Effects of Robots' Revoicing on Preparation for Future Learning

Hajime Shirouzu, NIER, 3-2-2 Kasumigaseki, Chiyoda-ku, Tokyo, shirouzu@nier.go.jp
Naomi Miyake, University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo, nmiyake@p.u-tokyo.ac.jp

Abstract: In the method of *preparation for future learning*, learners often engage in *constructive interaction* with expressing, listening to and integrating their own multiple voices. In order to identify a specific discourse structure underlying successful collaboration, we used a remotely controllable robot as a member of a small discussion group of college students who solved a challenging physics problem. For the robot to act as a listener who solicits voices from students in the group, we manipulated its ways of "revoicing": it performed *minimum revoicing* of students' keywords without evaluative comments in one condition and *guiding revoicing* towards scientific models in the other condition. Comparing these two conditions in addition to a human-only condition, we found that the robot's minimum revoicing fostered students' agency and reflection on their mental models, which prepared them to learn from a lecture and solve a transfer problem. The role of listener for PFL was discussed.

Introduction

In this paper, we examine if there is a specific structure of discourse, i.e., *multi-vocal* structure, in successful collaborative learning using a remotely operate-able robot. In successful classrooms, learners often expand their understanding adaptively by revising their folk knowledge into scientific understandings. This process often requires collaborative forms in which learners express their own multiple thoughts (voices), exchange and integrate them in *constructive interaction* (Miyake, 1986; Shirouzu, Miyake & Masukawa, 2002). As one of forms, we use the learning method of *preparation for future learning* (PFL) (Schwartz & Martin, 2004) or *productive failure* (Kapur & Bielaczyc, 2012) because it is now widely practiced and yields positive results in classrooms. With this method, researchers set up a small-group "groping" activity (e.g., ill-defined problem-solving) prior to a lecture or whole-class discussion, and examine how such an activity prepares students for future learning from instructional resources. Their analyses indicated that the groping fosters students' epistemic agency and drives their extensive search in the problem space. They also implied that constructive interaction took place there. We can identify constructive interaction in Schwartz and Martin's conversation excerpt of ninth graders who tried to invent a measure of reliability for pitching machines. One student first proposed a solution, and the other two students monitored it, commenting on its limitations and flaws, which triggered the first one's reflection. In this way, they took turns acting as *task-doer (talker)* and *monitor (listener)*, leading to the constructive proposal of newer and more general solutions, even if they did not reach canonical ones. In order to trigger such interaction rather than waiting for it to occur, we could put a listener into a group because the listener solicits others' active explanations and makes comments that foster talkers' reflection on their ideas.

Why do we try to put a listener instead of a leader who facilitates and demonstrates how to collaborate? It is because we believe all individuals have the potential to deepen their understanding through collaboration. Besides, observation of well-designed collaborative learning reveals that students who stay mostly silent during group work still listen to others' words and say simple but important words that provoke discussion and perspective shifting among the members. This concept is similar to Greeno and van de Sande's (2007) *constructive listening* and Engle and Conant's (2002) *problematizing*. However, such observation depends on "emergent" group conversation, which we do not know how to replicate systematically.

Revoicing is a good way both to demonstrate active listening and to mark ideas worth reflecting. In this study, robots react to paired students' discussions with different ways of revoicing. Revoicing originally refers to a teacher's repeating or rephrasing students' utterances. Many studies have focused on the pedagogical features of revoicing (e.g., guiding students' attention to critical contents, giving ownership to them, and positioning them to each other) (O'Conner & Michaels, 1996). In order to make the robot a listener, however, we must use student-like rather than teacher-like revoicing. Thus, our robots perform *minimum, stingy revoicing* that simply repeats students' utterances of keywords, regardless of their correctness. We compare this "minimum revoicing condition" with a "guiding revoicing condition" in which the robot adds evaluative phrases to keywords in order to guide students' attention to critical contents. At a glance, the latter condition seems to outperform the former, but we aim to demonstrate that the opposite is the case, since teacher, leader-like revoicing makes students passive and dependent on the robot, rather than promoting their own problem-solving.

