a flat display would be useful for communication using the pointing gestures. The projecting shape, like a human nose, help an information receiver perceive exactly what the informer points to, because pointing is a directional gesture that points to one of several objects.

Based on these analyses, we explored how the various modalities of robots affect users' perception towards the robot in the delivery of location information.

III. Seat Guide

The robot could verbally inform the user of the location of the object. In addition, the robot's pointing gesture could indicate the location of the object. A pointing gesture is an extension of eye pointing [28]. In other words, eye pointing is the basis of a pointing gesture. Therefore, we chose verbal types and pointing gesture types including eye pointing and nose pointing as independent variables (IV) in this study. In order to investigate the effect of the robot's verbal locative

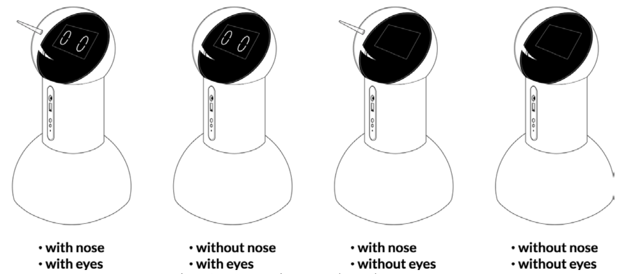

Fig. 1. Experimental Robot Types

deixis expressions, nose pointing, and eye pointing on users' perception towards the robot, and the degree of perception of the robot's direction representation, we designed a 2 (verbal types: deictic vs. descriptive) x 2 (nose pointing: with nose vs. without nose) x 2 (eye pointing: with eyes vs. without eyes) mixed-participants experiment (Fig. 1). In order to examine how effectively, socially, and naturally the robot provided location information to users according to each robot's modality, we measured the perceived effectiveness, sociability, and naturalness of the robot. In addition, we explored the effect of the location information provider robot's modalities on product evaluation to examine the overall impression towards the robot. We recruited forty-eight experiment participants aged 23 to 37 (22 males and 26 females).

A. Study Design

Task: For this experiment, the robot greeted participants in a laboratory and guided them to a seat to sit; thus, creating a situation in which the robot naturally pointed the direction to the user. The user was asked to sit in one chair, as pointed to by the robot, out of the five chairs placed in the laboratory.

Materials: Experiments were conducted with the use of Wizard of Oz (WOZ) technique, to investigate the effects of the robot's various modalities in delivering location information to users. A prototype robot was developed to implement the WOZ technique in the experiment [29].

Specification: The robot prototype could travel with direction changes based on TURTLEBOT3 [30], which is a ROS-based tele-operable mobile robot. Two active wheels and two passive wheels installed in TURTLEBOT3 permit the prototype robot to turn in place. A speaker controlled by laptop, using Bluetooth communication, and a 7" TFT LCD monitor controlled by Raspberry Pi allow the robot prototype to interact with the human with facial and verbal expressions. The movement of the robot is manually controlled via a joystick controller using Bluetooth communication (Fig. 2).

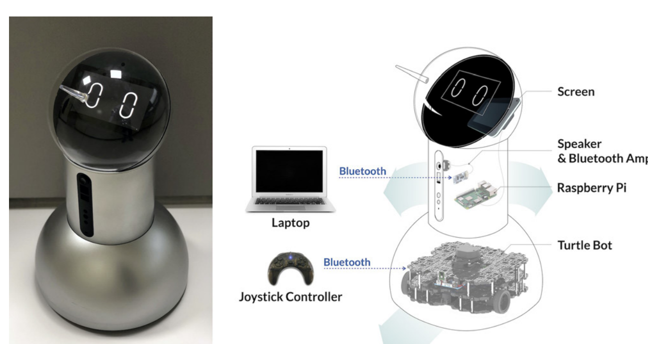

Fig. 2. Robot Used in the Experiment

Script and Gesture Design: To find out the effect of verbal types and specific pointing gesture types in the seat guide, the example of the robot's verbal and gesture expression was as follows (Fig. 3).

Experimental Setup: The participant stood facing the robot, and five chairs were set behind the participant. There was a partition behind the chairs. Experimenter 1 briefed the participant on the experiment and stood by next to the participant for emergencies. Experimenter 2 manipulated the robot's movement, and Experimenter 3 controlled the robot's utterance behind the partition (Fig. 4).

Procedure: The participants entered the laboratory, listened to the description of the experiment, and signed the experiment participation agreement. After that, they were asked to sit in a certain position by the robot. Each participant received a total of four seating requests, in random order, and after each seating, the questionnaire was administered to evaluate the impression of the robot.

Measurements: As a guide robot, we explored effectiveness [31] and sociability [32], to find out how effectively and socially the robot guided the user. In addition, the user was asked to measure naturalness [31], to find out how naturally acceptable the way of expressing the direction of the referent was perceived to be by the user. Finally, product evaluation [33], [34] was conducted to see how satisfied the user was with the robot. All the measures were rated by using seven-point Likert-type items. Furthermore, we checked if the participant sat in the chair indicated by the robot.

B. Results

We examined the effects of a guide robot's directional guidance on the users' perception of the robot. A two-way repeatedly measured analysis of variance (ANOVA) was conducted to verify the predictions (Fig. 5).

