The 2009 IEEE/RSJ International Conference on
Intelligent Robots and Systems
October 11-15, 2009 St. Louis, USA

Lexical Entrainment in Human-Robot Interaction: Can Robots Entrain Human Vocabulary?

Takamasa Iio, Masahiro Shiomi, Kazuhiko Shinozawa, Takahiro Miyashita, Takaaki Akimoto and Norihiro Hagita

Abstract— **A communication robot must recognize a referred-to object to support us in daily life. However, using our wide human vocabulary, we often refer to objects in terms that are incomprehensible to the robot. This paper focuses on lexical entrainment to solve this problem. Lexical entrainment is the phenomenon of people tending to adopt the terms of their interlocutor. While this has been well studied in human-computer interaction, few published papers have approached it in human-robot interaction. To investigate how lexical entrainment occurs in human-robot interaction, we conduct experiments where people instruct the robot to move objects. Our results show that two types of lexical entrainment occur in human-robot interaction. We also discuss the effects of the state of objects on lexical entrainment. Finally, we developed a test bed system for recognizing a referred-to object on the basis of knowledge from our experiments.**

I. Introduction

THE technical roadmap in robotics formulated by the Japanese government envisions a communication robot for a senior citizen living alone [1]. The robot can manipulate electric household appliances and take out and put away objects. To achieve such a robot, we need a wide range of technologies, such as structured environmental information, autonomous mobile control, manipulation, and human-robot interaction. In recent years, many studies have discussed human-robot interaction [2]. People will communicate with a robot intuitively even if they have no great knowledge of the robot because the robot can provide multi-modal interaction with voice or body motion like them.

Our objective is to design human-robot interaction that can facilitate a smooth conversation. In particular, we focus on the situation where a human refers to an object in the environment and the robot recognizes the referred-to object (Fig. 1), since this scenario is important for a communication robot that supports us in daily life. For example, when we instruct a robot to move one of the desks in a room, the robot has to recognize which desk we are referring to.

To recognize a referred-to object, a robot needs to combine various recognition technologies, such as speech recognition, pointing gesture recognition, and position detection of objects [18]-[19]. There are two approaches to improving the performance of these recognition technologies: the engineering approach and the entrainment approach. The former focuses on development of new devices or algorithms [5], [22]-[25], such as speech recognition using microphone arrays, multi-modal robot systems using other kinds of sensors, and detection of objects using ID-tags. The latter focuses on entrainment in human-robot interaction [15], [20]-[21]. In particular, there have been many reports of physical entrainment as conjugated gaze and nod.

In our work, we discuss lexical entrainment in human-robot interaction; accordingly, we investigate whether a robot can lead the utterance of a human. Lexical entrainment has been well studied in human-computer interaction. However, little attention has been given to lexical entrainment in human-robot interaction. We believe that the performance of speech recognition will improve if the robot can encourage the user to speak in easy terms through the interaction.

This paper reports the properties of lexical entrainment specific to human-robot interaction; furthermore, it describes our robot test bed system to encourage lexical entrainment on the basis of these properties. The remainder of the paper is organized as follows. Section 2 reports related works on physical entrainment in human-robot interaction and lexical entrainment in human-computer interaction. Section 3 hypothesizes about the properties of lexical entrainment in human-robot interaction and introduces our experimental methodology. Section 4 describes our experimental results, while Section 5 discusses lexical entrainment in human-robot interaction and the experimental limitations. Section 6 gives details of our test bed system, and finally in Section 7 we draw our conclusions.

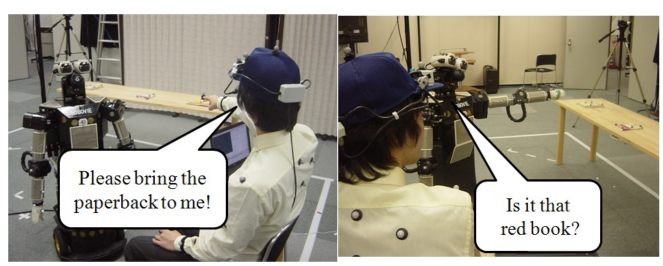

Fig. 1 Recognition of a referred-to object.

Manuscript received March 1, 2009. A part of this research was supported by KAKENHI, Grant-in-Aid for Young Scientists (B), 20700191, and a part of this research was supported by New Energy and Industrial Technology Development.

T. Iio, M. Shiomi, K. Shinozawa, T. Miyashita, T. Akimoto and N. Hagita are with the Advanced Telecommunications Research Institute International, Kyoto, JAPAN (corresponding author's phone: +81-774-95-2570; fax: +81-774-95-1401; e-mail: m-shiomi@atr.jp).