If revoicing also serves to mark important ideas, it may provoke conflicts among contrasting ideas of students who engage in PFL tasks. The remaining question about PFL is how collaborative groping leads to learning from a lecture: specifically, whether (1) the groping only has to raise students' agency, and the quality of the search does not need to be high, or (2) the search must cover critical points and raise their awareness closest to the threshold for receiving the lecture. By using a challenging physics problem as a groping task that solicits conflicting ideas (*mental models*), we analyze the collaborative process as search, proposal, question,

and criticism of mental models, and the effects of a robot's revoicing on this process. We also prepare a "human-only condition" of groups of three students and compare it with the minimum revoicing condition to see if a robot that does not express its own ideas but just listens and re-utters helps students' construction of mental models by problematizing the differences among their ideas and keeping their multi-vocality.

Method

Participants

Forty-five undergraduates of a Japanese university participated. We assigned six pairs (12 students) each to the minimum revoicing condition and the guiding revoicing condition, and seven groups of three students (a total of 21 students) to the control (human-only) condition. Members in a pair or group knew each other well. None of the students knew the tasks, and there were no remarkable differences in their physics ability among the groups. We used the desktop robot "Robovie-W" for the experiment. It was 30cm tall and had 17 degrees of freedom with a built-in camera, speaker, and microphone. Robot utterances and actions can be generated and adjusted with a remote control by an operator, and utterances are achieved by voice-synthesizing software (XIMERA).

Tasks

Bobbin Problem

We chose the "bobbin problem" from Anzai and Yokoyama (1984) as the main task. The problem is multiple-choice, to predict the direction of movement of a bobbin as illustrated in Fig. 1. It is quite difficult to determine the correct answer as (1). More than 90% of university students as physics beginners select incorrect answers (2) and (3). According to Anzai and Yokoyama's analysis, beginners tend to make mistakes by drawing on their experience that the bobbin rolls clockwise when the string is pulled from the "fixed" center of the bobbin.

"The centers of two circular frames are interconnected by an axle, and a string is wound around it like a bobbin, as illustrated in the figure below. What will happen if you pull the string as shown in the figure? The discs may roll, but never slide. Mark the number that you think is correct. Let's discuss your reason for selecting it."

> (1) The bobbin rolls to the left (counterclockwise).
> (2) The bobbin rolls to the right (clockwise).
> (3) The bobbin does not move.
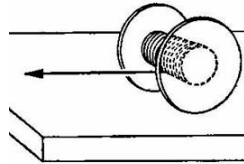> (4) Other. (Write your answer concretely.)

Fig. 1. Bobbin problem.

Transfer Problem: Toilet Paper Problem

We selected a transfer problem from Anzai (1991) (hereafter "the toilet paper problem"). This problem questions the movement of an object when the positions of the fulcrum and the power point are changed from those depicted in Fig. 1 and pulled as shown in Fig. 2. The correct answer is (2), but the participant who simply assumes that the object always rolls to the left when the string is pulled to the left will select the wrong answer.

"An axle was passed through the center of an object and a string was wound on it, as depicted in the figure below. What will happen if you pull the string as shown in the figure? The object may roll, but never slide. Mark the number you think correct, and discuss your reason for selecting it."

> (1) The object rolls to the left (counterclockwise).
> (2) The object rolls to the right (clockwise).
> (3) The object does not move.
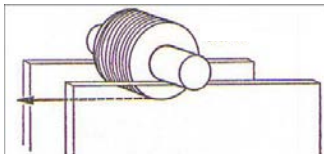> (4) Other. (Write your answer concretely.)

Fig. 2. Transfer problem: Toilet paper problem.

