Web application for recording learners' mouse trajectories and retrieving their study logs for data analysis

Mitsumasa Zushi*

Faculty of Informatics
Shizuoka University, Japan
E-mail: mzushi@inf.shizuoka.ac.jp

Yoshinori Miyazaki

Faculty of Informatics
Shizuoka University, Japan
E-mail: yoshi@inf.shizuoka.ac.jp

Ken Norizuki

School of Information Studies
Shizuoka Sangyo University, Japan
E-mail: norizuki@ssu.ac.jp

*Corresponding author

Abstract: With the accelerated implementation of e-learning systems in educational institutions, it has become possible to record learners' study logs in recent years. It must be admitted that little research has been conducted upon the analysis of the study logs that are obtained. In addition, there is no software that traces the mouse movements of learners during their learning processes, which the authors believe would enable teachers to better understand their students' behaviors. The objective of this study is to develop a Web application that records students' study logs, including their mouse trajectories, and to devise an IR tool that can summarize such diversified data. The results of an experiment are also scrutinized to provide an analysis of the relationship between learners' activities and their study logs.

Keywords: Recording; Mouse trajectories; Analysis data

Biographical notes: Mitsumasa Zushi is currently an Associate Professor of Faculty of Informatics, Shizuoka University, Japan. He earned his M.A. at Graduate School of Hiroshima University in Japan and then completed a doctoral course at the same graduate school without taking the degree. He specializes in cross-cultural communication and English education.

Yoshinori Miyazaki is currently an Associate Professor of the Faculty of Informatics in Shizuoka University, Japan. He earned his Ph.D. in the program of Engineering from the University of Tsukuba in Japan. Dr. Miyazaki specializes in numerical analysis and e-learning. His interests include development of educational software especially involving English and math.

Ken Norizuki is currently a Professor of the School of Information Studies in Shizuoka Sangyo University, Japan. He earned his M.Sc. at the University of Edinburgh, U.K. He specializes in language testing and learning.

1. Introduction

With the advancement of IT, the introduction of e-learning systems has accelerated in educational settings in recent years. This has enabled teachers to record detailed study logs of their students, which was hardly feasible with conventional paper-based exercises. Even as a large number of software products that collect such data have been merchandized as LMS (Learning Management System), a method of analysis for these study logs has not been researched in much depth. In this study, the authors prepared English jumbled-word (henceforth, EJW) questions, for which learners were asked to record the jumbled words to match appropriate Japanese translations. Based on this exercise, we sought to develop software that could retrieve the learners' mouse trajectories as well as their study logs for analysis. Such software would allow on-site teachers to identify online the problems each learner encounters in solving EJW questions, which would effectively help to determine whether the learner could benefit from specific additional instructional support. One good potential is that the system gives a hint for teachers to distinguish correct answers solved with confidence from accidental ones, by analyzing its mouse trajectory (or, to distinguish the extent of lack of understanding by the trajectory of each incorrect answer). In Japan, EJW questions have traditionally been used in sentence transformation exercises and remain in common use, typically presented in objectively-scorable high-stakes proficiency tests, such as in the National Center Test for University Admissions and Test in Practical English Proficiency (Eiken). However, as Shea (2009) points out, this type of question requires not only an understanding of structural accuracy but also test-taking tactics. Answering such a question successfully does not necessarily mean that the candidates have correctly and fully understood *per se* the sentence structure, syntax, grammar, or idioms.

2. Related research

Ohmori, Iibuchi, Horie, and Itoh (2006) analyzed the behavior of mouse movements during the course of learners' reading of materials. Their analysis led to the classification of learners' characteristics into 3 general patterns. This system, however, did not extract the specific characteristics of individual learners. Freeman and Ambady (2010) presented a software package named MouseTracker that allows the use of a computer mouse-tracking method for assessing real-time processing in psychological tasks. Mouse trajectories can be processed, averaged, visualized, and explored for analyzing users' behaviors. Arroyo, Selker, and Wei (2006) presented a web logging system that tracks mouse movements on websites to help web site administrators run usability tests and analyze the collected data. Tateda, Kotani, and Horii (2005) developed an authentication system which focused upon the analysis of mouse trajectories. This system was used to illustrate a number of technical properties of individual learners' mouse movements, including their performance speeds. Despite some similarities to our present study, their approach was primarily intended for a biometrics authentication system. Fujita, Zushi, Hayashi, and Yamasaki (2006) implemented an LMS with a function that detected students' learning plateaus (stages of stalled progress) by collecting study logs with psychoanalytic information. Several similar studies have featured the functions of study logs; however, to date, none of them seem to have incorporated mouse trajectory data for the purpose of analyzing learning processes during e-learning.

3. Development of software

3.1. Basic concepts

EJW questions that are presented in a conventional paper-based format present the following problems:

- The process of finding solutions is not traceable,
- The absence of a retrieval system causes difficulty in identifying real problems,
- Monitoring the entire process by which the learner reached a solution is tedious and complex.