T. Iio is also with Doshisha University, Kyoto, Japan.

II. Related Work

A. Physical Entrainment

Many kinds of physical entrainment have been observed in human-human interaction. Condon et al. reported that a baby's motion is engaged by his/her mother's speech [26]. Other researchers also reported entrainment of a posture in human-human interaction [27]-[28].

In human-robot interaction, physical entrainment has been often observed in the joint attention process as well as human-human interaction [29]. Ono et al. observed physical entrainment of a pointing gesture in the interaction between people and a guide robot [30]. Breazeal also found physical entrainment of head tilt and facial expression [31]. Ogawa et al. developed a robot that synchronizes to a human's behaviors and observed the entrainment of a nodding behavior [32].

Such physical entrainment is useful for recognizing a referred-to object in human-robot interaction; for example, a pointing gesture is helpful to recognize the object. However, it is difficult to identify an object by only a pointing gesture when many objects exist in one place. In such a situation, speech information is needed to identify the object; therefore, it is important to focus on lexical entrainment for evoking easy referential terms.

B. Lexical Entrainment

There is potential for enormous variability in people's lexical choices in dialog with a human, computer, or robot interlocutor. Furnas, Landauer, Gomez, and Dumais coined this variability as "the Vocabulary Problem" in their studies of command languages [6]. This problem has been well studied; it has also been revealed that this variability is lower within a conversation than between conversations [7]-[9]. When two persons repeatedly discuss the same object, they come to use the same referential term. This phenomenon has been called lexical entrainment.

Lexical entrainment has been studied in human-computer interaction. Through two Wizard-of-Oz experiments using a database query task [12], Brennan suggested that people adopt terms of their computer interlocutor. In a similar experiment, Gustafson simulated a tourist information system with a speech interface. When the system asked users a question containing an infrequently used verb, they used the verb in their answer [13]. These results showed that users of a spoken dialog system also adapt their lexical choices to match system vocabulary. Moreover, Tomko found that users tend to mirror the simple form of confirmation delivered by a system when the users' inputs are rejected by the system [14].

Some research efforts have focused on dialogue in human-robot interaction [11]. For example, a robot took the initiative in a conversation and could narrow down the vocabulary of a user [15]. Moreover, a tension-moderating technique was effective for promoting more natural pronouns [16]. Shinozawa et al. reported on a persuasive robot that led people to select a color name that the robot selected before through verbal interaction [17]. However, lexical entrainment has not been investigated in human-robot interaction.

In human-robot interaction, a human, a robot, and a referent occupy the same environment; therefore, we suggest that the state of the referent affects an occurrence of lexical entrainment because the human attends to not only the robot's speech but also the state (e.g. location). This aspect has never been discussed in human-computer interaction.

III. Methodology

According to previous works on lexical entrainment in human-computer interaction, people are likely to adopt the terms of their computer interlocutor. We should thus investigate whether people adopt the terms of their robot interlocutor. In particular, we are interested in whether lexical entrainment occurs when people and a robot refer to an object because it will be useful in recognizing a referred-to object.

We conducted experiments based on a Wizard-of-Oz method because it was difficult for the robot to automatically recognize a referred-to object. The difficulty arose for the following two reasons. When subjects referred to an object,

i. they could use not only their voice but also a pointing gesture and gaze, and
ii. they could refer to the object with various types of expressions: the name, the color, the size, the shape, or the position of the object.

Therefore, an operator played the role of some of the robot's sensors to avoid the difficulty of recognizing a referred-to object.

In this section, we hypothesize about the state of the referent affecting lexical entrainment, and we describe an experimental methodology for investigating our hypotheses.

A. Hypotheses

We tested five hypotheses on lexical entrainment:

Hypothesis 1: Subjects will adopt a referential term used by a robot when they refer to an object.

Hypothesis 2: Subjects will prefer the same type of referential term when the robot limits the referential term to a certain type; for example, they would adopt color adjectives if the robot always used color adjectives.

Hypothesis 3: Subjects will more likely adopt a referential term used by a robot when they refer to an unfamiliar object than to a familiar object.

Hypothesis 4: Subjects will more likely adopt a referential term used by a robot when they refer to a hidden object than to an object in their view.

Hypothesis 5: Subjects will more likely adopt a referential term of the robot when they refer to an unfamiliar, hidden object.

We focus on whether lexical entrainment occurs in human-robot interaction in hypothesis 1. The objective of hypothesis 2 is to explore the relationship between the way used to confirm a referred-to object and lexical entrainment.

For hypotheses 3, 4, and 5, we discuss the effects of the state of an object on the occurrence frequency of lexical entrainment.

B. Experimental Design

To verify the above hypotheses, we conducted a laboratory experiment in which a subject instructed the robot to move an object in the experimental environment. The robot first greets the subject and then introduces itself. The robot asks the subject which object he/she would like moved. After the subject chooses an object, the robot confirms the object. In the confirmation, the robot directs its gaze and pointed finger to the object; moreover, it speaks from a script prepared for each object. If the confirmation is correct, the subject specifies another object, otherwise he/she specifies the same object again.

