Personality Shifting Agent to Remove Users' Negative Impression on Speech Recognition Failure

Tatsuki Hori[1] and Kazuki Kobayashi[2]
[1]*Graduate School of Science and Technology, Shinshu University, Nagano 380-8553, Japan*
[2]*Academic Assembly, Shinshu University, Nagano 380-8553, Japan*

Keywords: Conversational Agents, Experiment, Voice Interfaces, User Experiences.

Abstract: In this study, we propose a method to shift an agent's personality during speech interaction to reduce users' negative impressions of speech recognition systems when speech recognition fails. Although spoken dialog interfaces, such as smart speakers, have emerged to support our daily lives and the accuracy of speech recognition has improved, users are burden with rephrasing commands for these systems because they fail. Speech recognition failure makes users uncomfortable, and the cognitive strain in rephrasing commands is high. The proposed method aims to eliminate users' negative impression of agents by allowing an agent to have multiple personalities and accept responsibility for the failure, with the personality responsible for failure being removed from the task. System hardware remains the same, and users can continue to interact with another personality of the agent. Shifting the agent's personality is represented by a change in voice tone and LED color. Experimental results with 20 participants suggested that the proposed method reduces users' negative impressions by improving communication between users and the agent, as well as the agent's sense of responsibility, and that users felt that the agent have emotions.

1 INTRODUCTION

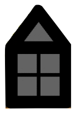
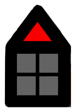
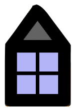
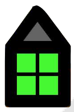
In recent years, spoken dialog technology has been widely used and is becoming a part of people's daily lives (Guamán et al., 2018). Smart assistants based on speech interaction technologies such as Google Assistant, Amazon Alexa, and Apple's Siri have been developed and installed on smartphones, smart speakers, cars, and so on. A survey on smart speaker penetration (Philpott, 2018) rates reported that 21% of homes in the US have smart speakers, and in the four years since the release of the first smart speaker, Amazon Echo, smart speakers have reached a level of adoption compared to other smart home devices such as security webcams and smart thermostats.

The research field of speech interaction includes speech synthesis (Hojo et al., 2018) (Bulut et al., 2002), the implementation of dialog models (Yamamoto et al., 2018), and speech recognition (Kato et al., 2008) (Le et al., 2019), all of which are primarily focused on improving the experience of natural interaction between users and devices. Although speech recognition techniques have progressed through various studies, it is difficult to completely prevent speech recognition failures because of noise in complex real-world environments, different accents, user speech problems (stuttering, word swallowing, etc), an excessively narrow database (in its vocabulary), and so on. When speech recognition fails, users have a negative impression of the speech recognition device, and it would have a negative impact on its continuous use. It is necessary to appropriately remove negative impressions of users when recognition fails because humans are more persistent in negative impressions than in positive impressions and are less likely to override negative impressions (Fiske, 1980). As a method to prevent users from having negative impressions of speech recognition devices, a method to present the internal state of a robot (Breazeal et al., 2005) (Komatsu et al., 2018) and a method to apologize when the robot fails (Engelhardt et al., 2017) have been proposed. However, other than apologies, no method for accepting responsibility through specific actions has been proposed. The ability of a system to perform its responsibilities is an important part of human-system interaction (Çürüklü et al., 2010) and is a complex issue that requires a contextual understanding of the interaction (Webb et al., 2019).

This paper proposes a speech interaction agent with multiple personalities within a single device to

Hori, T. and Kobayashi, K.
Personality Shifting Agent to Remove Users' Negative Impression on Speech Recognition Failure.
DOI: 10.5220/0010687400003060
In *Proceedings of the 5th International Conference on Computer-Human Interaction Research and Applications (CHIRA 2021)*, pages 181-188
ISBN: 978-989-758-538-8; ISSN: 2184-3244

Table 1: Agent's appearances and behavior in each state.