In order to solve these problems, our aims have been:

- To reproduce learners' solution processes with no loss of information,
- To enrich the information retrieval functions,
- To reduce the teacher's time-consuming workload by quickly reproducing mouse trajectory logs,
- To perform statistical analysis of a large volume of data, and
- To offer learners features that are easy to use.

The software we developed consists of a Study Module, a Retrieval & Analysis Module, and a Problem Constructing Module. These modules have been programmed using PHP, Visual Basic, and MySQL.

3.2. Study module

The aim of the Study Module (Fig. 1) is to facilitate the EJW-based problem-solving process by allowing the entire process to be performed as a D&D (Drag&Drop) solution. Also, the location for the insertion of word(s) is explicitly displayed with the prompt "|". The words to be moved are underlined. Along with these basics, additional functions were further implemented, following an investigation of how this type of problem is solved, based upon interviews with learners. These functions are:

- Word groupings ... An arbitrary number of words may be grouped together by mouse-dragging (rectangular selection) if it is deemed preferable to treat them together.
- Relocation to register(s) ... Areas called "registers" are provided as a temporary "shelter" for words. Learners can integrate a set of words into a few meaningful segments, which helps them reorganize their ideas. Their answers are evaluated for correctness only after the words are put back into the main line.

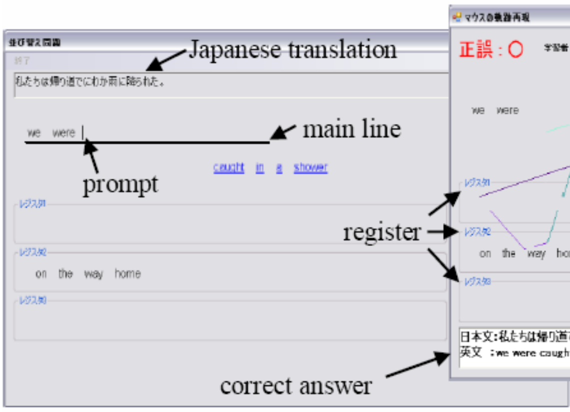

Fig. 1. Screenshot of study module

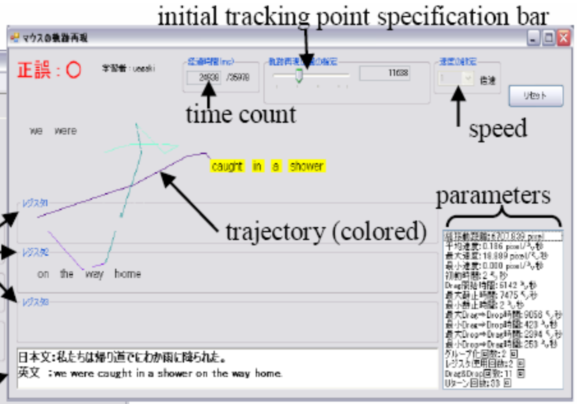

Fig. 2. Reproducing mouse trajectory

The student responses are then automatically graded. In the older version, the scores were binary (correct/incorrect), while in the latest version, partial credit can also be given, which reflects the learner's degree of grammatical accuracy, understanding of the sentence structure, or idioms required to generate a correct answer (e.g., a certain amount of credit is given if an answer contains a pre-defined series of words in the right order).

3.3. Retrieval & analysis module

This module performs the retrieval and analysis of the study log data obtained in the Study Module. This module has the following 6 menus, each of which has links to related pages of the other menus so that teachers can browse among relevant sets of data: the menus of Reproduction, Learner Analysis, Problem Analysis, Study Log Retrieval, Correlation Analysis, and Clustering; for more details see Fig. 11 in Appendix.

Reproduction

This menu, which is shown in the above Fig. 2 as supplemental description to the Study Module, stores and reproduces all the actions recorded in learners' mouse trajectories, such as D&Ds, word groupings, relocations to registers, etc. Furthermore, in order to lessen the burden on teachers, the following features were implemented:

- Change of speed of reproduction … to 50, 200, 300, and 500% of its original speed,
- Reproduction at any selected point … for concentration on only characteristic parts,
- Visualization of mouse trajectories in coloured lines: the colours change each time an independent set of D&D actions is detected, and then fade away in several seconds.

Learner analysis

This menu shows statistical data for specific learner characteristics, including each parameter and mouse trajectory (Fig. 3). The parameters are: # (the Number) of Problems Solved, % (the Ratio) of Correct Answers, Total Elapsed Logon Time, # of Access, and Average Speed Needed to Answer. Mouse trajectories are reproduced after a given study log is specified.