We used the confirmation stage to initiate lexical entrainment because this is a natural process in human-human interaction. If a subject adopted the terms of the robot's confirmation, we regarded it as lexical entrainment.

We chose books as objects used in our task because they are found in many households and, moreover, they involve various expressions for reference such as title, color, type, author, shape, and location. Table I shows a typical example of a human-robot dialogue in our experiment.

TABLE I
EXAMPLE DIALOGUE

Speaker	*Dialogue*
Robot	Please indicate a book.
Subject	That travel magazine, please.
Robot	The Red book?
Subject	Yes.
Robot	OK, please indicate a next book.
Subject	Hmm, carry this novel, please.
	(This task continues until all books are specified.)

1) Environment

Figure 2 depicts the experimental setup. The experiment was conducted in a rectangular room 7.5 m by 10 m. We used an area of 3.5 m by 3.5 m in the center of the room due to the restricted area covered by the video camera. A subject was seated in front of the robot. Five different books were positioned between the subject and the robot so that the subject could identify these books by sight.

a) Video camera and microphone

This is a Wizard of Oz experiment, in which an operator operates the robot remotely. We installed a video camera in the experimental room so that the operator could look at the body motion of the subject and the robot, and we attached a microphone to the subject's body so that the operator could listen to the subject's speech. The video image and voice were recorded for analysis after the experiment.

b) Operator role

The operator recognized the subject's reference instead of the robot and started up a behavior program for the robot on the basis of the voice and gesture of the subject.

We assumed that the robot had limited cognitive abilities, and thus the operator rejected all reference expressions except for the following three.

1. Reference by a book title.
2. Reference by a book color.
3. Reference by pointing a finger at a book.

The second rule was set from the supposition that a robot could pick up a characteristic color of a book by image recognition techniques in the future; therefore, the operator recognized references by a book color only when the predefined color terms were used. The third rule was also set on the basis of the feasibility of pointing-gesture recognition.

c) Robot

Robovie-R ver. 2 is a humanoid robot developed by the Intelligent Robotics and Communication Labs, ATR. It has a human-like upper body designed for communicating with humans. Figure 3 shows its overview. It has a head, two arms, a body and a wheeled-type mobile base. On its head, it has two CCD cameras for eyes and a speaker for a mouth. The speaker can output recorded sound files installed on the internal-control PC located on the body. The robot has several degrees of freedom (DOFs): two DOFs for the wheels, three DOFs for its neck, and four DOFs for each arm. Its body has sufficient expressive ability to perform human-like gestures. In addition, it has two wheels to move (forward-reverse travel and rotation). Its height is 1100 mm, its width is 560 mm, its depth is 500 mm and its weight is about 57 kg.

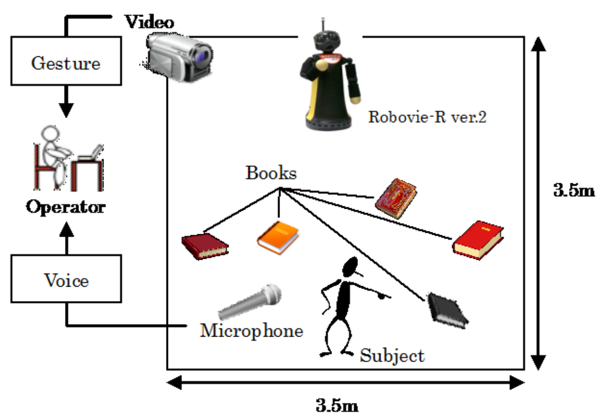

Fig. 2 Environment of the experiment.

Fig. 3 Robovie-R ver.2

2) Experimental Procedure

In all of the experiments, a subject was first given a brief description of the purpose and the procedure of the experiment. After this introductory explanation, the subject was asked to review and sign a consent form. The subject moved to the experimental laboratory and was then given the details of the task. We told the subject that we were developing a robot for recognizing an object and would like his/her help in evaluating the design. The subject was assigned to one condition in each experiment. After completing the task, the subjects answered a questionnaire that measured their impression of the robot. Figure 4 presents images of the experiment.

Fig. 4 Images of the experiment.

3) Experimental Conditions

We conducted five experiments, with two sessions (i.e. two conditions) in each experiment.

Experiment 1 was a within-subjects experiment. In this experiment, we investigated whether a subject adopted the terms of the robot's confirmation utterances. Table II lists the books used in the experiment. The subject had two dialogues with the robot, i.e. session 1 and session 2, in the following states.

Session 1: The subject did not know the referential terms of the robot.

Session 2: The subject knew the referential terms of the robot through the robot's confirmation.