Lecture

For the bobbin problem, we delivered a lecture, the key points of which are as follows: "First, let's regard the contact point between the bobbin and desk as the fulcrum, and the point from where the string comes as the power point (Fig. 3a). Next, consider the line segment that passes both of these points as the axis of the bobbin, and you may understand more easily that the line segment falls to the left around the fulcrum (Fig. 3b). However, the bobbin is circular, not being composed of a single line segment. Therefore, if you assume the circle contains innumerable line segments, you can see the axes will fall to the left, one after another. As a result, the circle rolls to the left (Fig. 3c)." In order to provide exactly the same explanation for every condition, we recorded the lecture beforehand, cut it, and attached voice files with a total of 13 power-point slides.
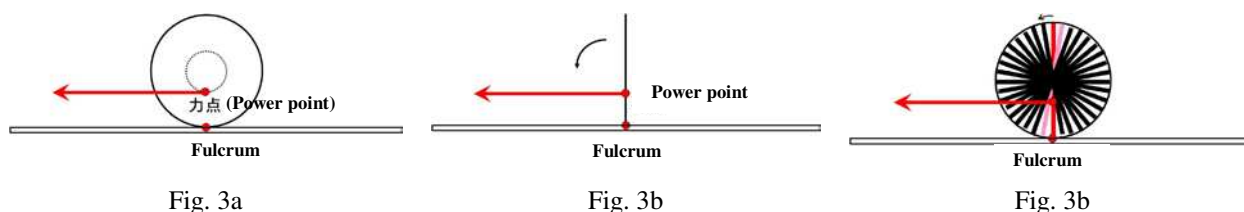
| Fig. 3a | Fig. 3b | Fig. 3b |

Fig. 3. Explanation slides for the bobbin problem.

Procedure

Minimum Revoicing Condition

First, each pair of students had 5min to interact freely with the robot. For ice-breaking, the robot asked the names of students and called them by their names, and they exchanged information about their origins. Other operations were left to the discretion of the operator of the remote control. After a brief explanation of the experiment, we distributed sheets of paper to the students describing the bobbin problem in Fig. 1. The students attempted to solve the problem for 10min while talking freely. The robot first asked students to read the problem aloud, in order to enhance their relationship. In the prediction phase, the robot revoiced the students' discussion, the details of which are presented in the next section. In the discussion phase, we distributed sewing machine bobbins. After conducting the experiment and confirming the result, the students discussed reasons for 6min (Fig. 4a). The robot also performed revoicing. An atmosphere was created so that the robot observed the experiment as the students discussed while pulling the bobbin and participated in their discussions. Thereafter, the students listened to the lecture for 5min, as displayed on the screen. The robot also turned its face to the screen to create the appearance of viewing it with the students (Fig. 4b). Upon completion of the explanations, the robot turned its face to the students again. If the students did not engage in discussion, it questioned, "Did you understand the explanations? Why did it go to the left?" If they were discussing the explanations, the robot waited a while and then asked these questions. After receiving answers from the students, the robot questioned, "What are the key points or keywords?" When only one student answered the first question, the robot asked the member who did not answer. Lastly, the robot asked students for summarization, "The lecturer mentioned the fulcrum, power point, and circle, didn't he? How did they relate to this problem?" Five minutes was allowed for the above discussion. Upon completion of discussion, power to the robot was turned off.

Next, we distributed sheets of paper describing the toilet-paper problem in Fig. 2. The students were told to talk freely for 5min, and to write down the answer and the reason for it. If the group did not reach a consensus, they were allowed different answers. These answers provide with primary data for deciding their performance of the transfer problem. After the teacher demonstrated the answer using toilet paper in front of the classroom, students spent another 5min discussing the reason and then answered the question again. Lastly, questionnaires and interviews about the impression of the robot were conducted for 10min. The questionnaires contained six items (e.g., "Do you want to learn together with Robovie again?" "Did you think that Robovie knew the answer to the bobbin problem?" and "What partner do you think Robovie would be if it were human?"). All these processes were recorded on video, ICR, and the log of the remote-control system. The groups were separated far enough from each other for clear recording and avoidance of contamination.

| Fig. 4a. Discussion | Fig. 4b. Listening to lecture |

Fig. 4. Humans and the robot on the bobbin problem.

Guiding Revoicing Condition

The procedures for the guiding revoicing condition were the same as those for the minimum revoicing condition, except for the method of revoicing, as described in the next section.

Control (Human-Only) Condition

The procedures for the control condition were the same as for the two conditions above, except that three students formed a group and the prompts for discussion after the lecture were given on a piece of paper. Questions about "the reason for rolling to the left," "key points/words," and "relationships among the fulcrum, power point, and circle" were printed, and students were instructed to discuss them at their own pace.

Experiment Manipulation: Robot's Revoicing

The robot merely repeated the keywords that students mentioned under the minimum revoicing condition; however, it reacted affirmatively to correct keywords and negatively to incorrect keywords as terms in physics under the guiding revoicing condition. For example, under the minimum revoicing condition if students said "It rolls to the right" (wrong prediction), the robot said only "Rolls to the right"; however, under the guiding revoicing condition it suggested a negative evaluation by adding a phrase like "'Does it roll to the right?"