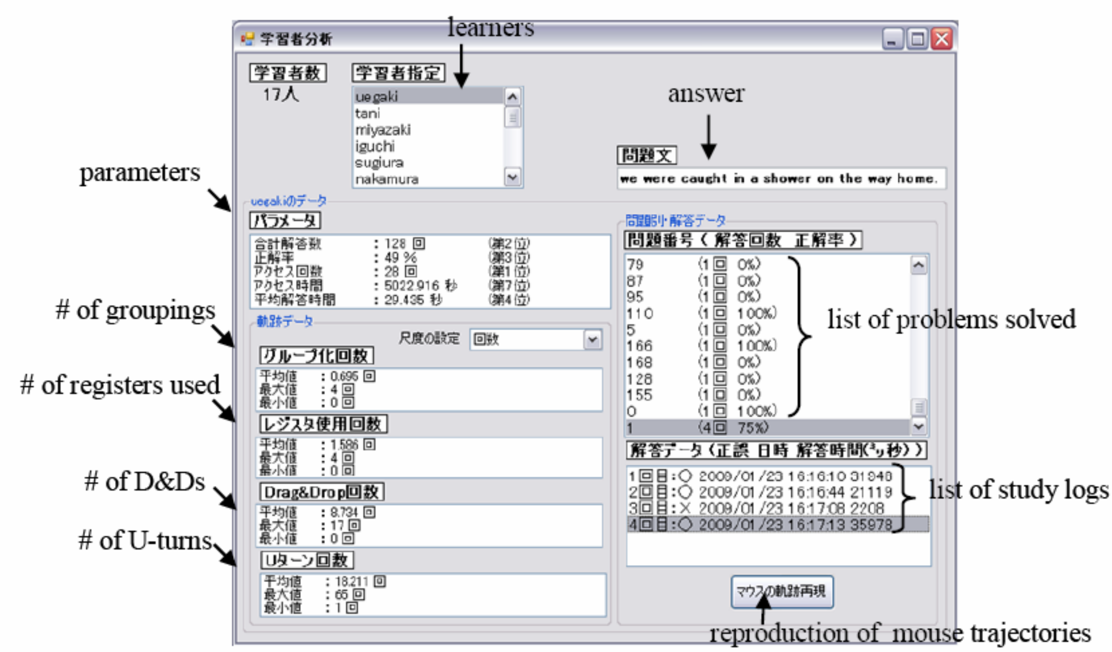

Fig. 3. Learner analysis menu

Problem analysis

This menu shows how individual problems were solved, including various statistical data on problem-response parameters as well as visual representations of mouse trajectories (Fig. 4). The parameters here include: # (the Number) of Student Attempts per Problem, % (the Ratio) of Correct Answers, Average Speed Needed to Answer, and # of Jumbled Words in the Problem. As is in the case of Learner Analysis, the selection of a study log leads to the reproduction process in order to see how the problem was solved by the learner.

Study log retrieval

This is the menu in which data meeting pre-specified criteria are retrievable through the following three types of data search:

- "Learners" Search : # of Answers, % of Correct Answers, # of Access, Total Elapsed Logon Time, Average Speed Needed to Answer
- "Problems" Search : # of Answers, % of Correct Answers, Average Speed Needed to Answer, # of Jumbled Words in the Problem

- "Individual Response Data" Search : Username, Binary Evaluation/Degree of Correctness, Time Used to Answer, Access Period by Learners