Experiment 2 was a between-subjects experiment. In this experiment, we investigated whether the subject used a certain type of term when the robot continued to give confirmation with the same type of term during session 1. The subject was assigned to either of two conditions and had one dialogue with the robot. The conditions were as follows.

Condition 1: Various types of confirmation terms were used, as listed in Table II.

Condition 2: The color type of confirmation terms was used, as listed in Table III.

Experiment 3 was a between-subjects experiment. In this experiment, we investigated whether the subject's reference varied according to the readability of the book title; we thought the English books were unreadable for our subjects because they were native speakers of Japanese. The subject was assigned to either of two conditions and had one dialogue with the robot. The conditions were as follows.

Condition 1: The subject referred to the Japanese books listed in Table III

Condition 2: The subject referred to the English books listed in Table IV.

Experiment 4 was a within-subjects experiment. In this experiment, we investigated how the subject's reference varied if the books were hidden. The books are shown in Table III. The subject had two dialogues with the robot; the first dialog was session 1, and the second dialog was session 2. The subject was in either of the following states.

Session 1: The books were around the subject.

Session 2: The books were removed and not in the area.

Experiment 5 was a between-subjects experiment. In this experiment, we investigated how the subject's reference varied if the books were unmemorable. The subject participated in Experiment 3 and then proceeded to participate in Experiment 5; we removed the books during the subject's rest time between these two experiments.

Condition 1: The Japanese books listed in Table III were removed and not around the subject.

Condition 2: The English books listed in Table IV were removed and not around the subject.

The color type of confirmation was used in Experiments 3, 4 and 5.

TABLE II
SCRIPTS OF CONFIRMATION UTTERANCES

Book	*Type*	*Confirmation*
Japanese-Book 1	Color	Is it the yellow book?
Japanese-Book 2	Title	Is it the "Descartes' sanctum"?
Japanese-Book 3	Size	Is it the large book?
Japanese-Book 4	Classification	Is it the comic book?
Japanese-Book 5	Title subset	Is it the book of Kyoto?

TABLE III
SCRIPTS OF CONFIRMATION (ONLY COLOR)

Book	*Type*	*Confirmation*
Japanese-Book 1	Color	Is it the yellow book?
Japanese-Book 2	Color	Is it the gray book?
Japanese-Book 3	Color	Is it the white book?
Japanese-Book 4	Color	Is it the purple book?
Japanese-Book 5	Color	Is it the red book?

TABLE IV
SCRIPTS OF CONFIRMATION (ONLY COLOR)

Book	*Type*	*Confirmation*
English-Book 1	Color	Is it the gray book?
English-Book 2	Color	Is it the brown book?
English-Book 3	Color	Is it the white book?
English-Book 4	Color	Is it the green book?
English-Book 5	Color	Is it the black book?

4) *Measurement*

We set two independent variables in each experiment. The dependent variables involved only one measurement: the number of references including the confirmation terms prepared for each book.

Experiment 1: In both sessions, we measured the number of references that include the confirmation terms in Table II.

Experiment 2: In both conditions, we measured the number of references that include the confirmation terms in Table III.

Experiment 3: In condition 1, we measured the number of references that include the confirmation terms in Table III. In condition 2, we measured the number of references that include the confirmation terms in Table IV.

Experiment 4: In both sessions, we measured the number of references that include the confirmation terms in Table III.

Experiment 5: In condition 1, we measured the number of references that include the confirmation terms in Table III. In condition 2, we measured the number of references that include the confirmation terms in Table IV.

IV. RESULTS

In Experiments 1 and 4, we applied the McNemar test to compare proportions in matched pairs of the subjects' references between sessions. Otherwise, we applied the Chi-square test.

A. *Experiment 1 (within-subjects)*

The results of analysis of 40 references from 8 subjects are shown in Table V. In session 1, subjects tended to read out the book titles. After subjects received confirmation from the robot, they were likely to adopt the confirmation terms prepared for each book (listed in Table II).

TABLE V
NUMBER OF REFERENCES THAT INCLUDE CONFIRMATION TERMS

Book	*Type*	*Number of indications including confirmative terms*	
		Session 1	*Session 2*
Book 1	Color	0	2
Book 2	Title	5	5
Book 3	Size	2	3
Book 4	Classification	0	3
Book 5	Title subset	1	4

Moreover, Figure 5 compares the references to books by users, i.e. using or not using the confirmation terms, between sessions. This figure shows a significant difference between the proportions of references that include the confirmation terms in session 1 and session 2 ($\chi^2 = 4.267, p < 0.05$). From these results, hypothesis 1 is verified.

