Social Facilitation Effects by Pedagogical Conversational Agent: Lexical Network Analysis in an Online Explanation Task

Yugo Hayashi
Ritsumeikan University
56-1 Kitamachi, Toji-in, Kita-ku
Kyoto, 603-8577, Japan

yhayashi@fc.ritsumei.ac.jp

ABSTRACT
The present study investigates web-based learning activities of undergraduate students who generate explanations about a key concept taught in a large-scale classroom. The present study used an online system with Pedagogical Conversational Agent (PCA), asked to explain about the key concept from different points and provided suggestions and requests about how to make explanations, and gave social facilitation prompts such as providing examples by other members in the classroom. A total of 314 learner's text based explanation activities were collected from three different classrooms and were analyzed using the social network analysis methods. The main results from the lexical analysis show that those using the PCAs with social feedback worked harder to use more various types of explanations than those without such feedback. Future directions on how to design online tutoring systems are discussed.

Keywords
Online tutoring; Explanation activities; Social Facilitation; Lexical Network Analysis.

1. INTRODUCTION
Studies on designing intelligent tutoring systems, such as Pedagogical Conversational Agents (PCAs), which autonomously engage in learning activities, have suggested its effective use for learning, much like a human tutor [12, 9, 1]. Still, few studies empirically investigate the use of such technology for large numbers of students in a class and investigate the learner's cognitive processes. The present study investigated the unique designs of the user interface for learners that use an online tutoring system guided by a PCA in three different types of classes. The study especially focused on the use of PCAs in a concept-explanation activity task, where the PCA asked several questions for explanation and provided feedback such as social information about other members who were engaging in the task. We focused on how such feedback can increase the learner's explanation behaviors during such activities.

1.1 Facilitating explanation activities using PCAs
Studies on collaborative problem solving in the field of cognitive science reveal how concepts are understood or learned [3, 5]. Studies have shown that asking reflective questions for clarification to conversational partners is an effective interactional strategy to gain a deeper understanding of a problem or a concept [15, 16]. It has also been demonstrated that the use of strategic utterances, such as asking for explanation or providing suggestions, can stimulate reflective thinking and meta-cognition involved in understanding a concept. Based on these theories, there have been many attempts in the learning sciences to use such methods in classrooms [17, 13]. However, in an actual pedagogical situation, as in a large classroom, it is often difficult for one teacher to monitor learners and supervise their explanations. Recent studies [2, 11] have shown that the use of conversational agents that act as educational companions or tutors can facilitate learning process. Study [10] have shown that using PCAs that provide suggestions about how to make effective explanations can facilitate better motivation and improve task performance. Moreover, in a series of studies by the author, it is shown that the use of PCAs can provide affective feedback and facilitate better outcomes [7, 8, 6]. More specifically, the results show that PCAs with positive emotion motivates the learners to work harder compared to those without any emotional expressions. In this report, the author further investigated the effects of using such PCAs in an online explanation task. The study focused on a classroom of more than one hundred students who were using an online explanation task, where individuals made explanations to the PCA on a one-on-one basis, as an after school work activity. In such activity, the PCA will play the role of questioner and ask the student to explain about the key concept. The learners were students enrolled in a psychology class where their task was to make explanations about a key concept taught in their class, as an after class exercise.

1.2 Using social facilitating effects
One of the important factors that strongly influence human behavior in groups is the effect of the social influence produced by other members. Studies in social psychology have suggested that work efficiency is improved when someone is watching a person, i.e., the presence of an audience facilitates the performance of a task. The impact that an audience has on a task-performing participant is called the "audience effect." Another relevant concept on task efficiency is called "social facilitation theory" [19]. The theory claims that people tend to do better on a task when they are doing it in the presence of other people in a social situation; it implies that personal factors can make people more aware of social evaluation.

