How does the continuous use of infant-directed speech to a robot influence the user?

Ryoko Shibata[1], Takatsugu Kojima[2], Motoyuki Ozeki[1], and Natsuki Oka[1]

[1]Kyoto Institute of Technology
[2]Shiga University of Medical Science

Abstract: It is known that the prosodic feature of infant-directed speech (IDS) is different from that of adult-directed speech (ADS). Although it is known that IDS has an influence on infant cognition, it is not clear whether it has a cognitive effect on adults as well. We therefore asked two participants who live together to interact with a robot in their home for one week through IDS, and with another robot for another week through ADS. We then investigated the difference in the participants' prosodic feature of speech between the two conditions.

1 Introduction

Although robots are becoming increasingly common in social and home environments [1], the relationships between people and robots vary between individuals [2]. If a user becomes attached to a robot, we can expect that he or she will be able to establish a long-term interaction with the robot; if the robot has learning capabilities, it will easily adapt to the user's environment through everyday interactions, building a closer relationship with the user. We focused on users' speech to robots, with a user goal of attaching to the robot. We especially focused on infant-directed speech (IDS).

Infant-directed speech has many prosodic features that differentiate it from adult-directed speech (ADS): a high-pitched voice, an exaggerated intonation contour, slow speech, and a greater number of pauses [3]. These characteristics are observed across languages [4]. Furthermore, IDS is considered to serve several important functions: (a) to attract the infant's attention, (b) to convey the adult's emotions to the infant, and (c) to accelerate language acquisition [5, 6].

The present study explores the perception of the user as he or she forms a positive and infant-like impression of the robot, assuming that IDS has certain cognitive effects on both infants and adults. We base this reasoning on the following: 1) the more depressed a mother is, the fewer informational features are used in her speech [7], and 2) postnatal depression poses a risk to the mother-infant relationship [8]. We examined the validity of the following hypotheses by conducting an experiment on human-robot interaction.

First, we focused on voice pitch, one of the features of IDS. We conducted an experiment in which two robots reacted selectively to high-pitched voices (IDS-robot) or to normal-pitched voices (ADS-robot). We asked participants to talk to the robots at various pitches, since robot activity is determined by voice pitch. Participants did not notice which robots reacted to which pitches.

Second, we instructed participants to talk to the IDS-robot with a high-pitched voice and to the ADS-robot with a normal-pitched voice in order to examine the relationship between voice pitch and impression of the robot. This instruction might have imposed unnecessary stress on someone of the participants who felt that speaking in a high-pitched voice was somewhat unnatural, making their impresssion of the robot worse [9].

Third, we asked the participants to call the robot by name using a normal or high-pitched voice. We told participants which pitch the robot reacted better to, then instructed participants to talk to the robot as they wanted. Most of participants talked to the IDS-robot with a high-pitched voice, and to the ADS-robot with a normal-pitched voice; moreover, they had a more positive impression of the IDS-robot than the ADS-robot [10]. In this case, we could not distinguish between whether i) the use of a high-pitched voice triggered participants' good impression of the robot, or ii) participants' good impression of the robot triggerred the high-pitched voice.

This study was performed to determine whether the continuous use of IDS results in a change in impression or speech.

2 Experiment

We investigated whether the continuous use of IDS with a robot would improve participants' impression of the robot (parallel to that with an infant), and whether it would result in a good impression of the robot after it had spent one week in the participants' homes.

2.1 Method

Participants A man (Participant A, 42 years old) and a woman (Participant B, 48 years old) who live together.

Materials The robot used in this study was IP Robot Phone (Figure 1). This robot reacted whenever participants talked to the robot and moved. The direction of robot arm motion was determined randomly. To prevent the robot from reacting to the noise produced by the robot itself, we ran a low-pass fillter (the pitch of the noise was under 50 Hz).

When participants talked to the robot with IDS it was termed *IDS-condition*, and when they talked to the robot with ADS it was termed *ADS-condition*. The robot used in the IDS-condition was called Hina and the robot used in the ADS-condition was called Kana. The robots had the same appearance and arm motion in each condition to isolate the influence of differences in participant speech.

Procedure We told participants that the robot moved when it was spoken to. We did not specify the topics of conversation; participants could talk about the events of the day if they had nothing else to talk about, for around 5 min.