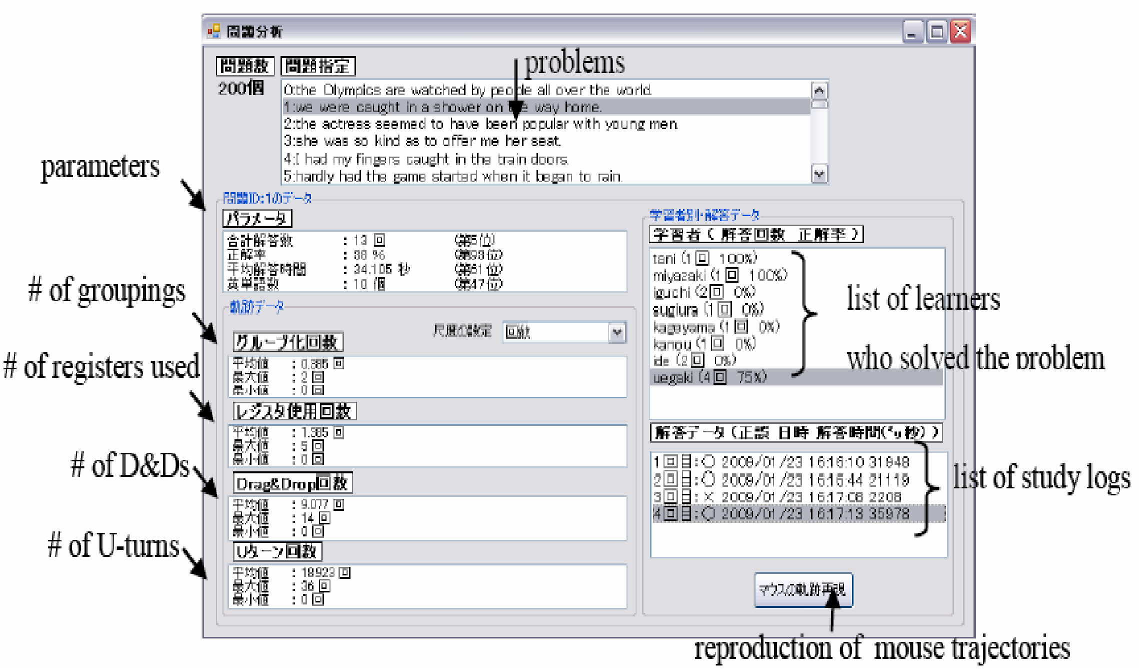

Fig. 4. Problem analysis menu

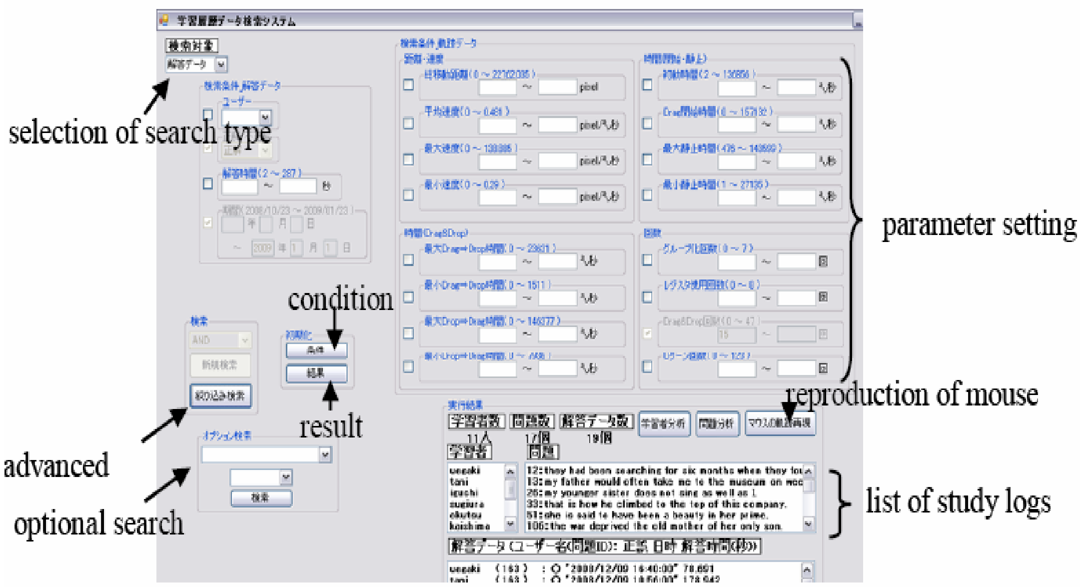

Fig. 5. Study log retrieval menu

Other statistics available are the distance of mouse movement, the average velocity of mouse movement, the duration of mouse suspension, # of D&Ds, # of U-turns (instances of student hesitation), # of uses of registers, and # of groupings. AND and OR search commands are also implemented (Fig. 5). Combining several search types and commands, the teacher can easily have an access to, for example, who needed U-turns more than 15 times to solve a particular problem, how many students have used this software in the last two months, which problems required D&Ds less than 10 times for a certain student to solve, and so on.

Correlation analysis

The computation of various study log statistics makes it possible for the correlation coefficients to be produced for all combinations of parameters of the learners or problems (Fig. 6): this figure is an example of correlation between # of D&Ds and the total distance of mouse movement, with each dot corresponding to one time solution of a problem.

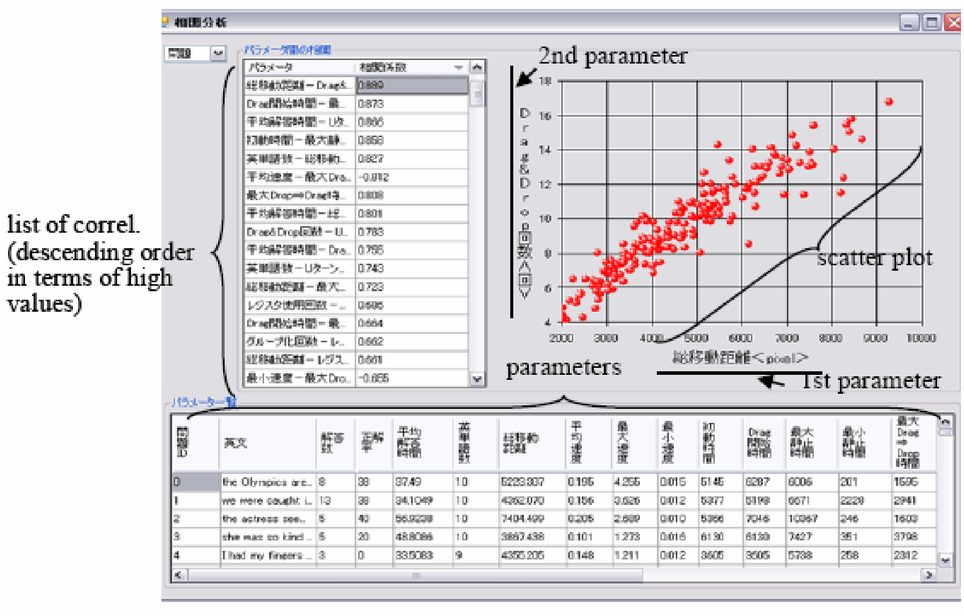

Fig. 6. Correlation analysis menu

Clustering

This function assembles similar learners in clusters. In this study, considering the large difference in the variances and scales of each parameter, a standardized Euclidean distance was used instead of the generally used (non-standardized) Euclidean distance. Ward's method was adopted for computing the distances among clusters. Teachers can choose the number of clusters and which parameters to be incorporated for clustering criteria, from "# of Answers," "% of Correct Answers," "# of Access," "Total Elapsed Logon Time," and "Average Answer Time."