Coming back to the present study, even though the students made explanations about a concept to the PCA in a one-on-one situation, it was extremely important that they were aware that they were working in a social situation. Studies in media-psychology have provided much evidence that people lack social awareness in computer-mediated communication, compared to face-to-face communication [4]. Thus, it is effective to give information about the awareness of other learners online and create social facilitations to make the learners become more active. One of the strong points of using online learning environments is that they are able to collect a huge amount of data from learners. A large database of dialogues of explanation texts may be reused for prompting hints or giving examples to learners who make explanations. It is also effective to provide information about the members who are working on the explanation task in real time or non-real time. If such kinds of feedback are used in online tutoring systems, it may facilitate learners' social awareness, and motivate their explanation activities.

Given all this, the present study investigated the effects of PCAs, which provide information about "other members", along with suggestions and comments about their explanations. The goal of the study is to investigate the how the quality of the learners explanations may change due to the facilitations from a PCA which encourages them to actively explain about key terms that were taught in class. The present study will use social network analysis method to capture the dynamics of diverse explanations during the online task. Unlike standard text analysis methods calculating the frequency of single important key terms that appear in the text, this method enables to detect different key terms that appear simultaneously in one explanation made by the learner. If the learner meets the expectations from the PCA, where it asks the learner to explain the key from various perspectives, different types of key terms should be used during their activity.

2. Method

The study was conducted in three large classes, each consisting of more than hundred students. We constructed an online web system that let learners make text-based explanations about key concepts taught in a psychology class. Students in an undergraduate psychology class used the system, and participated as part of their homework. A total of 30 different key terms (e.g., Gestalt, long-term memory, cognitive dissonance) were selected from the class and randomly assigned to each of the learners based on their IDs. On using the system, they were guided by a PCA that (1) instructed them on what to explain, (2) provided meta-cognitive suggestions, and (3) gave examples about how other members in the classroom made explanations.

2.1 Tutoring system for the experiment

A web-based tutoring system was developed only for the experiment using a web server, a database, and rule-based scripts. It was managed as a member-only system, and learners were required to login to the system for use. As mentioned in the previous section, each student was assigned to work on one randomly selected key term. As they logged into the system, a PCA appeared on the screen and stated the selected key concept, and gave him/her questions about how to explain it. The task was comprised by 17 trials with two major steps in each trial as follows: (a) text-input and, (b) feedback from the PCA.

On the first (Trial 1) and the final trials (Trial 17) the PCA asked the learner to input freely regarding whatever they knew about the key concept. These are taken as pre- and post- tests were they can freely input the messages as a free recall test. Through the 2nd and 16th trials, the learners were given specific questions about what to explain about the keyword. For example, the PCA may ask a series of question such as "How can it be used", "What is it similar to", or "In what period of time you use it" etc. These trials are considered as the explanation/training phase. The PCA also encourages the learner to think on their own way and input individual unique explanations.

On each trial, they were asked to do the following: (1) input explanations and click on the next button, (2) read the provided meta-suggestions from the PCA to make effective explanations, and depending on the experimental condition (explained in the next section), it provided information about other members who also responded for the given key concept.

To facilitate the social presence of the other members and make learners to think in their own way, the study uses two types of prompts. First, the utterances of other learners who had already inputted into the system were used. These messages were presented along with the initials of the person who answered the explanation. This enabled them to be aware how many in the class were working on the same key term. The utterances of other group members were only shown after the learner inputted his/her answers, and so the leaner couldn't simply copy and paste other's explanations during their trial.

2.2 Experiment design and learners

The experiment was conducted in three classes where each class was assigned to an experimental condition. In one class (the baseline condition), all learners were assigned to use PCAs without any social awareness functions or examples of other learners. The PCA only provided back-channel feedback and gave meta-suggestions about how to make explanations more effectively (e.g., Try to think from various viewpoints). These suggestions were compiled from a previous study [7]. In another class (the example condition), the learners were assigned to use the PCAs with additional functions, which provided examples of answers inputted by other members. The third class (the example+ condition) was assigned to those in the example condition with PCAs with additional functions. In other words, they were presented with examples with explanations of others, plus information about the number of members who were assigned to work on that key concept. There were 105 Japanese undergraduates (55 males, 50 females, mean age = 18.26 years) in the baseline condition. In the example condition, there were 105 Japanese undergraduates (55 males, 50 females, mean age = 18.46 years). Finally, in the example+ condition, there were 104 undergraduates (52 males, 52 females, mean age = 18.35 years).