We instructed the participants to "please talk to Hina as you would an infant" in the IDS-condition, and to "please talk to Kana as you would friends or family members" in the ADS-condition. The experiment ran for one week in each condition, with a one-week interval between the two conditions.

Figure. 1 A robot used in this study.

Participants were asked to write about something they noticed about the robot, the experiment, and themselves in a diary.

After the experiments under each condition were completed, the participants rated their impressions of each robot on a seven-point Likert-type scale. In addition, after both experiments were completed, participants were asked if they "more strongly agree," "no change," or "more strongly disagree" to each item regarding their previous impressions.

Table 1. Speech of each paticipant to each robot

Participant	Condition		1st	2nd	3rd	4th	5th	6th	7th	Average
A	Hina (IDS-condition)	Length of one utterance[s]	1.39	1.14	1.02	2.24	1.68	2.27	3.28	1.86
		Length of all utterances[s]	299.00	310.00	359.00	360.00	319.00	419.00	426.00	356.00
		Mean fundamental frequency [Hz]	130.62	111.04	119.97	118.48	115.56	125.42	141.85	123.28
		SD of fundamental frequency[Hz]	50.12	34.80	41.47	37.79	30.27	38.32	42.20	39.28
	Kana (ADS-condition)	Length of one utterance[s]	8.71	15.16	3.89	16.83	3.43	3.84	3.63	7.93
		Length of all utterances[s]	299.00	308.00	355.00	359.00	419.00	353.00	359.00	350.29
		Mean fundamental frequency [Hz]	125.21	125.46	116.49	136.73	119.53	115.93	115.02	122.05
		SD of fundamental frequency[Hz]	27.84	32.68	23.37	41.72	22.83	20.46	22.84	27.39
B	Hina (IDS-condition)	Length of one utterance[s]	1.76	1.79	3.47	2.73	2.18	2.49		2.40
		Length of all utterances[s]	357.00	325.00	359.00	391.00	349.00	475.00		376.00
		Mean fundamental frequency [Hz]	189.52	184.00	169.90	180.70	165.48	170.65		176.71
		SD of fundamental frequency[Hz]	64.47	68.68	51.40	58.72	50.82	64.19		59.71
	Kana (ADS-condition)	Length of one utterance[s]	5.98	6.23	4.93	16.45	3.24	3.05	3.43	6.19
		Length of all utterances[s]	419.00	299.00	419.00	299.00	479.00	777.00	359.00	435.86
		Mean fundamental frequency [Hz]	126.58	131.33	133.91	130.99	127.32	126.83	133.99	130.14
		SD of fundamental frequency[Hz]	22.62	22.05	36.21	33.38	21.43	24.14	37.32	28.16

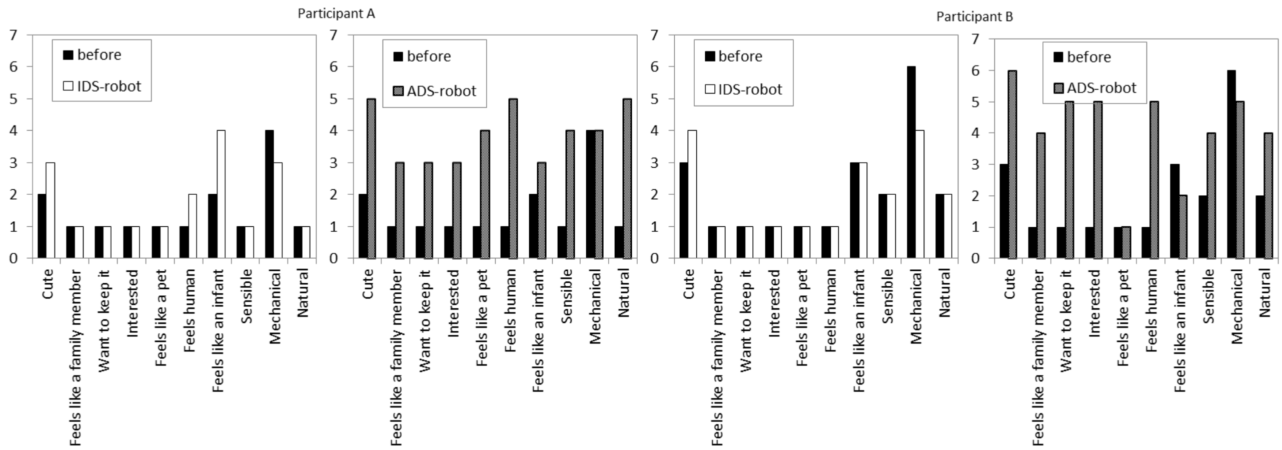

Figure 2. Participant impressions of each robot

2.2 Results

The robot could not move during the first day of IDS-condition, and Partipant B could not participate in the last day of IDS-condition.