State	Appearance	Personality	Behavior
Standby		All	The upper and bottom parts remain gray colored in each personality
Listening		All	The upper part blinks red at 30 fps in each personality
Speaking		Personality A_1	The bottom part blinks blue at 30 fps
Speaking		Personality A_2	The bottom part blinks yellow at 30 fps
Speaking		Personality A_3	The bottom part blinks green at 30 fps

properly handle the responsibility of failure. In the proposed method, when speech recognition fails, the personality of the agent is replaced with another one. Although speech recognition failure leaves a user with a negative impression of the system, it is possible to prevent users' impression from deteriorating because by replacing the agent's personality with another personality, the agent accepts responsibility for the failure. The proposed method can display a scene in which the current personality of the agent is changed to another personality during speech recognition failure, with an indication of the agent's internal state it and an apology as a specific action to accept responsibility. As interacting with multiple agents has a high cognitive strain (Yoshikawa et al., 2017) (Nishimura et al., 2013), our proposed method allows only one agent to respond at a time, which has the advantage of preventing confusion during speech interaction.

2 PERSONALITY SHIFTING AGENT

The proposed personality agent is a speech interaction agent that interacts with a user by representing different personalities by combining factors such as the agent's appearance color and voice tone.

2.1 Agent Behavior

The agent's personalities are implemented in a single device, and one personality is expressed at a time. The expressed agent's personality is replaced with another one when speech recognition fails. Speech recognition failures often leave users with a negative impression. However, personality shifting by the agent can suppress users' negative impression because the current personality accepts the responsibility for the failure and is dismissed from the task. Each personality expresses a combination of factors such as the agent's appearance color and voice tone and speaks a common predetermined content. The personality names are not presented to users and the agent always starts the personality change with "Next is my turn."

Table 1 shows the appearances and behavior of the agent in each state. The agent has three personalities and states, respectively, and expresses them by changing its color. The states are standby, listening, and speaking, and the personalities are A_1, A_2, and A_3. In the standby state, the color of the agent is gray because no color is changed. In the listening state, the upper part of the agent blinks (30fps) red according to a user's speech input. In the speech speaking state,

Table 2: Example of dialog between user and agent.

Recognition	Dialog
Success	User: Turn off the TV. Agent: Yes, I'll turn off the TV... It's off.
Failure	User: Turn off the air cleaner. Agent: I'm sorry. I did not catch that. (Then Shifting Personalities) Agent: Next is my turn. Your orders, please. User: Turn off the air cleaner.

Table 3: Audio settings of agent personalities.

Parameter	Personality A_1	A_2	A_3
Vocaloid Name	Takahashi	Tsudsumi Suzuki	Sasara Satoh
Gender	Male	Female	Female
Volume	0.00	0.00	0.00
Speed	1.00	1.21	1.00
Pitch	0	0	0
Quality	0.00	0.00	0.00
Intonation	1.00	1.00	1.00
Special features	Energetic: 0.00 Normality: 1.00 Depressed: 0.00	Cool: 0.47 Embarrassed: 0.53	Energetic: 1.00 Normality: 0.00 Anger: 0.00 Sadness: 0.00

the bottom part of the agent blinks (30fps) the color of the personality, i.e., blue, yellow, or green, according to the agent's speech. The color change represents the internal state of the agent, to avoid speech collisions, and expresses the differences in the personalities of the agent.

Table 2 shows examples of successful and unsuccessful speech recognition of dialog between users and the agent. Personality is shifted only when speech recognition fails, and the next personality appears and waits for commands. Personality shifting is cycled in one direction, $A_1 \rightarrow A_2$, $A_2 \rightarrow A_3$, and $A_3 \rightarrow A_1$. Personality shifting is immediately after the agent informs users of its failure and apologizes. Each personality is expressed through synthetic speech sounds generated by the speech synthesis software CeVIO Creative Studio Ver. 6.1, which includes three voices: Takahashi, Suzuki Tsudumi, and Sato Sasara. Table 3 shows speech sound settings for A_1, A_2, and A_3.