3.4. Problem construction module

This module supports teachers as they construct EJW problems on their own. This task is greatly simplified since they only need to write an English sentence with the words in the correct order (and the corresponding Japanese sentence), and the module then automatically breaks down the sentence into words and randomizes their order.

Automatic randomization of this module in our first software, however, had slight, but not trifling possibility of making an EJW problem whose word order was almost the same as the correct answer's, especially when the problem had as few as seven or eight words in a sentence. When this happened, the ratio of correctness dramatically rose irrespective of the understanding level of the learner, with fewer D&Ds demanded to solve the problem. This kind of unintended proximity to the target structure in constructing problems should be avoided in light of high negative correlation between the ratio of correctness and the number of D&Ds, which will be discussed in 4 (Experiment and Discussion).

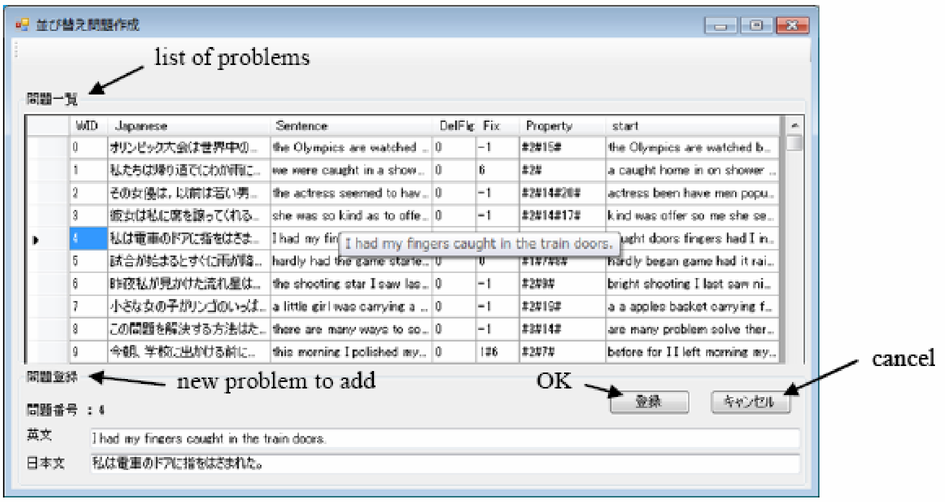

Fig. 7. Problem construction module

Fig. 8. Word position fixing

Our latest version of this module, however, is revised so as for the teacher to choose among three options of word arrangements of EJW problems: randomized, alphabetical, or optional order. Thus, if the teacher selects the alphabetical form, the problem whose answer is to be "I had my fingers caught in the train doors" is provided in the arrangement of "caught, doors, fingers, had, I, in, my, the, train," leaving no chance of getting near the correct sentence (Fig. 7, 8). Although there remains a very small chance that the original or randomly-rearranged word order coincides with the alphabetical order, this new module allows instructors to eliminate their anxiety through the use of optional word order, i.e. the rational selection of word arrangement.

Another difficulty in making EJW problems is how to avoid double/multiple correct answers to the same problem. Binary marking system of computer accepts only one generated sentence as a correct answer, so that the other correct one(s) will be esteemed to be wrong. For example, an original sentence of "We had a thunderstorm in Tokyo yesterday" from which a problem is constructed, has its alternative answer, "Yesterday we had a thunderstorm in Tokyo," and both answers should be evaluated to be correct. This often happens when the sentence involves an adverb/adverbial phrase in it. In the latest version, this module has a function to have the location of one or more words in a sentence fixed in order to prevent double/multiple answers (e.g., "yesterday" in the above sentence can be placed in the initial or end position in accordance with the teacher's choice). In the sentence shown in Fig. 8, the word "hardly" is placed at the beginning of the sentence by checking the "fixing label" in order to exclude "The game had hardly started when it began to rain" as an answer.

3.4. Data stored for analysis

In this software, several database tables are required in order to efficiently summarize different types of study log data. They include: "Member Table" to manage learner information, "Problem Table" to manage problems, "Access Table" to record access information, "AnswerData Table" and "MouseTrajectory Table". "AnswerData Table," which is the key data table for answer logs, and which also stores each piece of answer data, except for mouse trajectories, has the following fields: Learner ID, Problem ID, Accessed time, % of Correct Answers/Degree of correctness, Time spent solving the problem, Initial order of English words (placed at random, in alphabetical, or in optional order), and MouseTrajectory ID (foreign key to MouseTrajectory Table). "MouseTrajectory Table" manages mouse trajectories, which is comprised of various related information obtained by event-driven programming: MouseTrajectory ID, Time spent after beginning task, (X,Y)-coordinates, Event info (MouseMove, MouseUp, or MouseDown), Position of mouse (on mainline, in registers 1 to 3, or otherwise), English word being activated by mouse, and a set of English words being selected into groups.

