Using metaphors in collaborative problem solving: An eye-movement analysis

Asuka TERAI[1], Naoko KURIYAMA[1], Masaaki YASUHARA[1], Takenobu TOKUNAGA[1], Kimihiko YAMAGISHI[1] and Takashi KUSUMI[2]
Tokyo Institute of Technology[1], Japan, Kyoto University[2], Japan
asuka@nm.hum.titech.ac.jp, kuriyama@hum.titech.ac.jp, yasuhara@cl.cs.titech.ac.jp, take@cl.cs.titech.ac.jp, kimihiko@ky.hum.titech.ac.jp, kusumi@educ.kyoto-u.ac.jp

Abstract. The effectiveness of employing metaphors within problem solving is investigated quantitatively using eye-movement analysis. In the experiment, participant pairs attempted to solve Tangram puzzles. The pairs consisted of a solver who could see the goal shape but could not manipulate the pieces and an operator who could manipulate the pieces but could not see the goal shape. The analysis was conducted by comparing two groups of pairs contrasted in terms of their frequencies of using metaphors. The results show that (a) the rate of overlapping gazes and the ratio of the operator's speaking time were significantly lower for the low-frequency group, suggesting that the solvers in the low-frequency group could solve the puzzle with less communication with her/his operator; (b) in the high-frequency group, significantly more metaphors were used at the start of a trial to facilitate knowledge sharing for the goal shape; (c) metaphors evoke significantly longer periods of overlapping gazes between the participants for asymmetric goal shapes compared to symmetric goal shapes; and (d) the frequency of figurative utterances associated with overlapping gazes is lower than for literal utterances. These results indicate that metaphors do not function to convey information precisely but rather operate to facilitate bilateral communication.

Introduction

Metaphors are utilized in literature and in speeches for the purposes of emphasis, freshness and conciseness of expressions. Metaphors also aid comprehension and memorization within reading tasks (Reynolds and Schwartz, 1983). Furthermore, it has been shown that metaphors are effective in persuading people (Sopory and Dillard, 2002). Metaphors can, however, create misinterpretations between speakers and listeners due to a lack of shared background knowledge for conceptual structures (e.g. Lakoff and Johnson, 1980; Kusumi, 1995). Therefore, their use within collaborative work could be detrimental to task performance when there is an absence of shared knowledge between collaborators. For instance, (Kuriyama et al., 2007) reported that successful collaborative problem-solving requires the sharing of knowledge in terms of both language and culture in order to avoid the misinterpretation of metaphors on the part of the listener. Especially within collaborative problem-solving, clear communication might be required and misinterpretations should be avoided, so, metaphors may not always be effective.

Some studies have investigated the use of metaphors within collaborative tasks. For instance, (Nambu and Harada, 1998) investigated the relation between the use of metaphors and the quality of the communication channels. In their experiments, pairs of participants communicated through different channels that varied in quality to accomplish a figure-shape identification task. There were four levels of channel quality: direct-speech communication in the same room, communication over a telephone with three levels of noise. Even though task performance was not affected by the communication-channel conditions, more metaphors were used for the noisy channels. They concluded that higher cognitive costs are incurred in literally describing figure shapes in detail than explaining about them through metaphors. Therefore, their participants used metaphors in the noisy conditions to reduce their cognitive costs, despite the risks of listener misinterpretation.

A gender difference in terms of using metaphors has been reported by (Hussey and Katz, 2006). In a conversation task, the amount of metaphors produced by male pairs was not affected by the degree of intimacy with the partner, although female pairs tended to use more metaphors when talking to friends than to strangers. More literal expressions were observed between female strangers to avoid misinterpretations. Although these studies have investigated the tendency to use metaphors within a collaborative task, they have not, however, examined the effectiveness of metaphors for communication.