3 EXPERIMENTS

The experiment aims to evaluate users' impression of the personality shifting agent. The effectiveness of the proposed method is analyzed using the agent's personality shifting factors as independent variables and users' impression as a dependent variable.

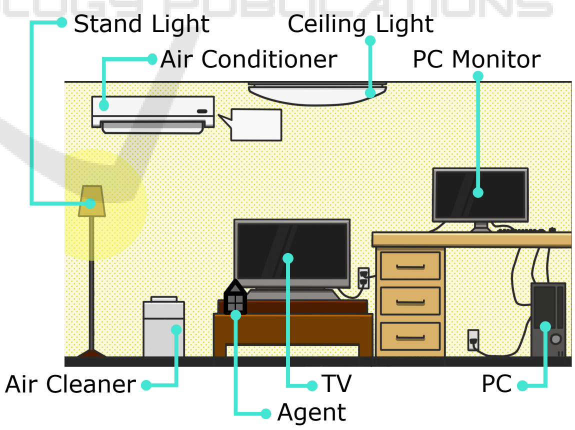


Figure 1: Voice control home appliance simulator.

3.1 Voice Control Home Appliance Simulator

Figure 1 shows the developed voice control simulator for simulating home appliance operation through voice commands. The simulator operates home appliances based on voice commands from a user and visu-

Table 4: Voice command list.

No.	Voice Command	Trigger Keyword	Recognition Feedback
1	Turn on the Ceiling Light	Ceiling / light / Lighting / Room/ Fluorescent / Turn / On / Power / Switch / Start / Begin / Bright	Success
2	Turn off the Stand Light	Stand / Indirect / Light / Off / Stop / End / Dark	Failure
3	Turn on the TV	TV / Television / Turn / On / Start / Switch / Project / Power	Success
4	Turn off the Air Cleaner	Air / Cleaner / Purifier / Turn / Off / Stop / End	Failure
5	Turn on the PC	PC / Personal / Computer / Turn / On / Start / Up / Boot / Fire / Power	Failure
6	Turn on the Air Conditioner	Air / Conditioner / AC / / Heater / Conditioning / Turn / On / Start / Beginning / Power	Success
7	Turn off the Ceiling Light	Ceiling / Light / Lighting / Room / Fluorescent / Turn / Off / Stop / Switch / End / Dark	Failure
8	Turn on the Stand Light	Stand / Indirect / Light / On / Switch / Start / Bright	Success
9	Turn off the TV	TV / Television / Turn / Off / End / Switch	Failure
10	Turn on the Air Cleaner	Air / Cleaner / Purifier / Turn / On / Power / Start / Clean	Success
11	Turn off the PC	PC / Personal / Computer / Turn / Off / Shut / Down / Shutdown / Power / End	Success
12	Turn off the Air Conditioner	Air / Conditioner / AC / Heater / Conditioning / Turn / Off / End / Stop / Down	Failure

alizes the behavior through animation. As a front-end system, Processing 3.5.3 is used to develop the simulator and as a back-end system, a web application is also developed with the Web Speech API for speech recognition. The Web Speech API is a JavaScript API that enables developers to embed speech recognition features in web applications. In our experiment, Google Chrome serves the JavaScript runtime to support the Web Speech API's speech recognition function. A web server runs on the Processing, and the developed web application calls the API to obtain the recognized text of a user's speeches. The obtained text data are sent to the front-end system through WebSocket communication. When the front-end system receives the text data from the back-end system, it checks the data against the predefined dictionary. If the text data contain keywords listed in the dictionary, the agent responds to the user, and the system begins to operate the home appliance according to the command. If the received text data do not contain the keywords, it is considered speech recognition failure, and then the agent's personality is shifted.