4. Experiment and discussion

An experiment was conducted to evaluate the software as it was used by learners (12 subjects) and by teachers (2 English teachers), who had been informed that their mouse trajectories would be recorded to be analyzed. The requirements set were that the subjects needed to solve at least 50 problems (provided at random) out of prepared 200 problems, which were labelled "elementary," "intermediate," and "advanced" levels (The 50 sampled problems are stored at http://lmo.cs.inf.shizuoka.ac.jp/~yoshi/sampled-50-problems.xlsx). In total, 628 sets of log data were collected. We conducted our preliminary evaluation and discussion in terms of the following:

General observations

First, it is expected that the "Retrieval & Analysis Module" will allow teachers to explore the process by which learners solve EJW problems via visual representations of their mouse movements. Several noticeable patterns of mouse movements were identified:

- Starting to work on the leftmost (first) word,
- Grouping words (e.g. for each idiom) with the use of registers,
- Long pauses followed by active mouse movements,
- Fluctuating trajectories during moments of hesitation.

Analysis of collected study logs

In the experiment, the average scores for individual learners presented pretty high negative correlations with the average speeds to solve problems, which implies that the higher the ratio of correct answers, the shorter the time for solving the problem (Table 1). It is conjectured that learners do not hesitate if the answer is obvious or somewhat easy. This is supported by the fact that 66.0% of the 50 fastest-answering students' responses were found to be correct. In contrast, using the same criterion, only 8.0% of the 50 slowest-answering students' responses were correct. This parameter might be a good predictor of learners' level of understanding. Similarly, these average scores had negative correlations with the average number of U-turns and D&Ds. It is easily assumed that mouse movements become unstable (e.g., excessive backward and forward movements) when learners are unable to confidently find a solution. This tendency may be reflected in the number of U-turns and D&Ds, which is our hypothesis. After retrieving the top 54 data in terms of the number of U-turns, we recognized that the correct ones totalled only 20.4%, while the bottom 70 data (small number of U-turns) show 52.9% were correct. As for the number of D&Ds, only 16.7% among the top 54 data reached the correct answer, while 52.6 % of the bottom 78 data found proper solution. Also, by reproducing the mouse movements, it seemed apparent when learners were tackling problems on a trial and error basis, by their up-and-down and side-to-side mouse movements, and by their frequent execution of the D&D functions. As is seen in Table 1, time used to solve the problem, the number of U-turns and D&Ds, they all have highly negative correlation with the ratio of correct answers. We computed the correlation for each group (formed by the number of words in the problem sentences); and the results were respectively -0.480 (7-8 words), -0.579 (9-11 words), and -0.679 (12-14 words). The more words the problems have, the higher the negative correlations are: this indicates the possibility that teachers can make use of mouse trajectory data as better criteria for the lack of understanding when they prepare EJW problems with words of about 10-14: however, far more than 14 words in a sentence will make the problem unanswerable for learners.

Table 1
Correlation with "% of Correct Answers" (by Binary Marking)
-- "Learners" Perspective

Parameter	Total	Upper	Lower
Time to Answer	-0.796	-0.873	-0.646
Distance (from start to end)	-0.718	-0.671	-0.789
# of D&Ds	-0.759	-0.606	-0.919
# of U-turns	-0.658	-0.577	-0.556
# of Groupings	0.429	0.683	0.080

Note: "Upper" means the top 6 subjects in terms of ratio of correct answers.

Compare this with Table 2, where the correlation coefficients were gained by computing the results of partial point marking. Here again, three major parameters (Time, U-turns and D&Ds) show quite definite, negative correlation with degree of correct answer.

Table 2
Correlation with “Degree of Correctness” (by Partial Point Marking)
-- “Learners” Perspective

Parameter	Total	Upper	Lower
Time to Answer	-0.771	-0.887	-0.575
Distance (from start to end)	-0.747	-0.732	-0.803
# of D&Ds	-0.776	-0.665	-0.888
# of U-turns	-0.697	-0.634	-0.625
# of Groupings	0.455	0.688	0.147

Note: “Upper” means the top 6 subjects in terms of degree of correct answers.

Closer examination reveals that the correlation coefficients in Table 2 are higher in general than those in Table 1 except two coefficients in the lower subjects, which seems to support that partial point marking may be better indicative of the degree of understanding of students. Partial point marking system employed here is to reflect different levels of achievement when answers contain pre-specified phrasal segments in the correct order. It should be admitted that this marking system is still inadequate, since it is not capable of dealing with the trivial mistakes which can be tolerated in writing when the sentence is almost fully understandable. Many of our subjects answered in solving a certain problem, “Olympics are watched by the people all over the world;” to which no points were given in binary marking because the correct answer should be “The Olympics are watched by people all over the world.” The former answer is not perfect but it may be worthy enough to award some partial points. With the development of reliable and valid marking criteria for awarding partial credit and proper mathematical analysis given to D&Ds and U-turns, it might be possible for these parameters to be far better indicators of students’ understanding.