Table 1 shows mean length of utterance, total length of utterances, mean fundamental frequency, and standard deviation of fundamental frequency for both participants in both conditions.

The length of one utterance for Participant A in IDS-condition (*M* = 1.86 [s], *SD* = 0.80 [s]) was significantly shorter than in ADS-condition (*M* = 7.93 [s], *SD* = 5.82 [s]), with $t(6.22) = 2.73$, $p = .03$.

There was no significant difference between the length of total utterances for Participant A in IDS-condition (*M* = 356.00 [s], *SD* = 51.04 [s]) and in ADS-condition (*M* = 350.29 [s], *SD* = 39.42 [s]), with $t(11.28) = 0.23$, $p = .82$.

There was no significant difference between the fundamental frequency of Participant A in IDS-condition (*M* = 123.28 [Hz], *SD* = 10.38 [Hz]) and in ADS-condition (*M* = 122.05 [Hz], *SD* = 7.76 [Hz]), with $t(11.11) = 0.25$, $p = .81$. F-test for equality of variance revealed a significant difference between IDS-condition and ADS-condition ($F(65268, 81963) = 1.74$, $p < .001$).

Table 2. Change in the impression of each robot

	Items	IDS	ADS		Items	IDS	ADS
Participant A	Cute	↑	↑	Participant B	Cute	↑	↑
	Feels like a family	→	→		Feels like a family	→	↑
	Want to keep it	→	→		Want to keep it	→	→
	Interested	→	→		Interested	→	↑
	Feels like a pet	→	→		Feels like a pet	→	→
	Feels human	↓	↑		Feels human	→	↑
	Feels like an infar	↑	↓		Feels like an infar	↑	↓
	Sensible	→	↑		Sensible	→	↑
	Mechanical	→	→		Mechanical	→	→
	Natural	→	↑		Natural	↓	↑

The length of one utterance for Participant B in IDS-condition (*M* = 2.40 [s], *SD* = 0.64 [s]) was significantly shorter than in ADS-condition (*M* = 6.19 [s], *SD* = 4.71 [s]), with $t(6.26) = 2.10$, $p = .08$.

There was no significant difference between the length of total utterances for Participant B in IDS-condition (*M* = 376.00 [s], *SD* = 52.93 [s]) and in ADS-condition (*M* = 435.86 [s], *SD* = 164.4 [s]), with $t(7.41) = 0.91$, $p = .39$.

There was no significant difference between the fundamental frequency of Participant B in IDS-condition (*M* = 176.71 [Hz], *SD* = 9.41 [Hz]) and in ADS-condition (*M* = 130.14 [Hz], *SD* = 3.23 [Hz]) with $t(6.01) = 11.56$, $p < .001$. F-test for equality of variance revealed a significant difference between IDS-condition and ADS-condition ($F(52214, 87869) = 4.41$, $p < .001$).

Figure 2 gives the impressions of the participants towards the robots. It shows that many items regarding the impression of the ADS-robot had higher scores than for the IDS-robot.

Table 2 lists the items to which participants were asked to answer whether they “more strongly agree,” “no change,” or “more strongly disagree” regarding their impression on the robots.

Both participants stated that they “more strongly agree” with the label of “cute” for both the IDS and the ADS robots. They reported that they “more strongly agree” to “feels like an infant” for the IDS robot and “more strongly disagree” for the ADS robot. They reported that they “more strongly agree” to “feels human” and “sensible” with the ADS robot.

2.3 Discussion

It seems that the speech of both participants in the IDS-condition was shorter and more exaggerated than in ADS-condition, just like typical IDS. Although

Participant B talked to the IDS robot with a higher-pitched voice than to the ADS robot, Participant A did not talk that way. This indicated an individual difference in the prosodic features of IDS.