The role of metaphors within collaborative problem-solving was investigated using a puzzle task (Kuriyama et al., 2007). They suggested that metaphors effectively facilitated smooth communication and problem solving based on qualitative analysis of participant utterances. They reported that metaphors describing the goal image for the puzzle were used to construct a common foundation for knowledge sharing during the early stages of problem solving, while metaphors describing a part of the goal figure were used to explain details during the later stages of problem solving. However, their study did not quantitatively examine the effectiveness of metaphor usage.

In order to quantitatively investigate the effectiveness of metaphor usage for problem solving, we need an index that reflects the success of communication. Problem-solving performance, such as task-completion times, is not sufficient because it is affected by participant ability and the strategies that participants adopt to problem solving. A more microscopic index is also necessary. A recent eye-movement study (Richardson and Rick, 2005) reports that the synchronization of speaker and listener eye movements may be a good indicator of communication success. They showed that the more closely a listener's eye movements were synchronized with the speaker's, the better the listener performed on a comprehension test following an explanation task. In the present study, the effectiveness of using metaphors within problem solving is quantitatively investigated using an analysis of eye-movements.

Experiment

An experiment of collaborative problem solving was conducted. The experimental setting was basically the same as (Spanger et al., 2010) except that we also recorded the gazes of both participants in synchronization with participant utterances and actions during problem solving.

Method

We recruited 16 undergraduates and graduates to make 8 pairs of the same gender friends (3 female pairs and 5 male pairs). They are paid for taking part in the experiment. Each pair was instructed to collaboratively solve Tangram puzzles on a Tangram simulator which displays a goal shape and a working area on a computer display. The goal of the Tangram puzzle is to construct an assigned goal shape by arranging all seven pieces in the working area. The pieces include two large triangles, a medium-sized triangle, two small triangles, a parallelogram and a square. These pieces can be manipulated by simple mouse operations on the Tangram simulator.

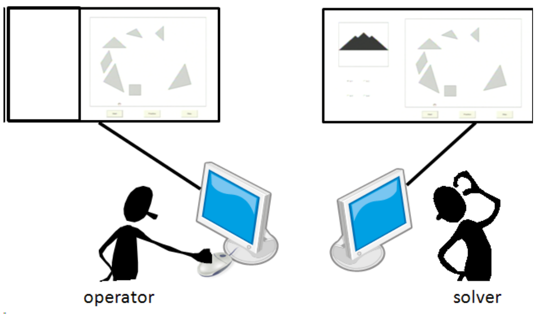


Figure 1. The two participant roles: *operator* and *solver*.

Each member of the pairs was assigned a different role: either *solver* or *operator* (Figure 1). The operator had a mouse to manipulate the Tangram pieces, but did not have the goal shape on the screen. The solver had the goal shape on the screen but did not have a mouse. This setting naturally leads to a situation where given a certain goal shape, the solver thinks about the necessary arrangement of the pieces and gives instructions to the operator about how to move them, while the operator manipulates the pieces with the mouse according to the solver's instructions. They sat side by side with their own computer displays showing the shared working area in real time. A room-divider screen was set between the solver (right side) and operator (left side) to prevent the operator from seeing the goal shape on the solver's screen, and to restrict their interaction to free speech only.

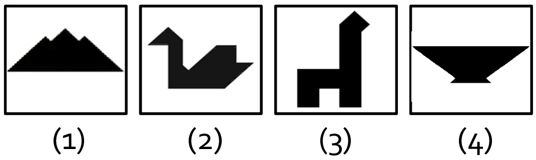


Figure 2. Goal shapes to solve.

Each pair was assigned 4 trials (symmetric (1), (4) and asymmetric (2), (3)) as shown in Figure 2, and they switched roles after two trials. Before starting the first trial as the operator, each participant had a short training exercise to learn how to manipulate the pieces with the mouse. The initial arrangement of the pieces was randomized every time. A time limit of 15 minutes was set for completion of each trial (i.e. construction of the goal shape). In order to prevent the solver from engaging in deep thought and keeping silent, the simulator is designed to provide a hint every five minutes by showing a correct piece position in the goal shape area. After 10 minutes have passed, a second hint is provided, while the previous hint disappears.