Graphs for relationship between average scores and other parameters

The key parameter we focused upon was the average score of individual learners, which was an indicator of their degree of understandings. After scrutinizing various other parameters, it was found that “number of U-turns” and “initial movement time,” each divided by the number of words in a problem for normalization, displayed a remarkable relationship.

Fig. 9 shows that the average scores decrease gradually as the number of U-turns / number of words (“*x*” in the graph) increases, and that regardless of the level of problem (Advanced/Intermediate/Elementary) the average scores were the highest between $0.5 \le x \le 1.5$. This would indicate that the number of U-turns (/ the number of words) serves as an index of hesitation and its resulting “low rates of correct answers.” Fig. 10 presents a graph of: x = the “initial movement time” / number of words; y = “the average scores” (the term, “initial movement,” indicates the time passed until the initial movement is made in one mouse trajectory). “A” through “N” represent each subject of this experiment. One can conjecture that there will be no mouse movement until learners come up with acceptable or partially acceptable solutions, and thus that the longer the time, the more difficult the learner finds the problem. This graph shows such a tendency regardless of the level of the learner.

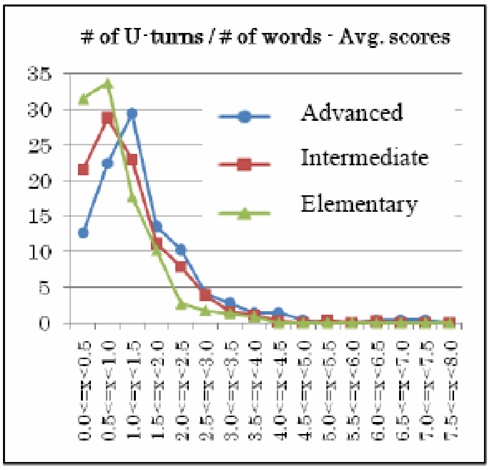

Fig. 9. U-turns - scores

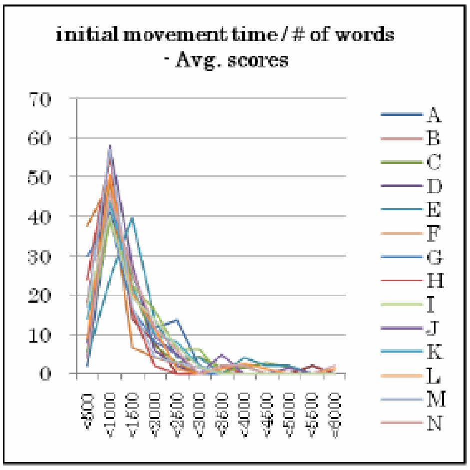

Fig. 10. Initial time - scores

When this is the case, on-site teachers who use our proposed system could review such problems with mouse trajectories indicating hesitation and uncertainty in greater detail with their students. Furthermore, this system could help teachers identify accidental correct answers reached by students who have not fully understood the target structures or idioms. In our experiment, some subjects were, for instance, found to be uncertain about the idiomatic phrase of “on the way home” in the sentence “We were caught in a shower on the way home” even though their answers were correct, because their trajectories showed many U-turns and D&Ds when forming the phrase. As far as this idiom is concerned, these students’ level of understanding would be estimated as almost the same as that of the group who, after replacing the word order many times, came to a wrong answer “. . . on the home way” in the end. With more elaborate scrutiny and mathematical analysis given to more data from more subjects, mouse trajectories may well be a good index which will help teachers identify the problems students have and to effectively determine what additional review or support to be offered.

5. Concluding remarks

In this study, we developed software to obtain study log information, including mouse trajectories. The usefulness of mouse trajectory information was verified as it displayed various facets of how problems were solved by individual learners. Tasks to be addressed in future research should include:

- improvement of partial credit system that reflects the quality of the solution as revealed by mouse trajectory data;
- adaptive selection of next problems to be presented in accordance with data from the learner’s study log;

- categorization of problems that need to be retrieved;
- more precise detection of hesitation from mouse trajectory data;
- many more samples to be investigated.

This module can be applied to other CALL modules since mouse trajectories are involved in all types of problems. The final goal of this study is to implement this module for multi-purpose use in e-learning, such as multiple-choice problems and logic-chart problems (problems that require the creation of schematic diagrams representing the objectives of a specific program, such as a flow-chart).

Acknowledgments

This work was supported by Grant-in-Aid (KAKENHI) for Challenging Exploratory Research (23652136), and we also owe special thanks to Ryosuke Miki, a student from Faculty of Informatics, Shizuoka University in Japan, for his help with data processing indispensable in our study.