3.2 Experimental Conditions

The experiment includes two conditions: an agent with personality shifting and without it. We used a between-participants experimental design. In the personality shifting condition, the agent shifts its personality when speech recognition fails, as shown in Table 2. In the non-shifting condition, the agent says ”I'm sorry. I didn't catch that.” if speech recognition is successful, the agents behave similarly in both conditions. In the experiments, we designed the simulator behavior so that speech recognition always fails at a specific time, even when it is successful. Twenty participants (19 male and 1 female) were used in the experiment. All participants were Shinshu University's students and the average age was 22.2 years old (S.D. = 0.73). Specifically, to ensure that the participants experience the agent's personality shifting in the personality shifting condition, the system was programmed to intentionally fail in speech recognition for 6 of the 12 commands that the participants were required to execute.

3.3 Procedure

The participants were briefed about the experiment on the simulator and only participated in the experiment if they agreed to it. They were led to a private room and given headphones equipped with microphones. A voice command list that contains the order of voice operations was given to them, and they were told to give commands to the simulator's home appliances based on the list. The voice command column in Table 4 shows the voice command list provided to the participants. The table also includes trigger keywords that trigger the agent to respond to the voice command and the success or failure result, as predesigned recognition feedback. Speech recognition always fails at least once for the entry labeled "Failure" in the table. The trigger keywords in the table such as nouns and verbs are originally in Japanese because the participants were Japanese. The same voice command list was used in both conditions. After instructing the simulator with the voice commands, the participants completed a questionnaire on their impressions of the system.

3.4 Evaluation Indices and Participants

Table 5 shows the questionnaire on their impression of the system, which consists of 26 items. Item 1 has a choice between 1 and 10. Items 2-26 are 7-point Likert-scale items. The number is large if the participant's impressions match the question; otherwise, the number is small. The questionnaire includes a section in which the participants can freely express their opinions and impressions of the experiment. Among the 20 participants, 10 people experienced the personality shifting condition and the other 10 people experienced the non-shifting condition.

4 RESULTS

Table 5 shows the result of the questionnaire and statistical analysis. The sample size for each condition is 10, which may not guarantee normality in the population. Therefore, we use Mann-Whitney's U test, which is highly reliable even when the sample size is small (Nachar, 2008).

The significant difference between the conditions in the mean value of Item 9 ($U = 23.50$, $p = 0.040$) and Item 17 ($U = 23.00$, $p = 0.041$) was found. This suggests that the personality shifting method leaves a stronger impression of good communication and shows a greater sense of responsibility than the non-shifting method.

In Item 1, the participants were asked how many talkers were in the system, and the results show that seven of ten participants answered 3, two answered 4, and one answered 2 in the personality shifting condition, implying that all participants were aware of the agents' personality shifting. Most of the participants in the non-shifting condition, on the other hand, thought the talker was one. This suggests that the participants were aware of the characteristics of each condition as we intended.

5 DISCUSSIONS

5.1 Effect of Personality Shifting

The statistical analysis suggests that personality shifting can provide stronger impressions of good communication (Item 9) and a greater sense of responsibility (Item 17) than the non-shifting method. For Item 9, notably, the participants interpreted that they could effectively communicate with the simulator even though the agent's extra speech during personality shifting may have negatively impacted dialog rhythm. For Item 17, the participants felt that the agent accepted the responsibility for failure in the personality shifting condition, which may have emphasized a sense of responsibility because what the agent says corresponds to what it does. In other words, saying sorry and resigning to accept the blame may give the impression that it is keeping to its word. However, this may be culturally dependent because all participants were Japanese.

Although there is no statistically significant difference in Item 2 (I felt emotions from the system), there is a significant difference between the experimental conditions. The score of the condition using the proposed method is higher than the condition not using it, which indicate that the proposed method may have emphasized emotional expressions. Agents that express emotions (George, 2019) are expected to play an important role in speech interaction in terms of empathizing with users and conveying non-verbal information. Although the proposed method does not directly deal with emotions, the behavior of personality shifting may affect emotional information in speech interaction. How such words and behaviors are interpreted as emotional information by users needs to be investigated in the future.

Multi-person dialog systems have limitations such

Table 5: Questionnaire items and experimental results.