We implemented revoicing in two ways: the operator pressed the button for a preloaded input prompt, or input the prompt by text. To load the keywords that would appear frequently in students' discussions into the operation system, we observed 24 extra juniors as they discussed the bobbin problem prior to this experiment in a similar experiment procedure. We identified frequently used words both in prediction and discussion phases. These "hot" words included correct physical terms, incorrect ones, and unrelated ones. The left-hand column of Table 1 lists candidate keywords for preloading in the prediction phase under the minimum revoicing condition. Under the guiding revoicing condition, we added phrases (in red, underlined letters in the right-hand column of Table 1) to these keywords (most of these phrases were expressed by Japanese sentence-ending particles such as *yo*, *ne*, *yone*, or *kana*, and were difficult to translate). We added three utterances of "evasion" to avoid questions from students in order to position the students, rather than the robots, as task-doers, along with three utterances for "the answer" and 15 utterances for "reasons." Eight of these utterances were manipulated across the conditions. For the discussion phase, we prepared 20 preliminary utterances for the minimum revoicing condition (1 for "answer" ("Rolls to the left!"), 12 for "reasons," 3 for "evasion," and 4 for "emotional expression"). Utterances of emotion were expression of surprise at the experiment result (e.g., "Oh, why?"). Because they were unnatural as utterances of the guiding-revoicing robot, two of the four utterances were deleted. Therefore, 18 utterances were used for the guiding revoicing condition. Among these, seven were manipulated across the conditions. We manipulated a total of 15 of 41 utterances (36.6%) in sum.

Table 1. Examples of revoicing utterances in the prediction phase of the bobbin problem.

	Minimum revoicing condition			Guiding-revoicing condition		
Ans-wer	Rolls to the left			Rolls to the left		
	Rolls to the right			<u>Does it</u> roll to the right?		
	Does not move			<u>It may</u> move (only when students said "move")		
Rea-son	Pull to the left			<u>Yes, we</u> pull it to the left		
	It goes forward			It goes forward, <u>doesn't it?</u>		
	The string goes out		Pull it upward	<u>Does</u> the string go out?		<u>Do you</u> pull it upward?
	Friction		Force	<u>Is it</u> friction?		<u>Is it</u> force?
	Rotates		It will come here	Rotates		It will come here
	Does not slide		If it does not slide	Does not slide		If it does not slide
	Pull the string	Pull	When pulled	Pull the string	Pull	When pulled
	The string comes from the bottom			The string comes from the bottom		
	Pull in parallel with the desk			Pull in parallel with the desk		
Eva-sion	Well, I can't understand. Please think together.			Well, I can't understand. Please think together.		
	Well, how does it move?			Well, how does it move?		
	I haven't decided the answer. Let me think more.			I haven't decided the answer. Let me think more.		

For free utterances, the minimum revoicing condition allowed free utterances (inputs) on the principle that "when a cluster of utterances is heard and the two students become silent, the robot will speak the keywords of each cluster." Under the guiding revoicing condition, the robot also performed utterances on the principle that "it may guide students to the correct answer by adding an evaluating phrase to the keyword, but never tell the answer." We appointed two postgraduates and four undergraduates who were familiar with the bobbin problem and well-trained as remote-control operators. To minimize the differences between the conditions, we had them take charge of the remote control under one condition as many times as under the other condition. Actually, no large difference in either the numbers or the ratios of preliminarily input sentences and freely input sentences was observed between the conditions. Revoicings were made 8.8 times in the minimum revoicing condition and 7.5 times in the guiding revoicing condition on average per pair in the experiment.

Results

In this section, we first confirm students' performance of the transfer task, then examine their feeling of agency through their perception of the robot, and finally analyze the collaborative process and the robots' effect on it.