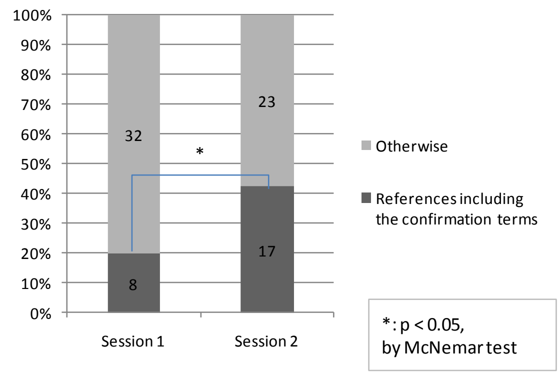

Fig. 5 Proportion of references that include the confirmation terms before subjects receive confirmation (Session 1) and after they received confirmation (Session 2).

B. *Experiment 2 (between-subjects)*

We obtained 39 references from the 8 subjects in condition 1 and 54 references from the 11 subjects in condition 2. Figure 6 shows the proportions of references that include the confirmation terms listed in Table III in condition 1 and condition 2. The proportion is significantly higher when the robot confirms the book by using the color type of terms rather than the various types of other terms ($\chi^2 = 8.092, p < 0.01$). Subjects came to adopt the color types of terms even if they referred to another book without knowing its confirmation terms. The results validate hypothesis 2.

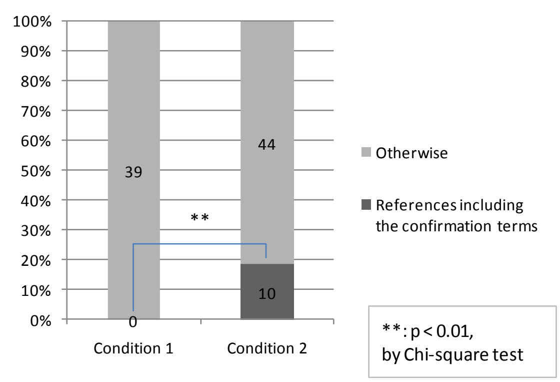

Fig. 6 Proportion of references that include the color type of terms when the robot confirmed by using all types of terms (Condition 1) and the color type of terms (Condition 2).

C. *Experiment 3 (between-subjects)*

Seven subjects participated in Experiment 3 and provided 35 references in each condition. All were native speakers of Japanese. Figure 7 shows the proportions of references that include the color type of terms. For English books (condition 2), subjects tended to adopt the color type of terms significantly ($\chi^2 = 5.851, p < 0.05$). As a result, hypothesis 3 is valid.

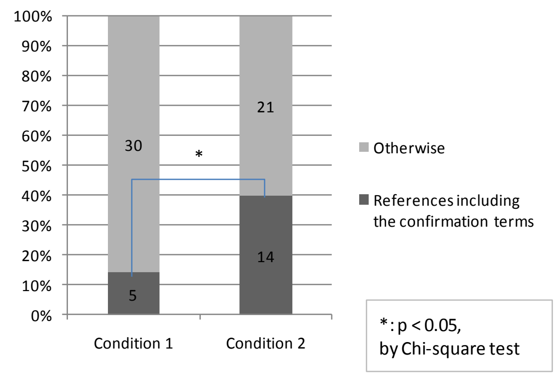

Fig. 7 Proportion of references that contain the confirmation terms when Japanese subjects referred to Japanese books (Condition 1) and to English books (Condition 2).

D. *Experiment 4 (within-subjects)*

Analyzing the 35 references from the 7 subjects revealed that subjects were more likely to adopt the confirmation terms when the books were removed and not around them ($\chi^2 = 8.100, p < 0.01$). Figure 8 shows the proportions of references that include the confirmation terms. The results prove the validity of hypothesis 4.

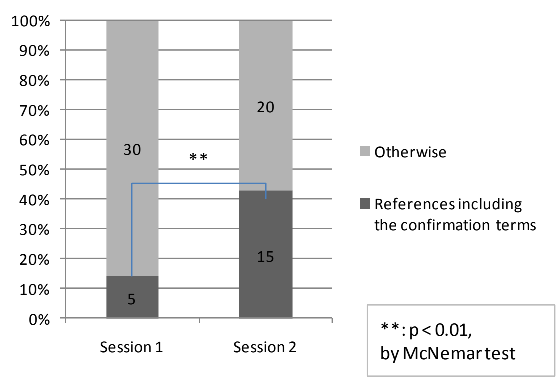

Fig. 8 Proportion of references that contain the confirmation terms before the books were removed (Session 1) and after (Session 2).

E. *Experiment 5 (between-subjects)*

The subjects were the same as those who participated in Experiment 3. Figure 9 shows the proportions of references that include the color confirmation terms. When the books were removed, the references of the English books included more of the confirmation terms than did those of the Japanese books ($\chi^2 = 8.571, p < 0.01$). As a result of the experiment, the validity of hypothesis 5 is proven.