A trial ends when the goal shape is complete or the time is up. Utterances by the participants were recorded separately in stereo through headset microphones in synchronization with the positions of the pieces, the mouse operations and the gazes of both participants. Piece positions and mouse actions were automatically recorded by the simulator at intervals of 1/65 second. A participant's gaze was captured by the Tobii T60 eye tracker at intervals of 1/60 second. The 9-point calibration of the eye tracker for both participants was conducted before starting the training exercises. The display size was $1,280 \times 1,024$ pixels and the distance between the display and participant's eye was maintained at about 45cm.

Results

The number of figurative utterances

The aim of the experiment was to investigate the effectiveness of metaphors for collaborative problem-solving by analyzing participant eye-gazes. Thus, we adopted 24 trials from six pairs (four male pairs and two female pairs) where the gazes of both participants were successfully captured for more than 60% of the overall trial period. Utterances including at least one metaphor such as "*Chawan mitai na kanji* (Something like a bowl)" were regarded as figurative utterances. Table I shows examples of figurative utterances from the collected data. The six pairs were divided into two groups based on the frequencies of figurative utterances which were calculated by the ratio of figurative utterances in a trial to the total number of utterances for the trial. The average frequency of figurative utterances over 24 trials was 5.25%. The pairs with an average frequency exceeding this overall average were classified as a high-frequency group and the others were classified as a low-frequency group. Each group consisted of three pairs including one female pair. The average frequency for the high-frequency group was 9.56% and that for the low-frequency group was 1.35%. The average total utterances for the high-frequency group was 347.4 and that for the low-frequency group was 201.7 for each trial.

Table I. Examples of figurative utterances with underlined metaphors.

Goal No.	Role	Figurative utterance
(1)	solver	*ma, Fujisan dane.* (Well, it is Mount Fuji.)
(1)	operator	*Tyôzyô ha sikaku nano?* (Is its top flat?)
(2)	solver	*Ahiru mitaina, Hakucho tte iuno.* (It looks like a duck or a swan.)
(2)	operator	*Sankakukei zya nai no, atama?* (The head, it's a triangle, isn't it?)
(3)	solver	*Sugge kakkowarui ne, kirin mitaina kanji.* (It's like a badly-shaped giraffe.)
(3)	operator	*Kore kubi?* (Is this its neck?)
(4)	solver	*Sono sokomen ni ...* (On the bottom of a cup ...)
(4)	operator	*Râmen no utsuwa tte kotoha ...* (A bowl for noodles means ...)

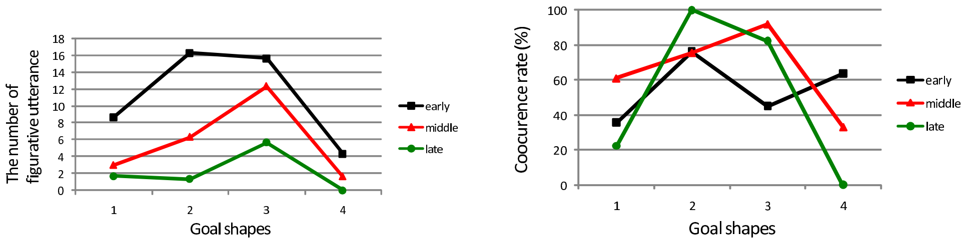
(Richardson and Rick, 2005) reported that overlapping gazes between collaborators indicates good communication between them. We used overlapping gazes as an index of communication quality. In this study, we define an overlapping gaze as when the two participant gazes rest within 100 pixels of each other for more than 0.1 seconds.

Table II. Comparison of the high- and low-frequency groups (average (standard deviation)).