3. RESULTS

3.1 Lexical Network Analysis

The text analysis was comprised by several steps such as (1) morphologically analyzing the text data, (2) developing a dictionary database using a thesaurus, and (3) conducting lexical network analysis to understand the usage of variety of different words during their final explanation. Recently, such social network analysis method is adopted to investigate the usage of important words in collaborative learning [8, 16].

3.1.1 Preprocessing

The recorded texts were broken down into morphemes with the Japanese morphological analysis tool MeCab (Java Sen port: *http://mecab.sourceforge.net* (accessed April 2015)). The objective of the first stage of the analysis was to extract the most frequent morphemes, such as the nouns and verbs through all learners textual inputs. 105,488 morphemes were collected and the most 28 frequent words were chosen as important words for explanations. Those were labeled based on the thesaurus dictionary database such as: 'presence', 'causal', 'relations', 'actions', 'thought', 'matters', 'case', 'conclude', 'understand', 'analogy', 'predict', 'logic', 'reason', 'hypothesis', 'convergence', 'explanation', 'intention', 'theory', 'relative', 'knowledge', 'explicate', 'transform', 'opposition', 'compliment', 'compare', 'inevitability', 'method', and 'reason' [14].

Additionally, based on the semantic hierarchical structure of the thesaurus, new keywords were added to the dictionary database that were related to the 28 keywords. This was done to capture all the semantically related words to these keywords. As a result, 2,722 new words that have relative meanings to the keywords were registered into the semantic dictionary database.

3.1.2 Network Analysis

Using the semantic dictionary database as training data set, the learners textual inputs were further analyzed. For each trial input, the number of appearing semantic keywords in the dictionary were counted. The data of these semantic key words were then analyzed by adopting the social network analysis method. This method was used to analyze the co-occurrence between keywords, i.e. *capturing the diversity of the types of words that were used during one explanation*. The network was developed based on a bipartite graph of keywords x explanations(trials). Since the PCA provided various questions and enforced them to explain uniquely along with their social feedbacks during their explanation activities(trial 2 to 16), their achievements should be reflected to their explanation activities. Learners should use more different types of key terms in the example+ condition since they are facilitated more strongly to take different perspectives by mentioning about other group members presence. Each node in a network was represented as the semantic category of the keyword that was frequently used during their explanation. The threshold of a node(semantic keyword) determining as frequently used or not was defined based on the comparison by the average of other nodes. The threshold of a node n was determined as follows:

$$\theta = \begin{cases} 1(n \geq \overline{n}) \\ 0(n < \overline{n}) \end{cases} \qquad (1)$$

On investigating the differences between conditions and over time, the number of links connecting each nodes were calculated. The following equation represents the amount of density where n stands for the number of nodes and l stands for the number of links:

$$d = \frac{l}{n(n-1)} \qquad (2)$$

Table 1 shows the quantitative results of the lexical network analysis. The results suggest that at the pre-test (1st trial), learners had only few connections between nodes, thus indicating that the variations of words were few in terms of semantic categories. On the post-test (17th trial), the connections of nodes increased due to conditions. This shows that learners used more variety of words during explanations in the post-test(17th trial) example+ condition(0.27) than example(0.24) and baseline(0.15) conditions. The results gives us a clear vision of the dynamics of explanations they gave to the agent differ due to the conditions using more social awareness designs.

Table 1. The score of density of each conditions performed by the lexical network analysis.