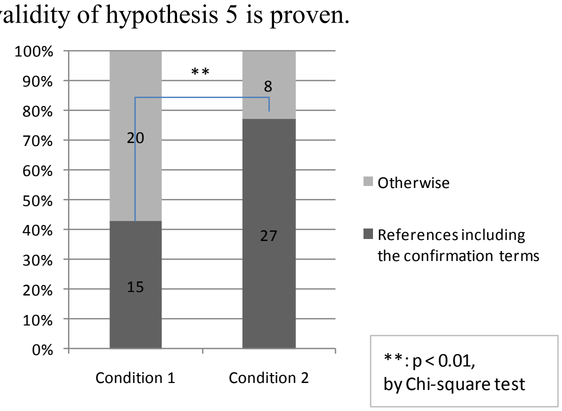

Fig. 9 Proportion of references that contain color words when subjects refer to Japanese books (Condition 1) and to English books (Condition 2) that have been removed.

V. Discussion

A. *Lexical entrainment in human-robot interaction*

The results provided strong support for our five hypotheses on lexical entrainment in human-robot interaction. We believe that subjects were induced to use the confirmation terms in their references because the results of Experiment 1 presented a significant difference in the number of references that include the confirmation terms between the sessions. The phenomenon was simple lexical entrainment in which subjects repeated terms used by the robot. This is similar to lexical entrainment observed in human-computer interaction. Meanwhile, according to the results of Experiment 2, the robot induced subjects to use color terms by limiting the confirmation terms to the color type even when they referred to a different book than the one they had chosen. This suggests that subjects were likely to not only repeat the confirmative terms but also take account of the type of term. Judging from the above, we may make the following two assumptions about lexical entrainment in human-robot interaction.

- **Entrainment per term**: People adopt a term of the robot in the next reference of the object.
- **Entrainment per type of term**: People adopt the same type of term as that used by the robot.

These two kinds of lexical entrainment will be useful in speech recognition. The robot will lead the user to say terms that are recognizable to the robot by uttering confirmation terms from the robot's dictionary for speech recognition. Furthermore, the robot can narrow the range of the type of referential terms chosen by the user if the robot limits the types of confirmative terms.

We illustrate in Figure 10 the proportions of the types of references used in Experiments 3, 4 and 5. When subjects referred to Japanese books, the proportion of references by book title in each session was about 30%. In particular, some subjects who used book titles in session 1 were likely to use book titles in session 2 as well. Meanwhile, other subjects who used book size or the reference terms in session 1 were influenced to use the confirmation terms in session 2. In the case of English books, both the subjects who use book titles and other terms in session 1 were more attracted to using the robot's confirmation terms in session 2. From these results, it may be difficult to lead subjects to use explicit terms that can distinguish an object uniquely from the confirmation terms. However, unmemorable or ambiguous terms in the confirmation terms of the robot may prove to be successful in evoking their use by the subject. This is a helpful suggestion because people do not always refer to an object by its official or formal name.

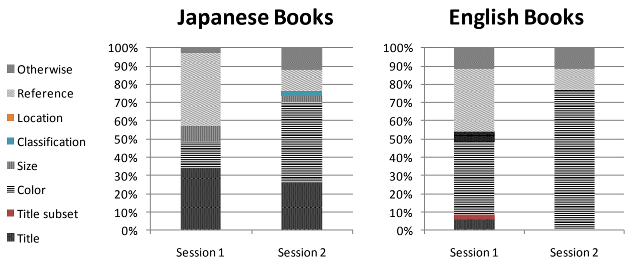

Fig. 10 Proportions of the types of references in Experiments 3, 4 and 5.

B. Limitations

We investigated lexical entrainment only in a book-reference task, so the generality of our findings might be limited. However, we believe that lexical entrainment will occur for various other objects as well from the results of Experiments 3, 4 and 5. These experiments showed that when a hidden book or an unfamiliar book was referred to, lexical entrainment was more likely to be observed. This tendency suggests that people are more likely to adopt terms that the robot uses when they do not understand an object clearly. An object whose name is unfamiliar might be more difficult to label with a proper name than a book, which has a title. Therefore, it is reasonable to assume that the validity we found for lexical entrainment with books will extend to various other objects.

VI. Test Bed System

Our results suggest that there are two types of entrainment in human-robot interaction: entrainment per term and entrainment per type of term. We believe that a robot can reduce the variability of reference terms by promoting lexical entrainment; therefore, we should discuss the framework for a system that can promote lexical entrainment and recognize a referred-to object automatically.

We conducted our experiments based on a Wizard-of-Oz method, where the operator manipulated the robot remotely. To automate the experimental system, we need to implement the following two functionalities: the recognition of a referred-to object and the production of a confirmation term (Fig. 11). In this work, we implemented the functionality of recognition.