Effectiveness: There was a significant main effect of nose pointing on effectiveness ($F_{(1,46)}$=21.171, p<0.0005). The robot that pointed with nose was rated as more effective than the robot that pointed without nose ($M_{withnose}$=5.32, SD=0.15 vs. $M_{withoutnose}$=4.54, SD=0.19). The significant main effects of verbal types ($F_{(1,46)}$=0.035, p=0.852) and eye pointing ($F_{(1,46)}$=0.181, p=0.673) on effectiveness were not found.

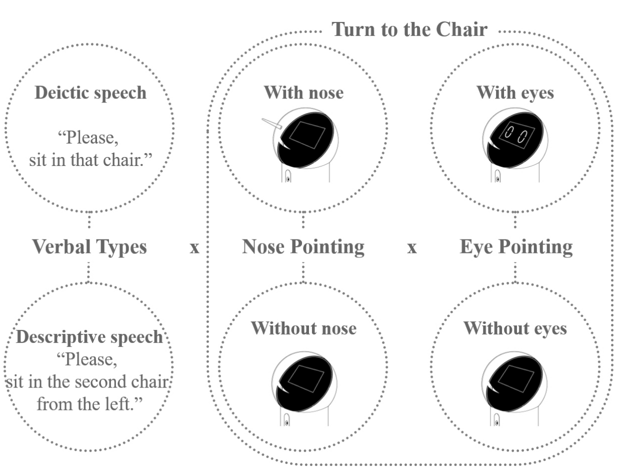

Fig. 3. Script and Gesture Design

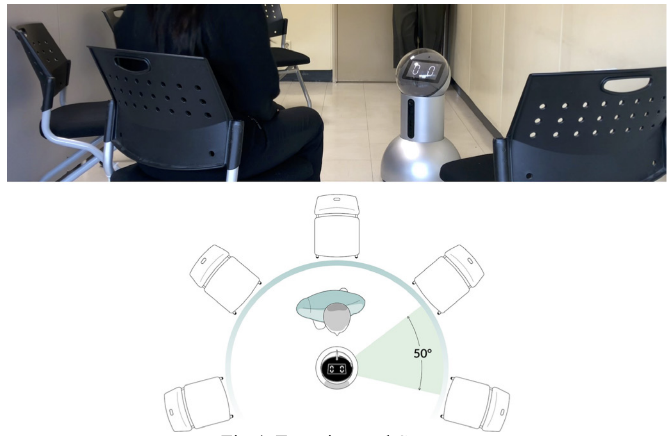

Fig.4. Experimental Setup

Sociability: There was a significant main effect of nose pointing on sociability ($F_{(1,46)}$=3.870, p=0.055). Participants evaluated the pointing gesture with nose as more social than that without nose ($M_{withnose}$=4.68, SD=0.16 vs. $M_{withoutnose}$=4.41, SD=0.17). The significant main effects of verbal types ($F_{(1,46)}$=1.535, p=0.222) and eye pointing ($F_{(1,46)}$=1.067, p=0.307) on the robot's sociability were not found.

Naturalness: There was no significant main effect of verbal types ($F_{(1,46)}$=0.096, p=0.758), nose pointing ($F_{(1,46)}$=2.161, p=0.148), and eye pointing ($F_{(1,46)}$=2.141, p=0.150) on naturalness.

Product Evaluation: There was a marginally significant main effect of verbal types on overall product evaluation ($F_{(1,46)}$=2.955, p=0.092). Participants perceived the descriptive speech robot more positively than the deictic speech robot ($M_{deictic}$=4.70, SD=0.17 vs. $M_{descriptive}$=5.00, SD=0.15). A significant main effect of nose pointing on product evaluation was also found ($F_{(1,46)}$=7.736, p=0.008). The robot's pointing gesture with nose was more positively evaluated than that without nose ($M_{withnose}$=5.01, SD=0.15 vs. $M_{withoutnose}$=4.70, SD=0.15).

Accuracy of Location Information Perception: We checked whether the user sat in the chair, exactly as guided by the robot for each condition, to examine how accurately the participants perceived the location information provided by the robot. The result is shown in Table 1.

C. Discussion

People evaluated the robot as more effective and social, and overall more positively when the robot pointed with its nose, compared to the case in which the robot pointed without nose. Moreover, they felt the descriptive speech robot was more positive than the deictic speech robot.

TABLE I. ACCURACY OF LOCATION INFORMATION PERCEPTION

IV	Verbal		Nose pointing		Eye pointing	
Condition	Deictic	Descriptive	With nose	Without nose	With eyes	Without eyes
Number of correct answers	94/96	78/96	86/96	86/96	84/96	88/96