Performance of the Transfer Problem

We examined the ratios of students who were able to correctly predict the transfer (toilet paper) problem and those who were able to explain the reason, based on the relationship between the fulcrum and the power point as described in the Lecture section (e.g., "The axis falls to the right because the power point is located under the fulcrum."). Fig. 5 indicates that the correct answer ratio was the highest under the guiding revoicing condition, followed by the minimum revoicing condition and then the human-only condition, but the difference is not significant. However, the ratio of description of the correct reason under the minimum revoicing condition exceeded those under other conditions. The chi square test indicated significant differences among conditions ($\chi^2(2) = 6.69$, $p < .05$), with residual analysis indicating that more students in the minimum revoicing condition gave correct explanations than expected. Considering that no students in the minimum revoicing condition correctly predicted the bobbin problem, they had no superior prior knowledge, but learned from the discussions and lecture, and transferred their knowledge appropriately to the subsequent, toilet-paper problem.
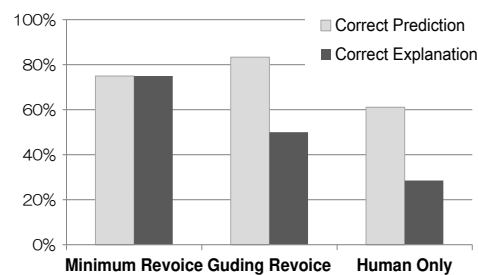

Fig. 5. Results by conditions on the transfer problem.

Perceived Roles of the Robot

Next, we examined the results of post-questionnaires and interviews to determine the students' perception of the robot. First, the ratio of students who felt that the robot knew the answer was higher under the guiding revoicing condition (75%) than under the minimum revoicing condition (58%). We classified their perceptions of the robot into five categories using their literal expressions (Table 2). While 58% indicated that the robot was like their friend under the minimum revoicing condition, nobody answered this way under the guiding revoicing condition. Instead, the robot was perceived as a heterogeneous, obtrusive being, such as a teacher or facilitator who gave guidance or as a child "who always asks 'why?'" Corresponding to this result, 83% under the minimum revoicing condition indicated that they wanted to learn with the robot again, whereas only 50% did so under the guiding revoicing condition. In summary, the students in the minimum revoicing condition tended to perceive the robot as their thinking partner who did not know the answer. The students did not seem to rely on the robot but to think of themselves as epistemic agents.

Table 2. Perceived roles of the robot (Number of persons).

	Teacher	Facilitator	Student/Friend	Listener	Children
Minimum	1	3	7	1	0
Guiding	2	4	0	3	3

Collaborative Groping Processes

These results imply that the subtle difference in revoicing made all the difference in test performance and perception of the robot. In order to determine how the students deepened their understanding through interaction with the robot, we analyzed the process of collaborative groping before the lecture. Tables 3 presents excerpts of typical pairs under the minimum revoicing condition, and Table 4 presents those under the guiding revoicing condition. Before tracing them in detail, we explain *types of utterance* and *mental models*.

Types of Utterance: We roughly classified students' and robots' utterances into six types: the students' "questioning" of answers to the robot, their "reasoning" of answers and justification, "observation" of experiment results, "revoicing-back (re-uttering)" to the robot's revoicing (e.g., line 20 of Table 3), and the robot's "evasion" and "revoicing." Using these types, we could determine the discourse pattern among the students and the robot (e.g., how the robot avoided students' questioning and let them act as reasoners).

Mental Models of the Bobbin Problem: We coded utterances and gestures of students' "reasoning" into five mental models (and their variations) in Fig. 6. These models are developed from Anzai and Yokoyama (1984) and our observations. The *tension-force model* (Fig. 6a) simply considers the tension force in the string that pulls the bobbin to the left (broad arrow in Fig. 6), resulting in the correct answer. Although this model has a weakness of not specifying how the string behaves as the bobbin moves, we often observed that it re-emerged