Conditions	Pre (1st trial)	Post (17th trial)
baseline	0.07	0.15
example	0.06	0.24
example+	0.06	0.27

The analysis above shows that learners were using more different key terms at the same time in each trial. However it lacks in evidence rather if they tried to use different key terms in their post test compared from those in the pre-test. They might have simply used the same words they inputted from their first trial. It is it is important in this learning context that to know if they changed their phrases or tried to use more sophisticated words from the initial state of the explanation activity. Therefore, additional analysis was conducted to investigating the network similarity between the pre(1st) and post(17th) trial. The following correlation index was adopted on calculating the similarity between the two networks.

$$c = \frac{\sum_{i=1}^{784}(a_i - \overline{a})(b_i - \overline{b})}{\sqrt{\sum_{i=1}^{784}(a_i - \overline{a})^2}\sqrt{\sum_{i=1}^{784}(b_i - \overline{b})^2}} \qquad (3)$$

a and b stands for the number of nodes in the bipartite graph each pre- and post-test respectively. Figure 1 indicates the results of c for each condition.

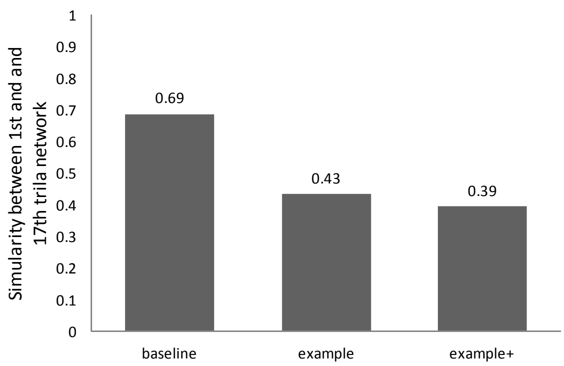

Figure 1. Results of similarity between the pre(1st) and post(17th) trial in each condition

The results indicate that learners in the baseline condition used more similar words from the pre-test on their final post-test explanations(0.69). On the other hand, learners in the example+ and example condition shows that they were using more different key terms compared to those from those in the 1st trial(0.43, 0.39 respectively).

The analysis from the series of analysis indicates that learners with social facilitation (1) used more different key terms simultaneously in their final explanation activities, and (2) those were different from those in the initial explanation activities. This analysis captures a new view from the study of [8] where it did not investigate the changes of the network over time.

4. DISCUSSION AND CONCLUSION

The present study investigated the use of PCAs in an online explanation activity where students were required to make explanations about a key concept. The focus here was to investigate the effects of social facilitations over time, using a large scale database collected during the online explanation task. These social facilitations were provided through a PCA during the learner's explanation activities and they were to enhance the co-presence of other classmates and motivate their activities by encouraging them. In the experiment, students enrolled in three psychology classes used an online explanation system and made explanations to the PCA. They also received comments on how to make effective explanations along with social feedbacks of other classmates. The results of the text analysis show that learners tend to input more important messages simultaneously in the final trial compared to the first trial when they received feedback about other group members (example and example+ condition). This indicates that this type of social feedback can motivate learners to work harder and facilitate effective explanation over time. An interesting point is that even though all the students were told that their answers would not be graded, they still tried harder when they were shown some of the other members' activities. This shows that the effects of the "audience" and "social facilitation" are quite strong in such situations. The results can be interpreted that the situation given to the learner are useful to make the learners aware that their messages could be seen by other in-group members and thus this might have made them work harder in their activities. Another interpretation is that showing others' comments might have allowed learners to avoid negative feelings and thoughts, such as he/she might have inputted something very out of line. As explained earlier in this paper, novice learners have difficulty making explanations to others [5]. Thus, it may be assumed that learners in the baseline condition experienced negative feelings, worrying that they were making mistakes about the text. On the other hand, the use of the examples and the social contexts in the example and example+ conditions may have eased such negative feelings, and thus, increased self-confidence compared to the baseline condition. This study provided implications about how to design effective online tutoring systems, incorporating PCAs with information about other working members, thus providing social facilitation.