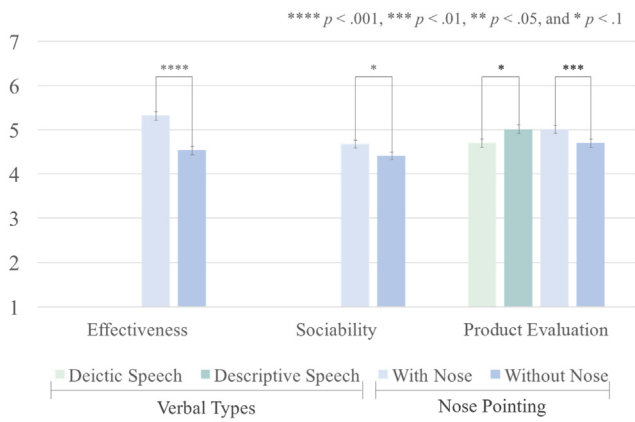

Fig. 3. Results of Seat Guide Experiment

There were several limitations to this study. First, there was a limit in the arrangement among the participant, the robot, and objects that the robot points to. As the chairs were placed around the robot, with each chair approximately 50 degrees apart, it could be easy for the participant to perceive the robot's pointing. Thus, it was necessary to conduct additional studies to find out which modality effectively informed the user of the direction that the robot pointed, in the case where the objects' placement angles were more subtle and not equally spaced. Furthermore, there was a turn-taking problem. The interaction between the robot and the participant was very short, and there was no turn-taking because the experiment was conducted to evaluate the robot after one seat guide. Finally, the pointing situation was limited. There are two types of pointing situation: imperative and declarative pointing [35]. Imperative pointing took place as an act to obtain an object [35]. Declarative pointing was expected to be held as an act to direct attention to the referred object, without desiring object retrieval [35]. This study was limited to the imperative pointing situation. The limitation of the first study led to the second study.

IV. Exhibition Guide

This experiment was designed to explore the effect of the robot's location information delivery modalities for objects placed at various angles in the declarative pointing situation with multiple turn-takings. We conducted a 2 (verbal types: deictic vs. descriptive) x 2 (nose pointing: with nose vs. without nose) x 2 (eye pointing: with eyes vs. without eyes) mixed-participant experiment. Forty-eight experiment participants were aged 23 to 37 (22 males and 26 females). Participants had bachelor's degrees to reduce errors due to differences in understanding of exhibition contents.

A. Study Design

Task: Participants were told about the robots on display at the robot showroom by a guide robot. As the guide robot explained the details of the robots, we asked the participants to briefly write the names and corresponding details of the robots on a memo pad, in the order the guide robot told them. Five robots were exhibited in the robot showroom. In each condition, the guide robot explained two characteristics for each exhibited robot. That is, there were 10 turn-takings for each condition, and the guide robot was evaluated after 10 turn-takings. The participant experienced 40 turn-takings

TABLE II. The Example of Script and Gesture Design

Verbal types	Speech	Gesture
Deictic Speech	"Hello. I'm GuideBot. Let's me explain the robots on display."	a turn to the participant
	"**This robot** is capable of dynamic walking based on real-time forward walking pattern generation technology. It has been used to develop the mechanism design and control algorithm for stable walking."	a turn to Robot A
	"May I explain the next robot?"	5 seconds later, a turn to the participant
	"Let me explain the next robot."	
	"**This robot** is a robot that can efficiently provide services using both arms and hands in home environment. 3D object recognition and vision-based control technology is applied."	a turn to Robot E
Descriptive Speech	"Hello. I'm GuideBot. Let's me explain the robots on display."	a turn to the participant
	"**The first robot from the left** is capable of dynamic walking based on real-time forward walking pattern generation technology. It has been used to develop the mechanism design and control algorithm for stable walking."	a turn to Robot A
	"May I explain the next robot?"	5 seconds later, a turn to the participant
	"Let me explain the next robot."	
	"**The first robot from the right** is a robot that can efficiently provide services using both arms and hands in home environment. 3D object recognition and vision-based control technology is applied."	a turn to Robot E

because each participant experienced a total of four conditions.

Materials: The same robot as the first study was used in this study.

Script and Gesture Design: To find out the effect of verbal types and specific nose pointing types in the exhibition guide, the example of the robot's verbal and gesture expressions are as follows (See Table 2).

Experimental Setup: The participant stood next to the guide robot and looked at the robots displayed. Five robots were displayed in front of the participant and the guide robot, and the robots were named Robots A, B, C, D, and E, respectively. Experimenter 1 briefed the participant on the experiment. Experimenter 2 and 3 controlled the robot in the room where they could observe the exhibition space (Fig. 6).

Procedure: The participant entered the laboratory, listened to the description of the experiment, and signed the experiment participation agreement. After that, he/she was

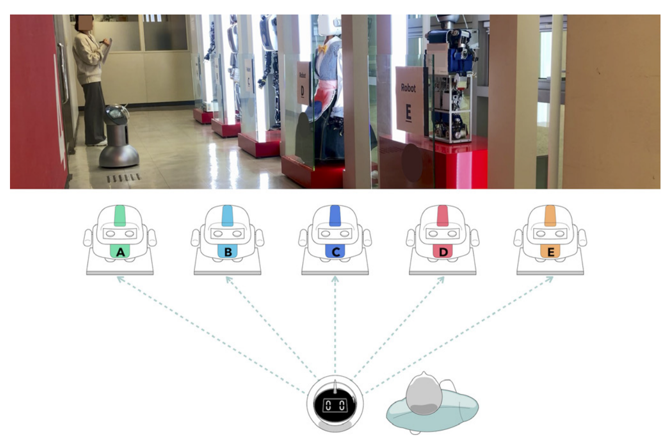

Fig. 4. Experimental Setup

asked to move to the robot showroom and write the name of the robots and the corresponding details, in the order the guide robot explained. After the participant experienced each condition, the questionnaire was administered to evaluate the impression of the robot.