No.	Questionnaire	Personality Shifting		Non-Shifting		U	$d.f.$	p
		Mean	S.D.	Mean	S.D.			
1	How many talkers are in the system?	3.10	0.57	1.20	0.63	–	–	–
2	I felt emotions from the system	3.90	2.02	2.20	1.62	28.50	18.0	0.089
3	I felt blamed by the system	1.20	0.63	1.20	0.63	50.00	18.0	1.000
4	I felt my instructions were well received by the system	4.70	1.25	4.50	1.65	48.00	18.0	0.905
5	My own instructions were appropriate	5.30	1.06	4.50	1.65	38.50	18.0	0.394
6	I felt that the system was stable	4.30	1.49	4.50	1.08	47.50	18.0	0.876
7	I felt that the system was confident	3.90	1.10	4.00	1.56	47.50	18.0	0.877
8	I felt that the system asked me so many time to repeat the command	4.50	1.35	3.90	1.52	37.50	18.0	0.353
9	I was able to communicate well with the system	5.30	1.34	4.10	1.29	23.50	18.0	0.040†
10	I felt that the system was friendly	5.40	1.35	4.60	1.51	34.50	18.0	0.246
11	I felt that the system was noisy	3.40	1.90	2.10	1.29	29.00	18.0	0.109
12	I felt frustrated using the system	2.40	1.51	3.30	1.70	33.50	18.0	0.216
13	I felt that the system was complicated	2.70	1.70	2.60	0.97	49.50	18.0	1.000
14	I felt that the system understood my instructions	5.70	1.16	5.00	1.25	33.50	18.0	0.216
15	I felt that the system was obedient to my instructions	5.70	1.34	6.00	1.05	44.50	18.0	0.692
16	I felt that this system was not performing well	3.30	1.77	2.90	1.29	45.50	18.0	0.757
17	I felt that the system had a strong sense of responsibility	4.50	1.90	2.60	1.71	23.00	18.0	0.041†
18	I could quickly figure out how to tell the system what to do	6.30	0.95	5.90	1.60	45.00	18.0	0.707
19	I could understand the meaning of the system's statements	6.70	0.67	6.60	0.70	45.50	18.0	0.690
20	I understand how to use the system	6.70	0.67	6.60	0.52	42.00	18.0	0.480
21	I felt that there was a calm atmosphere during the experiment	6.00	1.33	6.80	0.42	32.00	18.0	0.118
22	I felt the pleasant atmosphere during the experiment	5.00	1.63	4.40	1.17	36.50	18.0	0.308
23	I felt that I could easily control appliances in a similar experiment	4.70	1.77	5.20	1.69	40.50	18.0	0.485
24	I can trust this system	4.70	1.25	5.00	1.15	40.00	18.0	0.452
25	I felt the system considered my feelings	4.30	1.64	3.70	1.89	41.50	18.0	0.539
26	I want to use this system on a daily basis	4.50	1.65	5.00	1.33	38.00	18.0	0.367

†$p < 0.05$

as the loss of speech opportunities and the complexity of dialog (Nishimura et al., 2013). Although the proposed method also provides multi-personality, the mean values of the experimental results for the feeling of being blamed by the system (Item 3), frustration with the system (Item 12), and the complication of the system (Item 13) were all less than 4. Thus, the proposed method would avoid this type of problem because a user interacts with only one agent at a time. However, Item 11 suggests that the personality shifting condition may provide a noisier impression for users than the non-shifting condition. The average utterance time in the personality shifting condition was approximately 1 s longer than that in the non-shifting condition because of the additional utterance of "Next is my turn. Your orders, please." In addition to that, participants may have perceived a collective presence of multiple personalities even though they did not appear at the same time.