References

Arroyo, E., Selker, T., & Wei, W. (2006). Usability tool for analysis of web designs using mouse tracks. *Proceedings of CHI EA'06*, ACM, 484–489.

Freeman, J. B., & Ambady, N. (2010). MouseTracker: Software for studying real-time mental processing using a computer mouse-tracking method. *Behavior Research Methods, 42* (1), 226–241.

Fujita, N., Zushi, H., Hayashi, T., & Yamasaki, T. (2006). Learning management system including functions for detecting the learner's stalled situation in e-learning. *IEICE Technical Report, ET 105*(632), 63–68.

Ohmori, Y., Iibuchi, D., Horie, I., & Itoh, K. (2006). Understanding learning effect using MBC (1)-analysis of mouse behaviour. *Conference on IT Education*, 51–54.

Shea, D. P. (2009). The university English entrance exam and its implied effect on EFL pedagogy in Japan. *The Hiyoshi Review of English, 55*, 97–132.

Tateda, C., Kotani, K., Horii, K. (2005). Evaluation of an individual authentication system using the trajectory by the mouse. *IEICE Technical Report, HCS 105*(74), 17–20.

Appendix

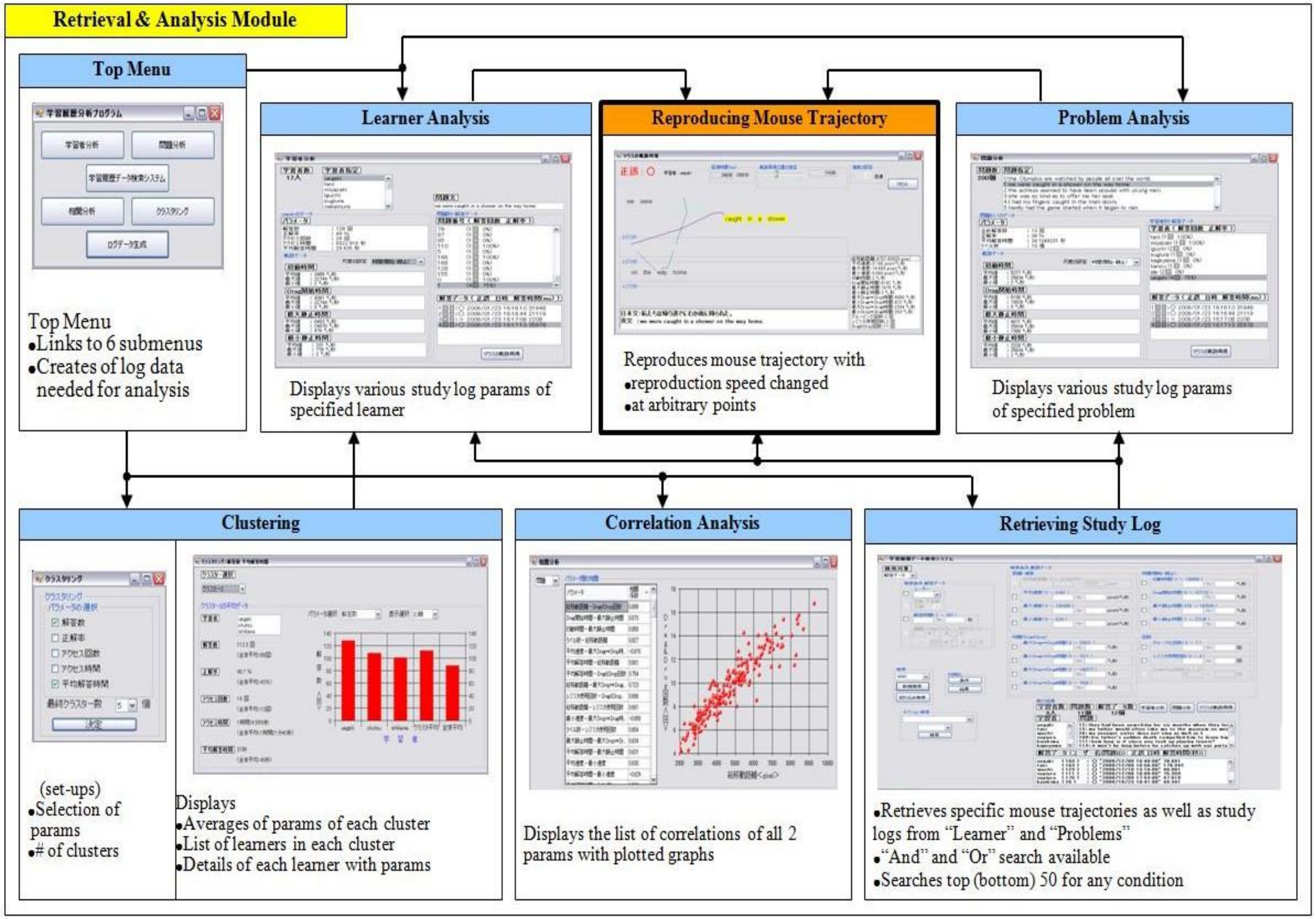

Figure 11. Structure of Retrieval & Analysis Module