Measurements: For the guide robot, we explored effectiveness [31], sociability [32], naturalness [31], and product evaluation [33], [34] in the same method as the first study. In the case of the exhibition guide task, the guide robot provided the user with professional information about displayed robots. Therefore, we asked the participant to evaluate competency [32] and trustworthiness [32], in order to find out how competently the robot could convey information to the users and whether the users could trust the information given by the guide robot. All measurement items were rated by using seven-point Likert-type items. In addition, in order to investigate how accurately people perceived the location information that the guide robot conveyed, the accuracy of deictic information perception was measured by comparing the order of the robot names they wrote on the memo pad and the order of the robot names the robots described. For each condition, participants were asked to list the total of ten robot names, because five robots were described twice in random.

B. Results

We examined the effects of a guide robot's directional guidance on the users' perception towards the robot. A two-way repeatedly measured ANOVA was conducted to test the predictions (Fig. 7).

Effectiveness: The significant main effects of verbal types on effectiveness was found ($F_{(1,46)}$=94.242, p<0.0005). People felt that the descriptive speech robot was more effective than the deictic speech robot (M_{deictic}=4.29, SD=0.17 vs. $M_{\text{descriptive}}$=5.63, SD=0.12). A significant main effect of nose pointing on effectiveness was found ($F_{(1,46)}$=53.908, p<0.0005). The robot that pointed with nose was rated as more effective than the robot that pointed without nose (M_{withnose}=5.43, SD=0.15 vs. $M_{\text{withoutnose}}$=4.48, SD=0.14). A significant main effect of eye pointing ($F_{(1,46)}$=0.532, p=0.469) on effectiveness was not found.

Sociability: There was a significant main effect of verbal types on sociability ($F_{(1,46)}$=54.904, p<0.0005). Participants perceived that the descriptive speech robot was more sociable than the deictic speech robot (M_{deictic}=4.33, SD=0.13 vs. $M_{\text{descriptive}}$=4.99, SD=0.14). In addition, a significant main effect of nose pointing on sociability was found ($F_{(1,46)}$=8.595, p=0.005). Participants evaluated the robot's pointing with nose as more social than that without nose (M_{withnose}=4.82, SD=0.14 vs. $M_{\text{withoutnose}}$=4.50, SD=0.13). A significant main effect of eye pointing on robot's sociability was not found ($F_{(1,46)}$=0.204, p=0.654).

There was a significant interaction effect between verbal types and nose pointing ($F_{(1,46)}$=4.213, p=0.046) on sociability. When the guide robot spoke in the deictic way, the robot's pointing with nose was considered more social than that without nose (M_{withnose}=4.58, SD=1.03 vs. $M_{\text{withoutnose}}$=4.08, SD=1.05, t=3.500, df=47, p=0.001). Conversely, when the guide robot spoke in the descriptive way, there was no significant difference by nose pointing (t=1.074, df=47, p=0.288). This indicates that in order to increase sociability of a guide robot, the descriptive speech robot could be designed irrespective of nose pointing types while the deictic speech robot needs to be designed with nose pointing.

Naturalness: The significant main effects of verbal types on the robot's naturalness was found ($F_{(1,46)}$=35.353, p<0.0005). People felt that the descriptive speech robot was more natural than the deictic speech robot (M_{deictic}=3.95, SD=0.15 vs. $M_{\text{descriptive}}$=4.73, SD=0.16). A significant main effect of nose pointing on naturalness was found ($F_{(1,46)}$=35.160, p<0.0005). The robot that pointed with nose was rated as more natural than the robot that pointed without nose (M_{withnose}=4.68, SD=0.16 vs. $M_{\text{withoutnose}}$=4.00, SD=0.15). A significant main effect of eye pointing ($F_{(1,46)}$=1.379, p=0.246) on naturalness was not found.

There was a marginally significant interaction effect between nose pointing and eye pointing ($F_{(1,46)}$=2.914, p=0.095) on naturalness. During interaction with the guide robot with eye pointing, the robot's pointing with nose was perceived as more natural than that without nose (M_{withnose}=4.61, SD=1.35 vs. $M_{\text{withoutnose}}$=3.73, SD=1.41, t=5.770, df=47, p=0.001). In addition, during interaction with the guide robot without eye pointing, the robot's pointing with nose was also considered as more natural than that without nose (M_{withnose}=4.75, SD=1.12 vs. $M_{\text{withoutnose}}$=4.26, SD=1.12, t=3.264, df=47, p=0.004). The robot with nose was measured as more natural than that without nose, for all eye pointing types; however, the differences in naturalness according to the nose pointing were more prominent for the robot with eye pointing.