5. ACKNOWLEDGMENTS

This work was supported (in part) by 2012 KDDI Foundation Research Grant Program and the Grant-in-Aid for Scientific Research (KAKENHI), The Ministry of Education, Culture, Sports, Science, and Technology, Japan (MEXTGrant), Grant No. 25870910.

6. REFERENCES

[1] Baylor, A.L. and Kim, Y. Simulating instructional roles through pedagogical agents. *International Journal of Artificial Intelligence in Education 15*, 1 (2005), 95-115.

[2] Baylor, A.L. and Ryu, J. The API (Agent Persona Instrument) for Assessing Pedagogical Agent Persona. In *Proc. World Conference on Educational Multimedia, Hypermedia and Telecommunications*, D. Lassner and C. McNaught (eds.) (2003), 448-451.

[3] Chi, M., Leeuw, N., Chiu, M. and Lavancher, C. Eliciting self-explanations improves understanding. *Cognitive Science 18*, 3, (1994), 439-477.

[4] Clark, H.H. and Brennan, S.E. Grounding in communication. In B. L. Resnick, M. R. Levine, and D. S. Teasley (Eds.), *Perspectives on socially shared cognition*, APA Press, 1991, 127-149.

[5] Graesser, A. and McNamara, D. Self-regulated learning in learning environments with pedagogical agents that interact in natural language. *Educational Psychologist 45*, 4 (2010), 234-244.

[6] Hayashi, Y. Learner-support agents for collaborative interaction: A study on affect and communication channels. In *Proc. 10th International Conference on Computer Supported Collaborative Learning* (2013), 232-239.

[7] Hayashi, Y. On pedagogical effects of learner support agents in collaborative interaction. In *Proc. 11th International Conference on Intelligent Tutoring Systems, Lecture Notes in Computer Science* (Cerri, A. S. and Clancey, B. eds.), Springer-Verlag, 7315 (2012), 22-32.

[8] Hayashi, Y. Explanation activities with a pedagogical agent in an online task: Lexical network analysis. In *Proc. CHI 2015 Works-in-Progress*, 1457-1460.

[9] Heidig, S. and Clarebout, G. Do pedagogical agents make a difference to student motivation and learning? *Educational Research Review 6*, 1 (2011), 27-54.

[10] Holmes, J. Designing agents to support learning by explaining. *Computers & Education 48*, 4 (2007), 523-547.

[11] Kim, Y., Baylor, A.L. and Shen, E. Pedagogical agents as learning companions: The impact of agent emotion and gender. *Journal of Computer Assisted Learning 23*, 3 (2007), 220-234.

[12] Kumar, R. and Rose, C. Architecture for building conversational agents that support collaborative learning. *IEEE Transactions on Learning Technologies 4*, 1 (2011), 21-34.

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

[14] National Institute for Japanese Language and Linguistics, *Bunruigoihyou*, Dainippon publications, (2004). (In Japanese)

[15] Okada, T. and Simon, H. Collaborative discovery in a scientific domain. *Cognitive Science 21*, 2 (1997), 109-146, 1997.

[16] Oshima, J., Matsuzawa, Y., Oshima, R., Chan, C. K. K., & van Aalst, J. Social Network Analysis for Knowledge Building: Establishment of Indicators for Collective Knowledge Advancement. In J. van Aalst, K. Thompson, M. J. Jacobson, & P. Reinmann (Eds.), *The Future of Learning: Proceedings of the 10th International Conference of the Learning Sciences (ICLS2012)*, 2 (2012), 465-466.

[17] Salomon, G. *Distributed Cognition: Psychological and Educational Considerations*. Cambridge University Press, New York, USA, 2001.

[18] Shirouzu, H., Miyake, N. and Masukawa, H. Cognitively active externalization for situated reflection. *Cognitive Science 26*, 4 (2002), 469-501.

[19] Zajonc, R.B. Social facilitation. *Science 149*, (1965) 271-274.