Since questionnaire results in this experiment were rated on a 7-point Likert-scale, a rating of 4 is interpreted as neutral. Although neither Item 8 nor Item 25 showed a significant difference in the statistical significance test, notably, the results of each of them are divided between the conditions with the evaluation value of 4 as the boundary. Item 8 is a question about the annoyance of the system. The personality shifting condition may be rated more negatively if the number of times the agent asks the participants the same question in both conditions is the same. On the other hand, Item 25 is a question about the system's consideration of user's feelings, and it might be seen that the personality shifting condition was rated more positively. A participant responded, "I felt like the agents were helping each other well by shifting personalities" as a free comment. By shifting personalities, it appears as if agents recover the failure caused by other agents. In the experiment, the relationship between the agents is not clearly expressed; however, it may be interpreted on the basis of Balance Theory for agents (Nakanishi et al., 2003). The design of the relationship between agents as well as the expression of individual agents may bring beneficial effects to users.

This suggests that the agent's personality shifting improved the impression of speech recognition failure and users' interaction experience.

5.2 Limitations

The results of this experiment were obtained under the constraint that the participants only interacted with the agent for a short period, approximately 10 min. Therefore, it is unclear whether the same effect will be observed when the proposed method is used for a longer period. The effect of habituation through long-term interaction between users and agents (Leite et al., 2013) should be investigated in the future, considering real-world scenarios.

The agent personalities used in the experiment had one female and two male voices, with the combination and shifting order fixed. Therefore, the effect of the ratio and gender distribution of the agent's personalities, as well as the order of alternation, is unknown. It is technically possible to shift the agent personality according to a user's gender, and such design guidelines and effective functions should be further investigated in the future. The results of this experiment were obtained under some constraints that the combination and order of the male and female voices of the agents were fixed, that the interaction period was short, insufficient number of participants, and that the participants were immersed in Japanese culture. These are subjects for further study.

The experiment was restricted to Japanese participants and the Japanese language. Different nuances of cultural pragmatics and politeness are emphasized in different cultural spheres in terms of their influence on experimental results (Haugh, 2004). Therefore, the experimental results may be strongly influenced by Japanese culture and language. In particular, it is contentious how personality shifting behavior is interpreted in other cultures, and this is a subject for further study.

6 CONCLUSION

This study proposed a method for shifting an agent's personality during speech interaction to reduce users' negative impressions of speech recognition systems when speech recognition fails. We developed a voice control simulator to simulate the operation of home appliances through voice commands and implemented the agent's personality shifting through the change of voice tone and LED color of a smart speaker. The experimental results suggested that the proposed method can provide stronger impressions of good communication and a greater sense of responsibility than the non-shifting method. Although speech recognition failure is uncomfortable and the cognitive strain of rephrasing commands is high for users, personality shifting could solve this type of problem. The proposed method has the advantage of allowing the agent to apologize and accept responsibility for speech recognition failure by taking the concrete action of dismissing the failed personality. This would be an important point for users and may become an essential element for various interactive systems.

REFERENCES

Breazeal, C., Kidd, C. D., Thomaz, A. L., Hoffman, G., and Berlin, M. (2005). Effects of nonverbal communication on efficiency and robustness in human-robot teamwork. In *2005 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 708–713. Institute of Electrical and Electronics Engineers.

Bulut, M., Narayanan, S. S., and Syrdal, A. K. (2002). Expressive speech synthesis using a concatenative synthesizer. In *Seventh International Conference on Spoken Language Processing*.

Çürüklü, B., Dodig-Crnkovic, G., and Akan, B. (2010). Towards industrial robots with human-like moral responsibilities. In *2010 5th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, pages 85–86. Institute of Electrical and Electronics Engineers.

Engelhardt, S., Hansson, E., and Leite, I. (2017). Better faulty than sorry : Investigating social recovery strategies to minimize the impact of failure in human-robot interaction. In *WCIHAI 2017 Workshop on Conversational Interruptions in Human-Agent Interactions : Proceedings of the first Workshop on Conversational Interruptions in Human-Agent Interactions co-located with 17th International Conference on International Conference on Intelligent Virtual Agents (IVA 2017) Stockholm, Sweden, August 27, 2017.*, volume 1943 of *CEUR Workshop Proceedings*, pages 19–27. KTH, Robotics, perception and learning, RPL, CEUR-WS.