in students' discussions as if they could not ignore it to the end. The *rotation-only model* (Fig. 6b) exploits experiential knowledge of something rotating around a fixed axle (e.g., toilet paper). This model does not predict the direction of the bobbin movement. Expanding the rotation-only model, the *rotation-to-right model* (Fig. 6c) assumes that the rotation creates the force to move to the right, resulting in the wrong answer. The tension-force model and the rotate-to-right model predict opposite directions of the bobbin; however, the students referred to them naturally, as if they believed both. Thus, these models are *fragmental*, *multiple*, *and autonomous objects* (Williams, Holland & Stevens, 1983). However, the students sometimes struggled to integrate them, through which two other models developed. One is the *correct model* (Fig. 6d), which predicts that the bobbin rotates to the left with winding the string. Interestingly, students often referred to it while denying its possibility. In this sense, this model is not fully scientific but only phenomenological, yet closest to scientific models such as that delivered by the lecture. The other is the *force-balance model* (Fig. 6e), which claims that the tension force balanced with the rotation (or friction) force, resulting in no movement (wrong answer). Using this framework, we can analyze how widely the students explored the problem space and how closely they approached the scientific models.
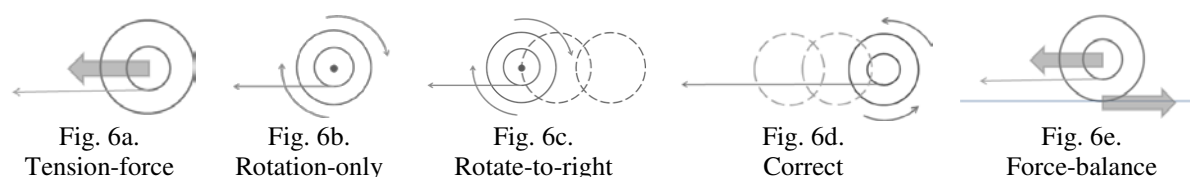
| Fig. 6a. Tension-force | Fig. 6b. Rotation-only | Fig. 6c. Rotate-to-right | Fig. 6d. Correct | Fig. 6e. Force-balance |

Fig. 6. Mental models of the bobbin problem.

Pair in the Minimum-Revoicing Condition: As indicated in Table 3, Pair A under the minimum revoicing condition, especially Student A1, first asked the robot for the answer, but the robot gave evasive responses (lines 1-6). The students then began to reason and express various models (lines 7-11), which caused conflicts between the models (italicized in lines 10-11). After the robot gave evasive responses again (lines 12-13), A1 developed his model into the rotate-to-right model, against which Student A2 protested from using the tension-force model (lines 14-15). Here, the robot first revoiced keywords, which were revoiced back by Student A1 and integrated into his explanation (underlined in lines 18-20). That is, A1 used the robot's subjective mode (if-clause) at line 19, and continued reasoning about the importance of the rotation (since preloaded revoicings are not always exactly the same as students' expressions, some students utilized such *new* forms of expressions). As a result, Student A1 leaned to the rotate-to-right model; in contrast, A2 seemed to focus on the tension force (line 15) and simulated the leftward movement again (lines 21-22), resulting in externalization of the correct model. Even though he might not have believed this model, it is important that the students were exposed to it. As a result, they focused on the movement of the string as soon as they observed the experiment results (line 102), the utterance of which was also revoiced by the robot and revoiced back by themselves (lines 103-104) and developed into another model (line 105). In summary, the robot's revoicing neither happened often nor guided students to scientific models (line 20); instead, it indirectly problematized the difference and conflict among the models and pushed the students towards integration.

Pair in the Guiding-Revoicing Condition: In contrast, Pair B under the guiding revoicing condition first talked about reasons (Table 4, lines 1-2); however, with the robot's guiding revoicing at line 6, Student B1 noticed the sentence ending particle of the question ("kana") and interpreted that the robot knew the answer (line 9). They then engaged in obtaining the answer from the robot.

In sum, the interactional patterns of both pairs rapidly formed through interactions in the early stage of the prediction phase, where participants neither knew the correct answer nor knew if the robot knew it, and might expect to gain it from the robot. In that stage, guiding revoicing tended to deprive the students of agency.

Mental Models and Subsequent Learning

We analyzed the quantity and quality of mental models to which the students referred. As indicated in Table 3, Pair A referred to more kinds of models more times (seven kinds, ten times) than Pair B in Table 4 (one kind, two times). Qualitatively, Pair A questioned the rotate-to-right model and referred to the correct model, which was not observed in Pair B. Table 5 summarizes all the results of the three conditions during the prediction and discussion phases, broken into correct and incorrect explainers of the transfer problem. The number indicates the average per student, and that in parentheses indicates standard deviation per student. "Referring to Correct model or/and Questioning Rotate-to-right Model" means percentages of the pairs/groups who referred to the correct model or questioned the rotate-to-right model in the total of pairs/groups in each cell. Because such references were infrequent and we assumed that exposure to them was important, we took the pair or group as a unit of analysis and examined whether at least one member made such references. We also grouped a pair or group as "correct" if at least one member correctly explained the transfer problem.