Competency: There was a significant main effect of verbal types on the competency of the guide robot ($F_{(1,46)}$=18.574, p<0.0005). Participants perceived that the descriptive speech robot was more competent than the deictic speech robot (M_{deictic}=4.37, SD=0.19 vs. $M_{\text{descriptive}}$=4.87, SD=0.17). In addition, a significant main effect of nose pointing on competency was found ($F_{(1,46)}$=8.184, p=0.006). Participants evaluated the robot's pointing with nose as more competent than that without nose (M_{withnose}=4.76, SD=0.17 vs. $M_{\text{withoutnose}}$=4.48, SD=0.18). The significant main effect of eye pointing on the robot's competency was not found ($F_{(1,46)}$=0.378, p=0.542).

Trustworthiness: There was a significant main effect of verbal types on the trustworthiness of the guide robot ($F_{(1,46)}$=43.864, p<0.0005). The descriptive speech robot was perceived as more trustworthy than the deictic speech robot (M_{deictic}=4.53, SD=0.17 vs. $M_{\text{descriptive}}$=5.28, SD=0.15). A significant main effect of nose pointing on trustworthiness was found ($F_{(1,46)}$=8.799, p=0.005). Participants felt that the robot's pointing with nose was more trustworthy than that without nose (M_{withnose}=5.09, SD=0.17 vs. $M_{\text{withoutnose}}$=4.72, SD=0.16). The significant main effect of eye pointing on the robot's trustworthiness was not found ($F_{(1,46)}$=0.126, p=0.724).

Product Evaluation: There was a significant main effect of verbal types on the product evaluation ($F_{(1,46)}$=56.358, p<0.0005). Participants rated the descriptive speech robot more positively than the deictic speech robot ($M_{deictic}$=4.44, SD=0.14 vs. $M_{descriptive}$=5.14, SD=0.12). A significant main effect of nose pointing on product evaluation was also found ($F_{(1,46)}$=14.507, p<0.0005). The robot's pointing with nose was more positively evaluated than that without nose ($M_{withnose}$=5.00, SD=0.14 vs. $M_{withoutnose}$=4.59, SD=0.13).

There was an interaction effect between verbal types and eye pointing ($F_{(1,46)}$=4.150, p=0.047) on product evaluation. When the robot was pointing with eyes, participants evaluated the descriptive speech robot as being more positive than the deictic speech robot ($M_{deictic}$=4.31, SD=1.25 vs. $M_{descriptive}$=5.20, SD=1.04, t=5.895, df=47, p=0.0005). Furthermore, when the robot was pointing without eyes, the descriptive speech robot was evaluated as more positive than the deictic speech robot ($M_{deictic}$=4.57, SD=0.96 vs. $M_{descriptive}$=5.08, SD=0.86, t=5.002, df=47, p=0.0005). The descriptive speech robot was evaluated more positively than the deictic speech robot for all eye pointing types. However, the differences in product evaluation by verbal types were more dominant for the robot with eye pointing.

Accuracy of Location Information Perception: There were significant main effects on the accuracy of location information perception by participants according to verbal types ($F_{(1,46)}$=8.452, p=0.006), nose pointing ($F_{(1,46)}$=4.699, p=0.035), and eye pointing ($F_{(1,46)}$=7.614, p=0.008). Participants perceived the position of the displayed robot more accurately when the guide robot referred to the displayed robot in a descriptive manner, rather than when the guide robot referred to the displayed robot in a deictic manner ($M_{deictic}$=9.58, SD=0.11 vs. $M_{descriptive}$=9.91, SD=0.04). In addition, they perceived the displayed robot's position more accurately when the robot was pointing with nose, rather than when the robot was pointing without nose ($M_{withnose}$=9.88, SD=0.06 vs. $M_{withoutnose}$=9.62, SD=0.11). When interacting with a robot without eye pointing, participants more accurately perceived what the robot referred to, as compared to a robot with eye pointing ($M_{witheyes}$=9.58, SD=0.08 vs. $M_{withouteyes}$=9.91, SD=0.08).

There was an interaction effect between verbal types and nose pointing on the accuracy of location information perception ($F_{(1,46)}$=5.914, p=0.019). When the guide robot referred to displayed robots with a deictic speech, participants perceived the position of the displayed robot, which the guide robot with nose pointing explained more correctly than the guide robot without nose pointing explained ($M_{withnose}$=9.85, SD=0.62 vs. $M_{withoutnose}$=9.31, SD=1.45, t=2.349, df=47, p=0.023). On the contrary, when the guide robot referred to displayed robots with a descriptive speech, no significant effect was found by nose pointing (t= 0.443, df=47, p=0.659). This result means that when the robot uses descriptive speech, the robot could be designed regardless of nose pointing types. In contrast, when the robot uses deictic speech, the robot should be designed in consideration of nose pointing types.

V. General Discussion and Conclusion

A. Overview and Interpretation of Results

First, regarding the impression evaluation of the robot, both the descriptive verbal type and pointing with nose were more positively evaluated for both the situation of seat instruction and explaining a specific object as shown in Table 3. This result implies that the user feels that the robot that gives location information by using descriptive speech and nose pointing is better at acting as a guide robot, regardless of whether it is an imperative pointing situation or a declarative pointing situation.