Group	Task completion time	Overlapping gazes rate	Ratio of speaking time (solver)	Ratio of speaking time (operator)
High-frequency	692.6 sec (188.3)	30.6% (7.24)	41.3% (6.40)	36.6% (19.4)
Low-frequency	489.9 sec (183.1)	27.7% (8.05)	40.8% (12.1)	17.1% (11.9)

Table II presents comparisons for several data points for the high- and low-frequency groups for figurative utterances. They include the average times for task completion, the average rates of overlapping gazes which is a trail-based ratio of the total time for overlapping gazes to the total task completion time, the average a ratio of the solver's (or operator's) speaking time, which are a trail-based ratio of the solver's (or operator's) total speaking time to the total task completion time.

A two-way ANOVA was performed to compare the values in Table II for group and the different goal shapes. An angular transformation was applied to the values before the ANOVA in order to resolve the heteroscedasticity of the ratio-scale. There were main effects for group for the task completion time ($F(1, 12) = 20.7$, $p < .05$), for overlapping gazes ($F(1, 12) = 10.9$, $p < .05$) and for the ratio of the operator's speaking time ($F(1, 12) = 15.8$, $p < .05$). There was no effect of



(a) The number of figurative utterances

(b) Co-occurrence rates for figurative utterances and overlapping gazes

Figure 3. Analysis based on the temporal phase of a trial and the goal shape.

the different goal shapes and no interaction for task completion times and the other values.

The correlation coefficient between the number of figurative utterances and the task completion time was $r = 0.46 (p < .05)$.There were no significant correlations between the number of figurative utterances and overlapping gazes ($r < -0.01$), ratio of the solver's speaking time ($r < 0.19$), and ratio of the operator's speaking time ($r < 0.30$). These correlation coefficient results indicate that it is necessary to investigate the data for the high- and low-frequency groups separately.

Temporal analysis of the use of metaphors

We analyzed the occurrence distribution for metaphors over a trial period in the high-frequency group data. A trial period was divided into three phases of equal time spans: early, middle and late phases. Co-occurrences of figurative utterances and overlapping gazes were calculated for each phase. We consider a figurative utterance as co-occurring with an overlapping gaze when the overlapping gaze has any temporal overlap with a period starting from the onset of a figurative utterance to 1 sec after offset of the utterance.

A two-way ANOVA was performed to compare the number of figurative utterances and the co-occurrence rates for figurative utterances and overlapping gazes with the factors of temporal phase during a trial and the different goal shapes. The co-occurrence rates for figurative utterances and overlapping gazes were calculated from the trial-based ratio for the number of figurative utterances co-occurring with overlapping gazes to the number of total figurative utterances. There was a main effect of temporal phase on the number of figurative utterances ($F(2, 12) = 13.7, p < .05$) (Figure 3 (a)). More metaphors were used during early phases of trials. There is a main effect of the different goal shapes for the co-occurrence rates of figurative utterances and overlapping gazes ($F(3, 12) = 9.92, p < .01$) (Figure 3 (b)). Metaphors co-occur significantly more frequently with overlapping gazes when the goal shape is asymmetric than when it is symmetric.

Discussion

In order to investigate the influence of metaphors on collaborative problem solving, analysis was conducted by dividing the participant pairs into a high- and low-frequency group based on the frequencies of their metaphor use. The results indicate that the rate of overlapping gazes and the ratio of the operator's speaking time were significantly lower in the low-frequency group than in the high-frequency group. The low-frequency group, however, completed the tasks significantly faster, suggesting that the solvers could solve the puzzles by communicating less with her/his counterpart in this group. In addition, the low-frequency group used fewer metaphors to avoid misinterpretations. On the other hand, the high-frequency group tried to share knowledge about the goal shape by using metaphors, especially during the early phases of trials. Furthermore, their metaphors functioned more effectively for trials with asymmetric goal shapes. The asymmetric goal shapes required more detailed descriptions in order to correctly transfer a representation from the solver to the operator. They may have attempted to reduce their cognitive loads by using metaphors, in a similar manner to the noisy condition in (Nambu and Harada, 1998).