Table 3. Collaborative groping by a pair under the minimum revoicing condition.

Line	Talker	Utterance	Utterance type	Mental model
1	A1	What do you think?	Questioning	
2	R	Well, I can't understand. Please think together.	Evasion	
3	A1	Do you think if it rolls to the left?	Questioning	
4	R	Well, how does it move?	Evasion	
5	A1	Does it roll to the right?	Questioning	
6	R	Well, how does it move?	Evasion	
7	A1	If it doesn't rotate, it comes… [gesture]	Reasoning	Tension-force
8	A2	The problem says "it never slides." [gesturing the leftward-move]	Reasoning	Tension-force
9	A2	When we pull the string this way, the bobbin rotates like this…	Reasoning	Rotation-only
10	A1	And the string pulls the bobbin this way?	Reasoning	*Tension-force?*
11	A2	Coming with rotating like this? [gesturing the bobbin's leftward movement with the string coming untied]	Reasoning	*Tension-force + Rotation-only?*
12	A1	Do you think the bobbin remains still?	Questioning	
13	R	Well, how does it move?	Evasion	
14	A1	The bobbin rotates like this to that way. [gesture]	Reasoning	Rotate-to-right
15	A2	But the pulling force works this way. [gesture]	Reasoning	Tension-force
18	A1	Because the problem says "it never slides," so…	Reasoning	
19	R	If it does not slide	Revoicing	
20	A1	If it does not slide, the point is the rotation of the axle	Revoicing-back	
21	A2	If we pull the string, it rotates like this,	Reasoning	Rotation-only
22	A2	unless the bobbin rotates like this to the left [gesturing the bobbin's leftward movement with winding the string]	Reasoning	*Correct*
		<At the discussion phase (after the bobbin delivered)>		
102	A1/A2	Oh, it (the string) winds!	Observation	
103	R	The string gets wound.	Revoicing	
104	A1/A2	Yes, it gets wound…	Revoicing-back	
105	A2	There are only two forces leftwards and downwards, so it comes...	Reasoning	To Correct one

Table 4. Collaborative groping by a pair under the guiding revoicing condition.

Line	Talker	Utterance	Utterance type	Mental model
1	B1	If it does not slide, it goes to the right.	Reasoning	Rotate-to-right
2	B2	Because the string is wound clockwise, when we pull it, the bobbin goes that way.	Reasoning	Rotate-to-right
3	B1	The answer may be remains still or moves to the right. Do you think if it remains still?	Questioning	
4	R	Well, how does it move?	Evasion	
5	B1	Or, rolls to the right.	Questioning	
6	R	Does it roll to the right? (Migi-ni-korogaru-kana)	Guiding revoicing	
7	B1	Does it? (kana)	Questioning	
8	R	Well, how does it move?	Evasion	
9	B1	You said the answer is the right!	Blaming	

Table 5. Mental models referred by correct and incorrect explainer of each condition.

Condition	Kinds	Times	Referring to Correct model and/or Questioning Rotate-to-right Model
Minimum: Correct (n = 9)	3.78 (1.87)	6.22 (3.22)	80% (4 out of 5 pairs)
Minimum: Incorrect (n = 3)	2.67 (1.25)	3.33 (2.05)	0% (0 out of 1 pair)
Guiding: Correct (n = 6)	2.33 (1.11)	4.00 (2.74)	75% (3 out of 4 pairs)
Guiding: Incorrect (n = 6)	2.00 (1.00)	1.83 (1.07)	50% (1 out of 2 pairs)
Human: Correct (n = 6)	1.83 (0.69)	2.50 (0.76)	100% (3 out of 3 groups)
Human: Incorrect (n = 15)	1.60 (1.62)	2.00 (1.90)	25% (1 out of 4 groups)