Unlike the impression evaluation of robots, accuracy of location information perception by participants differed depending on the pointing situation. In the imperative pointing situation (seat guide experiment), the verbal types considerably affected the accuracy of location information perception (percentage of correct answers: deictic 98% and descriptive 81%). The deictic speech robot referred to the chairs using the words "this" and "that", whereas the descriptive speech robot used the direction indicators "left" and "right", to refer to the chair, such as the third chair from the left. Moreover, the "left" and "right" directions could be interpreted in two referent ways: a) the guide robot (speaker) and b) the participant (listener). When the speaker and listener faced each other, the referent points would be opposite. When

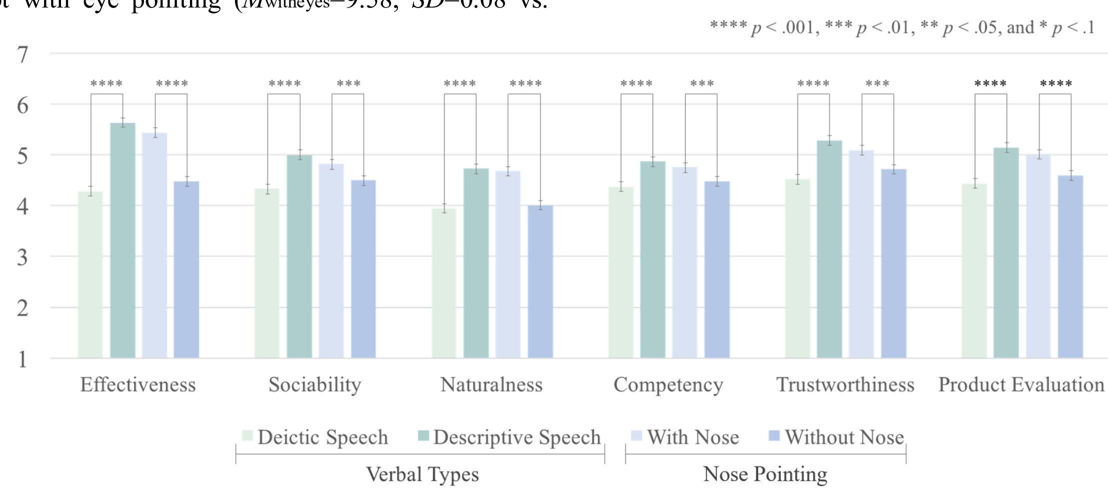

Fig. 5. Results of Exhibition Guide Experiment

TABLE III. SUMMARY

	Seat Guide						Exhibition Guide					
	Verbal Type		Nose Pointing		Eye Pointing		Verbal Type		Nose Pointing		Eye Pointing	
	Deictic	Descriptive	With Nose	Without Nose	With Eyes	Without Eyes	Deictic	Descriptive	With Nose	Without Nose	With Eyes	Without Eyes
Effectiveness	-	-	5.32 (0.15)	**** 4.54 (0.19)	-	-	4.29 (0.17)	**** 5.63 (0.12)	5.43 (0.15)	**** 4.48 (0.14)	-	-
Sociability	-	-	4.68 (0.16)	* 4.41 (0.17)	-	-	4.33 (0.13)	**** 4.99 (0.14)	4.82 (0.14)	*** 4.50 (0.13)	-	-
Naturalness	-	-	-	-	-	-	3.95 (0.15)	**** 4.73 (0.16)	4.68 (0.16)	**** 4.00 (0.15)	-	-
Competency							4.37 (0.19)	**** 4.87 (0.17)	4.76 (0.17)	*** 4.48 (0.18)	-	-
Trustworthiness							4.53 (0.17)	**** 5.28 (0.15)	5.09 (0.17)	*** 4.72 (0.16)	-	-
Product Evaluation	4.70 (0.17)	* 5.00 (0.15)	5.01 (0.15)	*** 4.70 (0.15)	-	-	4.44 (0.14)	**** 5.14 (0.12)	5.00 (0.14)	**** 4.59 (0.13)	-	-

Mean (Standard Deviation), **** $p < .001$, *** $p < .01$, ** $p < .05$, and * $p < .1$

asking participants who faced the robot to sit down, participants sitting in the wrong chair perceived the right and left sides based not on the robot (speaker), but rather based on themselves (listener). Although the chairs were placed at equal interval angles in the laboratory and all the utterances of the robot were made with pointing to a specific chair, in the context of descriptive speech, few participants ignored the robot's nose pointing and took the location information from their perspective. However, the participants and the robot stood abreast in the declarative pointing situation (exhibition guide experiment). Therefore, the referent points for the speaker and listener were identical. Consequently, the location information from the descriptive speech robot was more accurately perceived by the participants than that from the deictic speech robot, indicating that both speech types could be effectively applied to the guide robot, according to the speaker-listener referent points.

B. Limitations

First of all, the environment setting might affect the users' perception and evaluation. Only a single environmental setup which the distance between the user and the robot is close was used for each experiment. A guide robot would be possible to explain different types of objects in various spaces. This study could be more persuasive if we would conduct in a more diverse and natural experimental setting considering the distance between the user and the robot.

Second, the accuracy of location information perception was measured to figure out how effectively the robot delivered locational information in this study. Besides the accuracy of location information perception, if the amount of time it took a participant to find the chair would be measured, it would be able to fully evaluate the effectiveness of using verbal, pointing, and gaze.