We compared the co-occurrence rates for figurative utterances and overlapping gazes, and rates for literal utterances and overlapping gazes. Here, literal utterances refers to utterances without any metaphors. On average, more than 50% of the figurative utterances co-occurred with overlapping gazes, irrespective of the trail phase. Interestingly, literal utterances co-occurred more frequently with overlapping gazes, at 73.2% during the early phase, at 80.5% during the middle phase and at 78.6% during the late phase. This result suggests that metaphors do not always facilitate the direct and precise transfer of information, although metaphors can make for smooth communication between participants, and, especially, increase the frequency of operator utterances, as noted in previous research (Kuriyama et al., 2007). Metaphors, however, were useful for conveying a rough sense of the goal shapes in our task setting. Actually, all the first metaphors during the trials, except for one, were produced by the solvers to describe the goal shape. Their usefulness is also supported by the fact that metaphors were more frequently used during trials with asymmetric goal shapes, which are difficult to describe literally.

The present study quantitatively investigated the effectiveness of metaphors within collaborative problem solving by analyzing eye-movements. The results indicate that pairs use fewer metaphors in order to avoid misinterpretations when they adopt a master-slave strategy, where the solver takes the initiative for the problem solving and the operator just performs the instructions given by the solver, which leads to fewer operator utterances and less bilateral communication. On the other hand, pairs use more metaphors that allow for more cooperative strategies to problem solving. This tendency is supported by the fact that more operator utterances were observed in such cases. In summary, metaphors are useful within cooperative problem-solving for facilitating bilateral communication. We do not have sufficient data yet to investigate the kind of gender differences reported in (Hussey and Katz, 2006), but it will be an interesting direction to investigate with larger data.

Acknowledgments

This research is supported by the MEXT program ”Promotion of Environmental Improvement for Independence of Young Researches” and by the MEXT Grant-in-Aid for Scientific Research (B), 21300049.

References

Hussey, K. A. and A. N. Katz (2006): ‘Metaphor production in online conversation: Gender and friendship status’. *Discourse processes*, vol. 42, no. 1, pp. 75–98.

Kuriyama, N., K. Funakoshi, T. Tokunaga, and T. Kusumi (2007): *kyodo mondai kaiketsu niokeru Metaphor no yakuwari (The role of metaphors in collaborative problem solving)*. Tokyo: Hituzi Shobo Publishing Ltd.

Kusumi, T. (1995): *Hiyu no syori katei to imikouzou (Metaphor processing and semantic structure)*. Tokyo: Kazama Syobo.

Lakoff, G. and M. Johnson (1980): *Metaphors we live by*. Chicago: University of Chicago Press.

Nambu, M. and T. Harada (1998): ‘Cognitive artifacts and conversation in referential communication tasks’. *Cognitive Studies*, vol. 5, no. 1, pp. 39–50.

Reynolds, R. and R. Schwartz (1983): ‘Relation of metaphoric processing to comprehension and memory’. *Journal of Educational Psychology*, vol. 75, no. 3, pp. 450–459.

Richardson, D. C. and D. Rick (2005): ‘Looking to understand: The coupling between speakers’ and listeners’ eye movements and its relationship to discourse comprehension’. *Cognitive Science*, vol. 29, pp. 1045–1060.

Sopory, P. and J. P. Dillard (2002): ‘The persuasive effects of metaphor: A meta-analysis’. *Human Communication Research*, vol. 28, no. 3, pp. 382–419.

Spanger, P., M. Yasuhara, R. Iida, T. Tokunaga, A. Terai, and N. Kuriyama (2010): ‘REX-J: Japanese referring expression corpus of situated dialogs’. *Language Resources & Evaluation*. DOI:10.1007/s10579-010-9134-8.