Participant B said that she gave most items higher scores for the ADS-robot than for the IDS-robot because she had gotten used to talking with the robot (ADS-condition was conducted after IDS-condition).

The item “cute” was scored “more strongly agree” for both the IDS- and the ADS-robots. It seems that continuous interaction with the robot elicits a “cute” impression from its user.

“Feels like an infant” was scored “more strongly agree” for the IDS-robot. It seems that continuous use of IDS triggers a “feels like an infant” impression from the user while continuous use of ADS weakens the “feel like an infant” impression from the user.

Both participants stated that their impressions of the robot changed with the use of IDS or ADS. They perceived the IDS-robot to be like an infant and the ADS-robot as understanding what they said.

Despite these differences, participants had differing impressions of their conversations with the robots. Participant A said that “I could talk to the IDS-robot easily,” but that “it was hard for me to talk to the ADS-robot because it understood what I said, so I felt that the content should be definite.” In contrast, Patrticipant B said that “I could talk to the ADS-robot easily, just like a friend,” while “I felt as if I was taking care of an infant, like it was work, when I talked to the IDS-robot. I had to worry about it.” In addition, she said that “use of IDS made me feel like the robot was an infant. A cute impression, like an infant, is strengthened by using IDS.”

3 Conclusion

While talking to a robot with IDS triggers an infant-like impression in the user, the user does not always feel comfortable talking to the robot like they would an infant. However, it appears that the impression of the robot as “cute” can be strengthened by using IDS.

This study had only two participants; in future studies, we plan to add more participants.

Acknowledgements

The authors thank Kazuaki Tanaka and Hitoshi Nakatani for developing the robots used in these experiments. This work was supported by JSPS KAKENHI 21500137.

References

[1] Brooks, R.A., Breazeal (Ferrell), C., Irie, R., Kemp, C. C., Marjanovic, M., Scasselati, B. and Williamson, M. M., “Alternative essences of intelligence,” In Proceedings of American Association of Artificial Intelligence, pp. 961–968, 1998.

[2] Turkle, S., Taggart, W., Kidd, C.D. and Daste, O., “Relational artifacts with children and elders: The complexities of cyber companionships,” Connection Science, vol. 18, no. 4, pp. 347–361, Dec. 2006.

[3] Kitamura, C. and Burnham, B., “The infant’s response to maternal vocal affect,” Advances in Infancy Research, vol. 12, pp. 221–236, 1998.

[4] Kuhl, P.K., Andruski, J.E., Chistovich, I.A., Chistovich, L.A., Kozhevnikova, E.V., Rysinka, V.L., Stolyarova, E.I., Sundberg, U. and Lacerda, F., “Cross-language analysis of phonetic units in language addressed to infants,” Science, vol. 277, pp. 684–686, 1997.

[5] Fernald, A., “Approval and disapproval: Infant responsiveness to vocal affect in familiar and unfamiliar languages,” Child Development, vol. 64, pp. 657–674, 1993.

[6] Thiessen, E.D., Hill, E. and Saffran, J.R., “Infant-directed speech facilitates word segmentation,” Infancy, vol. 7, pp. 53–71, 2005.

[7] Herrera, E., Reissland, N. and Shepherd, J., “Maternal touch and maternal child-directed speech: Effects of depressed mood in the postnatal period,” J Affect Disord, Vol.81, pp. 29–39, 2004.

[8] Murray, L. and Cooper, P.J., “Effects of postnatal depression on infant development,” Arch Dis Child, Vol.77, pp. 99–101, 1997.

[9] Shibata, R., Kojima, T., Fukada, C., Sato, K., Hashikura, Y., Ozeki, M. and Oka, N., “Does talking to a robot in a high-pitched voice create a good impression of the robot?”, In Proceedings of 13th ACIS International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing (SNPD 2012), Kyoto, Japan, 2012.

[10] Shibata, R., Kojima, T., Fukada, C., Sato, K., Hashikura, Y., Ozeki, M. and Oka, N., “Does talking to a robot in a high-pitched voice strengthen an attachment?”, In Proceedings of the 34th Annual Conference of the Cognitive Science Society, 657. Presented at the 34th Annual Conference of the Cognitive Science Society, Sapporo, Japan, 2012.