C. Contributions

This study has contributions in three aspects.

First, our findings in this study could contribute to human social interaction studies. Although this study was designed based on ideas from human studies, e.g., anthropology and sociology, the results of this study could also provide suggestions to human studies. In human studies, the effect of the nose on head pointing has not been revealed because people cannot attach or detach their noses. However, we could figure out the effect of the nose on head pointing because our study dealt with an artificial agent, a robot.

Second, the results and interpretation contribute to the design of the robot's speech style. When the robot uttered descriptively, the referent point influenced the participants' perception of the location information indicated by the guide robot. This means that the robot should either inform the referent point to the user (e.g., the third from the left of me) or match the referent point with the user when the robot gives the user the location information. Moreover, the results of the study imply that information should be provided in the deictic speech type when referent points between the robot and the user are difficult to match.

Third, our study contributes to the improvement of robot design. The results of the robot impression evaluation showed that the descriptive speech robot was overall more positive than the deictic speech robot; however, in terms of the accuracy of location information perception by the user, the deictic speech robot was more effective than the descriptive speech robot when the robot and the person had different referent points. For tasks where the information provided by the robot does not have a critical effect on the user, it is suggested that the robot be designed to give users location information descriptively, rather than deictically. On the contrary, we suggest that for tasks where information misperception leads to critical damage, the robot should be designed to utter differently depending on the referent point.

ACKNOWLEDGMENT

This work was supported by the KIST Institutional Program (2E30280). This work was also supported by the Technology Innovation Program (10077553, Development of Social Robot Intelligence for Social Human-Robot Interaction of Service Robots) funded by the Ministry of Trade, Industry Energy (MOTIE, Korea).

REFERENCES

[1] T. Fong, I. Nourbakhsh, and K. Dautenhahn, "A survey of socially interactive robots," *Robotics and Autonomous Systems*, vol. 42, no. 34, pp. 143–166, 2003.

[2] J. Rubba, "Alternate grounds in the interpretation of deictic expressions," in *Spaces, Worlds, and Grammar*, 1st ed., G. Fauconnier and E. Sweetser, Eds. Chicago, IL: University of Chicago Press, 1996, pp. 227–261.

[3] S. C. Herring, "Discourse functions of demonstrative deixis in tamil," in *Proc. of the 20th Annual Meeting of the Berkeley Linguistics Society*, Berkeley, CA, February 1994, pp. 246–259.

[4] S. Kita, "Pointing: A foundational building block of human communication." in *Pointing: Where Language, Culture, and Cognition Meet*, 1st ed., S. Kita, Ed. Mahwah, NJ: Lawrence Erlbaum Associates, 2003, pp. 171–215.

[5] S. Goldin-Meadow, "The role of gesture in communication and thinking," *Trends in Cognitive Sciences*, vol. 3, no. 11, pp. 419–429, 1999.

[6] C. M. Murphy, "Pointing in the context of a shared activity," *Child Development*, vol. 49, no. 2, pp. 371–380, 1978.

[7] A. S. Clair, R. Mead, and M. J. Mataric, "Investigating the effects of´ visual saliency on deictic gesture production by a humanoid robot," in *Proc. of the 20th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*, Atlanta, GA, August 2011, pp. 210–216.

[8] Y. Hato, S. Satake, T. Kanda, M. Imai, and N. Hagita, "Pointing to space: Modeling of deictic interaction referring to regions," in *Proc. of the 5th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, Osaka, Japan, March 2010, pp. 301–308.

[9] E. McClave, H. Kim, R. Tamer, and M. Mileff, "Head movements in the context of speech in arabic, bulgarian, korean, and africanamerican vernacular english," *Gesture*, vol. 7, no. 3, pp. 343–390, 2007.

[10] N. J. Enfield, "'lip-pointing' : A discussion of form and function with reference to data from laos," *Gesture*, vol. 1, no. 2, pp. 185–211, 2001.

[11] D. Wilkins, "Why pointing with the index finger is not a universal (in sociocultural and semiotic terms)," in *Pointing: Where Language, Culture, and Cognition Meet*, 1st ed., S. Kita, Ed. Mahwah, NJ: Lawrence Erlbaum Associates, 2003, pp. 171–215.

[12] K. Cooperrider and R. Nu´nez, "Nose-pointing: Notes on a facial˜ gesture of papua new guinea," *Gesture*, vol. 12, no. 2, pp. 103–129, 2012.

[13] C. Goodwin, "Human sociality as mutual orientation in a rich interactive environment: Multimodal utterances and pointing in aphasia," in *Roots of Human Sociality: Culture, Cognition and Interaction*, 1st ed., N. J. Enfield and S. C. Levinson, Eds. New York, NY: Berg, 2006, pp. 97–125.

[14] S. Goldin-Meadow, C. Mylander, J. de Villiers, E. Bates, and V. Volterra, "Gestural communication in deaf children: The effects and noneffects of parental input on early language development," *Monographs of the Society for Research in Child Development*, vol. 49, no. 3/4, pp. 1–151, 1984.

[15] M. Imai, T. Kanda, T. Ono, H. Ishiguro, and K. Mase, "Robot mediated round table: Analysis of the effect of robot's gaze," in *Proc. of the 11th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*, Berlin, Germany, August 2002, pp. 411–416.