Fiske, S. T. (1980). Attention and weight in person perception: The impact of negative and extreme behavior. *Journal of personality and Social Psychology*, 38(6):889.

George, S. (2019). From sex and therapy bots to virtual assistants and tutors: How emotional should artificially intelligent agents be? In *Proceedings of the 1st International Conference on Conversational User Interfaces*, CUI '19, New York, NY, USA. Association for Computing Machinery.

Guamán, S., Calvopiña, A., Orta, P., Tapia, F., and Yoo, S. G. (2018). Device control system for a smart home using voice commands: A practical case. In *Proceedings of the 2018 10th International Conference on Information Management and Engineering*, ICIME 2018, page 86–89, New York, NY, USA. Association for Computing Machinery.

Haugh, M. (2004). Revisiting the conceptualisation of politeness in english and japanese. *Multilingua*, 23(1/2):85–110.

Hojo, N., Ijima, Y., and Mizuno, H. (2018). DNN-based speech synthesis using speaker codes. *IEICE TRANSACTIONS on Information and Systems*, 101(2):462–472.

Kato, T., Okamoto, J., and Shozakai, M. (2008). Analysis of drivers' speech in a car environment. In *{INTERSPEECH} 2008, 9th Annual Conference of the International Speech Communication Association, Brisbane, Australia, September 22-26, 2008*, pages 1634–1637. ISCA.

Komatsu, T., Kobayashi, K., Yamada, S., Funakoshi, K., and Nakano, M. (2018). Vibrational Artificial Subtle Expressions: Conveying System's Confidence Level to Users by Means of Smartphone Vibration. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, CHI '18, pages 1–9, New York, NY, USA. Association for Computing Machinery.

Le, T., Gilberton, P., and Duong, N. Q. K. (2019). Discriminate Natural versus Loudspeaker Emitted Speech. In *ICASSP 2019 - 2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 501–505. Institute of Electrical and Electronics Engineers.

Leite, I., Martinho, C., and Paiva, A. (2013). Social Robots for Long-Term Interaction: A Survey. *International Journal of Social Robotics*, 5(2):291–308.

Nachar, N. (2008). The mann-whitney u: A test for assessing whether two independent samples come from the same distribution. *Tutorials in Quantitative Methods for Psychology*, 4.

Nakanishi, H., Nakazawa, S., Ishida, T., Takanashi, K., and Isbister, K. (2003). Can software agents influence human relations? balance theory in agent-mediated communities. In *Proceedings of the Second International Joint Conference on Autonomous Agents and Multiagent Systems*, AAMAS '03, page 717–724, New York, NY, USA. Association for Computing Machinery.

Nishimura, R., Todo, Y., Yamamoto, K., and Nakagawa, S. (2013). Chat-like Spoken Dialog System for a Multiparty Dialog Incorporating Two Agents and a User. In *Proc. of iHAI2013 : The 1st International Conference on Human-Agent Interaction*, pages II–2–p13.

Philpott, M. (2018). 2019 trends to watch: Smart home. Ovum Ltd.

Webb, H., Jirotka, M., F.T. Winfield, A., and Winkle, K. (2019). Human-Robot Relationships and the Development of Responsible Social Robots. In *Proceedings of the Halfway to the Future Symposium 2019*, HTTF 2019, pages 1–7, New York, NY, USA. Association for Computing Machinery.

Yamamoto, K., Inoue, K., Nakamura, S., Takanashi, K., and Kawahara, T. (2018). Dialogue Behavior Control Model for Expressing a Character of Humanoid Robots. In *2018 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC)*, pages 1732–1737. Institute of Electrical and Electronics Engineers.

Yoshikawa, Y., Iio, T., Arimoto, T., Sugiyama, H., and Ishiguro, H. (2017). Proactive conversation between multiple robots to improve the sense of human-robot conversation. In *AAAI 2017 Fall Symposium Series*, pages 288–294.