To recognize a referred-to object, the following three kinds of recognition were implemented in the system: pointing-gesture recognition, gaze recognition, and speech recognition. For pointing-gesture recognition, we used a motion-capture system (Fig. 11), which could obtain 3-dimensional numerical data on a subject's motion from markers attached to the body. We also used a cap-style wearable eye-tracking device for capturing eye direction (Fig. 12). To recognize a subject's speech, we used Julius, which is free speech-recognition software[33]; moreover, we attached a microphone to a subject to avoid the problem of acoustic noise.

Using these functionalities of recognition, we designed the system to recognize a referred-to object as follows. The system recognizes the direction of the pointing and gaze while the subject refers to an object, and then it selects the object having a high probability from the results of speech recognition and the distance relationship between the directional line and the positions of objects.

We investigated whether the system could recognize a referred-to book by using the same task as described above (Fig. 14). As a result, the system could recognize 22 out of the 24 references when two subjects referred to a book 12 times each.

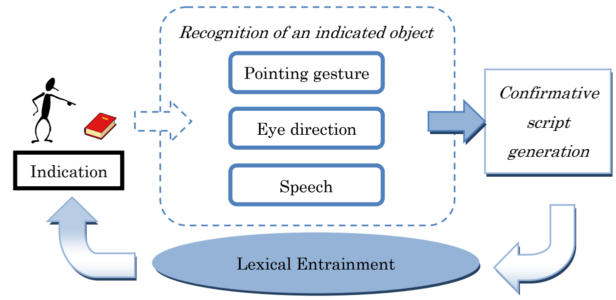

Fig. 11 System framework.

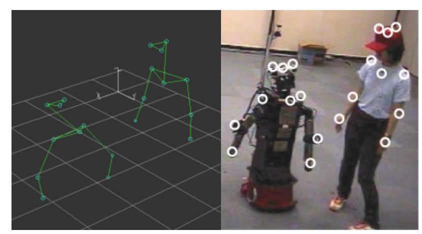

Fig. 12 Position information from the motion-capture system.

Fig. 13 Eye-tracking device.

Fig. 14 Robot system for recognizing a referred-to object.

The robot in the system confirmed a referred-to object with predetermined terms. Our experiment's results suggest that confirmation with a simple feature of the object promoted lexical entrainment. In future work, the system will have a function to automatically produce confirmation terms based on a simple feature corresponding to the recognized objects.

VII. Conclusions

We focused on lexical entrainment in human-robot interaction as an approach to improving the performance of recognizing a referred-to object. We tested hypotheses on lexical entrainment through experiments in which people referred to multiple objects via conversation with a robot. The experimental results reveal three important properties:

1. People adopt terms used by a robot in its confirmation utterances when they subsequently refer to an object.
2. People use the same type of terms in their references as the type used by a robot.
3. Lexical entrainment is accelerated when people refer to an object that is not around them or to an object that is unfamiliar to them.

We believe that our approach enables robots to improve recognition performance for a referred-to object. Moreover, we developed a test bed system for a robot to recognize a referred-to object, based on the results of our experiments, and verified that the system works.

Acknowledgments

We wish to thank Dr. Berthin and Mr. Yoshii for their help.

References

[1] Research study for the scope of the strategy map of robotics technology, 2008. (Available at http://www.nedo.go.jp/roadmap/2008/sys1.pdf). (in Japanese)

[2] H. Ishiguro, T. Ono, M. Imai, T. Maeda, T. Kanda and R. Nakatsu: “Robovie: an interactive humanoid robot,” Int. J. Industrial Robotics, vol. 28, no. 6, pp. 498-503, 2001.

[3] R. J. Fontana and S. J. Gunderson, Ultra-Wideband Precision Asset Location System, IEEE Conference on Ultra Wideband System and Technologies, pp. 147-150, 2002.

[4] M. Tüchler, V. Schwarz and A. Huber. Location accuracy of an UWB localization system in a multi-path environment, IEEE Conference on Ultra Wideband System and Technologies, pp. 414-419, 2005.

[5] A. Haasch, N. Hofemann, J. Fritsch and G. Sagerer: “A multi-modal object attention system for a mobile robot,” IEEE/RSJ Int. Conf. Intelligent Robots and Systems, pp. 1499-1504, 2005.

[6] G. W. Furnas, T. K. Landauer, L. M. Gomez and S. T. Dumais: “The vocabulary problem in human-system communications,” Communications of the ACM, vol. 30, pp. 964-971, 1987.

[7] S. Garrod and A. Anderson. Saying what you mean in dialog: A study in conceptual and semantic co-ordination. Cognition 27, 1987.

[8] S. E. Brennan and H. H. Clark: “Lexical choice and conceptual pacts in conversation,” Journal of Experimental Psychology: Learning, Memory, and Cognition, vol. 22, pp. 1482-1493, 1996.

[9] R. Porzel et al. "How Entrainment Increases Dialogical Effectiveness," in Proceedings of the IUI'06 Workshop on Effective Multimodal Dialogue Interaction, 2006.