[16] H. M. Taylor, "Japanese kinesics," *The Journal of the Association of Teachers of Japanese*, vol. 9, no. 1, pp. 65–75, 1974.

[17] S. R. Langton, "The mutual influence of gaze and head orientation in the analysis of social attention direction," *The Quarterly Journal of Experimental Psychology: Section A*, vol. 53, no. 3, pp. 825–845, 2000.

[18] K. Yamazaki, M. Kawashima, Y. Kuno, N. Akiya, M. Burdelski, A. Yamazaki, and H. Kuzuoka, "Prior-to-request and request behaviors within elderly day care: Implications for developing service robots for use in multiparty settings," in *Proc. of the 10th European Conference on Computer-Supported Cooperative Work (ECSCW)*, Limerick, Ireland, September 2007, pp. 61–78.

[19] C. Chen and J.-M. Odobez, "We are not contortionists: Coupled adaptive learning for head and body orientation estimation in surveillance video," in *Proc. of the 2012 IEEE Conference on Computer Vision and Pattern Recognition*, Providence, RI, June 2012, pp. 1544–1551.

[20] C. Breazeal and B. Scassellati, "How to build robots that make friends and influence people," in *Proc. of the 1999 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, Kyongju, Korea, October 1999, pp. 858–863.

[21] K. Dautenhahn, S. Woods, C. Kaouri, M. L. Walters, K. L. Koay, and I. Werry, "What is a robot companion-friend, assistant or butler?" in *Proc. of the 2005 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, Edmonton, Canada, August 2005, pp. 1192–1197.

[22] C. Diana and A. L. Thomaz, "The shape of simon: Creative design of a humanoid robot shell," in *Proc. of the 2011 CHI Conference on Human Factors in Computing Systems, Extended Abstracts*, Vancouver, Canada, May 2011, pp. 283–298.

[23] D. Li, P. P. Rau, and Y. Li, "A cross-cultural study: Effect of robot appearance and task," *International Journal of Social Robotics*, vol. 2, no. 2, pp. 175–186, 2010.

[24] M. Lohse, F. Hegel, A. Swadzba, K. Rohlfing, S. Wachsmuth, and B. Wrede, "What can i do for you? appearance and application of robots," in *Symposium at AISB'07: The Reign of Catz and Dogz? The Role of Virtual Creatures in a Computerised Society*, Newcastle, UK, 2007, pp. 121–126.

[25] D. S. Syrdal, K. Dautenhahn, S. N. Woods, M. L. Walters, and K. L. Koay, "Looking good? appearance preferences and robot personality inferences at zero acquaintance." in *AAAI Spring Symposium: Multidisciplinary Collaboration for Socially Assistive Robotics*, Palo Alto, CA, 2007, pp. 86–92.

[26] M. L. Walters, K. L. Koay, D. S. Syrdal, K. Dautenhahn, and R. te Boekhorst, "Preferences and perceptions of robot appearance and embodiment in human-robot interaction trials," in *Symposium at AISB'09: New Frontiers in Human-Robot Interaction*, Edinburgh, Scotland, 2009.

[27] A. Kalegina, G. Schroeder, A. Allchin, K. Berlin, and M. Cakmak, "Characterizing the design space of rendered robot faces," in *Proc. of the 13th Annual ACM/IEEE International Conference on HumanRobot Interaction (HRI)*, Chicago, IL, March 2018, pp. 96–104.

[28] T. Bejarano, *Becoming Human: From Pointing Gestures to Syntax*. Amsterdam, The Netherlands: John Benjamins Publishing, 2011.

[29] L. D. Riek, "Wizard of oz studies in hri: A systematic review and new reporting guidelines," *Journal of Human-Robot Interaction*, vol. 1, no. 1, pp. 119–136, 2012.

[30] (2020) Turtlebot3. ROBOTIS. Retrieved February 28, 2020 from https://www.turtlebot.com/.

[31] A. Sauppe and B. Mutlu, "Robot deictics: How gesture and con-´ text shape referential communication," in *Proc. of the 9th Annual ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, Bielefeld, Germany, March 2014, pp. 342–349.

[32] S. Andrist, E. Spannan, and B. Mutlu, "Rhetorical robots: Making robots more effective speakers using linguistic cues of expertise," in *Proc. of the 8th Annual ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, Tokyo, Japan, March 2013, pp. 341–348.

[33] M. Herzenstein, S. S. Posavac, and J. J. Brakus, "Adoption of new and really new products: The effects of self-regulation systems and risk salience," *Journal of Marketing Research*, vol. 44, no. 2, pp. 251–260, 2007.

[34] M. Zhao, S. Hoeffler, and D. W. Dahl, "The role of imaginationfocused visualization on new product evaluation," *Journal of Marketing Research*, vol. 46, no. 1, pp. 46–55, 2009.

[35] U. Liszkowski, "Infant pointing at 12 months: Communicative goals, motives, and social-cognitive abilities," in *Roots of Human Sociality: Culture, Cognition and Interaction*, 1st ed., N. J. Enfield and S. C. Levinson, Eds. New York, NY: Berg, 2006, pp. 153–178.