As shown in Table 5, first, the numbers of models depended on the conditions, but the numbers of times of reference were constantly high for the correct explainers, regardless of the conditions (shaded in Table 5). Thus, successful members repeatedly referred to the particular models. Second, the successful pairs/groups

tended to question the wrong, rotate-to-right model and/or verbalize the correct model (also shaded in the table). These results imply that the quality of the collaborative groping should be as high as possible for receiving the lecture. In addition, the minimum revoicing condition in total contributed to raising this quality (see "*n*" of Table 5). Space prevents us from describing how the revoicing functioned in all the pairs, but we often observed "revoicing-back." For example, in one pair, when one member referred to the rotate-to-right model as "If we pull it this way, it produces the <u>force</u> for the bobbin to rotate to that way," the robot just revoiced "<u>Force</u>" and the other member said "<u>Force</u>, pull, force, pull… So if we pull it this way, the <u>force</u> works this way too." The latter member, having the tension-force model in mind, seemed to take advantage of the robot's revoicing and put that model on the table again, which kept their multi-vocality and sustained the discussion.

Discussion

This study included a robot in a discussion group of collaborative learning, and let it perform revoicing to the other members in a minimum, stingy way or in a facilitative, guiding way. We found that minimum-revoicing promoted students' performance of the transfer problem and egalitarian perception of the robot, which implies students' agency as *task-doers* of problem solving. Process analyses indicated that guiding-revoicing made the students perceive the robot as a *knower*; whereas minimum revoicing caused them to perceive it as a *co-solver*. Even though the minimum-revoicing robot provided no new information, it listened to students' words heedfully and revoiced them. Reciprocally, the students did not ignore the robot but listened to it and took advantage of the revoiced words. Since the task provoked multiple mental models, the students often confronted conflicts with one another. The robot's revoicing contributed to making explicit such conflicts, letting role-exchange happen in the students, and sustaining the discussion, which forced students to integrate these models.

This paper provides PFL researchers with the finding that collaborative groping should prepare students at the highest levels of understanding for receiving a lecture. Schwartz and Bransford (1998) once reported that groping activity creates "a time for telling" for teachers; however, this study implies that such a time may be very limited. A time for telling comes diversely, depending on students' readiness, and thus repeated chances to access instructional resources are needed.

The present study has many limitations in addition to the small sample size. The experiment situation in which the students first worked with the robot might cause them to focus too much on the robot's words (the revoicings were often made but ignored in the human-only condition), to which the second author is now conducting experiments that include successive opportunities for children to interact with the robot. Some operators reported difficulty in remote-controlling (e.g., when to avoid, when to revoice, and how to revoice). With more findings of HRI and HRL, we should design a longer collaborative learning environment.

References

Anzai, Y. (1991). Learning and use of representations for physics expertise. In K. A. Ericsson & J. Smith (Eds.), *Toward a General Theory of Expertise* (pp. 64-92). New York: Cambridge University Press.

Anzai, Y., & Yokoyama, T. (1984). Internal models in physics problem solving. *Cognition and Instruction*, 1, 397-450.

Engle, R. A., & Conant, F. R. (2002). Guiding principles for fostering productive disciplinary engagement. *Cognition and Instruction*, 20, 399–483.

Greeno, J., & van de Sande, C. (2007). Perspectival understanding of conceptions and conceptual growth in interaction. *Educational Psychologist*, 42, 9-23.

Kapur, M., & Bielaczyc, K. (2012). Designing for Productive Failure. *The Journal of the Learning Sciences*, 21, 45-83.

Miyake, N. (1986). Constructive interaction and the iterative process of understanding. *Cognitive Science*, 10, 151-177.

O'Conner, C., & Michaels, S. (1996). Shifting participant frameworks: Orchestrating thinking practices in group discussion. In D. Hicks (Ed.), *Discourse, learning and schooling* (pp. 63-103). New York: Cambridge University Press.

Schwartz, D. L., & Martin, T. (2004). Inventing to prepare for learning: The hidden efficiency of original student production in statistics instruction. *Cognition & Instruction*, 22, 129-184.

Shirouzu, H., Miyake, N., & Masukawa, H. (2002). Cognitively active externalization for situated reflection. *Cognitive Science*, 26, 469-501.

Williams, M. D., Hollan, J. D., & Stevens, A. L. (1983). Human reasoning about a simple physical system. In D. Gentner & A. L. Stevens, (Eds.), *Mental models* (pp.131-153). Hillsdale, N.J.: Lawrence Erlbaum Associates.

Acknowledgments