[10] T. Kanda, M. Shiomi, L. Perrin, T. Nomura, H. Ishiguro and N. Hagita, “Analysis of people trajectories with ubiquitous sensors in a science museum,” Proc. 2007 IEEE Int. Conf. Robotics and Automation (ICRA’07), no. FrE8.4, pp. 4846-4853, 2007.

[11] K. Shinozawa, T. Miyashita, M. Kakio and N. Hagita: “User specification method and humanoid confirmation behavior,” IEEE International Conference on Humanoid Robots, 2007

[12] S. E. Brennan: “Lexical Entrainment in spontaneous dialog.” In Proceedings of the International Symposium on Spoken Dialogue, pp. 41-44, 1996.

[13] J. Gustafson, A. Larsson, R. Carlson and K. Hellman: “How Do System Questions Influence Lexical Choices in User Answers?” In Proceedings of Eurospeech, Rhodes, Greece, pp. 2275-2278, 1997.

[14] S. Tomko and R. Rosenfeld: “Shaping spoken input in user-initiative systems,” In The 8th International Conference on Spoken Language Processing, pp. 2825-2828, 2004.

[15] T. Kanda, H. Ishiguro, M. Imai and T. Ono, Development and Evaluation of Interactive Humanoid Robots, Proceedings of the IEEE, Vol. 92, No. 11, pp. 1839-1850, 2004.

[16] T. Kanda et al. A tension-moderating mechanism for promoting speech-based human-robot interaction, IEEE/RSJ International Conference on Intelligent Robots and Systems, pp. 527-532, 2005.

[17] K. Shinozawa et al. Difference in effect of robot and screen agent recommendations on human decision-making, International Journal of Human-Computer Studies, Vol. 62, pp. 267-279, 2005.

[18] Z. M. Hanafiah, C. Yamazaki, A. Nakamura and Y. Kuno: ”Understanding Inexplicit Utterances Using Vision for Helper Robots,” Int. Conf. on Pattern Recognition, 2004.

[19] T. Inamura, M. Inaba and H. Inoue: “PEXIS: Probabilistic Experience Representation Based Adaptive Interaction System for Personal Robots,” Systems and Computers in Japan, Vol. 35, No. 6, pp. 98-109, 2004.

[20] B. Scassellati: “Investigating Models of Social Development Using a Humanoid Robot,” Biorobotics, MIT Press, 2000.

[21] C. Breazeal, C. D. Kidd, A. L. Thomaz, G. Hoffman and M. Berlin: Effects of nonverbal communication on efficiency and robustness in human-robot teamwork,” Int. Conf. on Intelligent Robots and Systems, pp. 383-388, 2005.

[22] T. I. Carlos et al., A Robust Speech Recognition System for Communication Robots in Noisy Environments, IEEE Transactions on Robotics, 24(3), pp. 759-763, 2008.

[23] K. Nakadai et al. Real-Time Auditory and Visual Multiple-Object Tracking for Robots, Int. Joint Conf. on Artificial Intelligence (IJCAI’01), pp. 1425-1432, 2001.

[24] Y. Matsusaka et al., Multi-person Conversation Robot using Multimodal Interface, Proc. World Multiconference on Systems, Cybernetics and Informatics, pp. 450-455, 1999.

[25] S. Yamamoto et al. “Making a robot recognize three simultaneous sentences in real-time,” IEEE/RSJ Int. Conf. on Intelligent Robots and Systems, pp. 897-902, 2005.

[26] W. S. Condon and L. S. Sander, “Neonate movement is synchronized with adult speech: Interactional participation and language acquisition,” Science, 1974, 183, pp. 99-101.

[27] A. E. Scheflen, “The significance of posture in communication systems,” Psychiatry, 27, pp. 316-331.

[28] A. Kendon, “Movement coordination in social interaction: some examples described,” 1970, Acta Psychologica, 32, pp. 1-25

[29] M. Imai, T. Ono, and H. Ishiguro. “Physical relation and expression: Joint attention for human-robot interaction,” IEEE Transaction on Industrial Electronics, Vol. 50, No. 4, pp. 636-643, 2003.

[30] T. Ono, M. Imai and H. Ishiguro: A Model of Embodied Communications with Gestures between Humans and Robots, Proc. of Twenty-third Annual Meeting of the Cognitive Science Society, pp. 732-737, 2001.

[31] C. Breazeal “Regulation and entrainment in human-robot interaction,” International Journal of Robotics Research, 21 (10-11), 883-902, 2002.

[32] H. Ogawa and T. Watanave, “InterRobot: speech-driven embodiment interaction robot,” Advanced Robotics, Vol. 15, No. 3, pp. 371-377, 2001.

[33] Julius, http://julius.sourceforge.jp/en/julius.html