The International Journal of E-Learning and Educational Technologies in the Digital Media (IJEETDM) 4(4): 131-137
The Society of Digital Information and Wireless Communications (SDIWC), 2018 ISSN: 2410-0439 (Online)

University Students Using SWISS to Display Images Searched by Keywords Extracted from Utterances

Yutaka Yamaguchi, Daisuke Shibata, Chika Oshima, and Koichi Nakayama
Graduate School of Science and Engineering, Saga University
Honjo 1, Saga City, Saga Pref., Japan
knakayama@is.saga-u.ac.jp

ABSTRACT

In this paper, a system, Search Websites' Images using Search Suggestions (SWISS), was improved based on the results of an experiment in which university students discussed their research theme using SWISS. SWISS displayed images based on the extracted words from the students' utterances and another word predicted from a letter of the alphabet or Hiragana via an autocomplete function. The experiment showed that some entertaining ideas were generated from the students who used SWISS in their brainstorming. On the other hand, SWISS did not operate when no one was speaking, even when the students wanted to see some images to expand their ideas. Therefore, SWISS was improved so that the associated images are displayed even without word inputs to SWISS.

KEYWORDS

divergent thinking, expand idea, brainstorming, autocomplete, Watson NUL.

1 INTRODUCTION

University students who engage in the human-computer interaction arena often discuss their research themes in regular seminars hosted by the associated laboratory. They develop ideas to resolve an existing problem. For example, what could they construct as a support system for a couple in a long-distance relationship to keep getting along well [1]? First, it is important to expand their ideas, which is called divergent thinking.

There are many methods and systems that encourage people to expand their ideas [2][3][4][5][6][7]. Montag said that typical information will not help to achieve the goal of novelty [8]. Diverse opinions lead to an increased breadth of idea production [9].

Wang [10] stated, "Using pictures as extra stimuli may be more effective than language in stimulating idea generation." IdeaExpander [11] is a tool that supports group brainstorming by displaying pictures based on a chat conversation. IdeaExpander monitors the chat conversation and determines whether a remark contains an idea by using a machine learning classifier [10]. Then, the keywords drawn from the remarks are used to search pictures on Flickr. Especially, "pictures that contain multiple idea categories or categories that are less commonly discussed" help to convert the diversity of generated ideas [10].

Therefore, when people are brainstorming, an image that illustrates not only the meaning of word "X", which they have already put forward, but also the meaning of another word, "Y", with which they have not yet associated the word "X", may encourage their idea generation.

Shibata [12] presented a system, Search Websites' Images using Search Suggestions (SWISS), which continuously extracts words from the utterances of participants in a discussion. Then, the system constantly displays images based on the extracted words and other words predicted from a letter of the alphabet or Hiragana via an autocomplete function.

Shibata [12] examined whether images searched by SWISS allowed participants in a brainstorming activity to generate new ideas. In the experiment, a participant could come up with a new idea from an unexpected image generated by SWISS.

On the other hand, SWISS does not operate when no one is speaking. We believe it is necessary for SWISS to display the images for the participants during periods of silence.

In this paper, SWISS is improved to display the images even when no one is speaking.

2 SWISS

At the start of the SWISS application, voice data are input to SWISS. Then, SWISS recognizes audible words and converts them to text. SWISS conducts internet searches based on the text data and displays six images of the search result on the screen, as shown in Figure 1. It takes about one minute from the voice input for the images to disappear from the display.

It is assumed that a smartphone with SWISS installed is placed on a table at which participants of a discussion are sitting. If there is a wide display available, it is better to connect it with the smartphone to display the smart phone's images on the larger screen.

It is not always necessary for the discussion participants to see the displayed images. Then observation of the images is dependent on the participants' intent.

Figure 1. Example of result of searching images

3 SYSTEM SET UP

Figure 2 shows the construction of the SWISS system. The participants' utterances are input into the Android Speech Recognizer and converted to text. The text data are input to Bing Web Search API. Bing Web Search API searches websites using the text data. Then, the URL of the website at the top of the search results is extracted. IBM Watson's Natural Language Understanding (Watson NUL) analyzes the website and extracts four words that represent its features. The four words are input to Bing Autosuggest API. Each word is listed with a letter of the alphabet or a Hiragana character, which is selected at random, and receives another word that begins with that letter or Hiragana character from Bing Autosuggest API. Figure 3 shows an example of the autosuggest function. Finally, Bing Image Search API searches images with one of the four words and the word suggested by Bing Autosuggest API. Table 1 shows the development environment of the application.

4 EXPERIMENT

4.1 Aim

This section examines whether images searched by SWISS allow participants in a brainstorming activity to express a greater variety of ideas than do images searched by keywords extracted from only the participants' utterances [12].

4.2 Method

Twelve university students participated in the experiment. One of them was female. They were divided into four groups of three participants each.

Figure 4 shows the positions of the participants. One of the three participants assumed the role of a facilitator.

Table 2 shows two conditions for the experiment. In Condition A, SWISS searched images based on each word of the four words that Watson NUL extracted from the websites and the word suggested by Bing Autosuggest API. In Condition B, SWISS searched images based on each word of the four words. In other words, SWISS did not use the Bing Autosuggest API.

Each group discussed generating a product idea twice in both conditions. Two themes (Q and R) were provided for the discussion.

Theme Q: People in a long-distance relationship tend to become emotionally unstable and/or suspect infidelity because they miss their partner. Therefore, please suggest a tool or application that will allow couples to convey a sense of presence and/or express their affections to each other for reassurance.

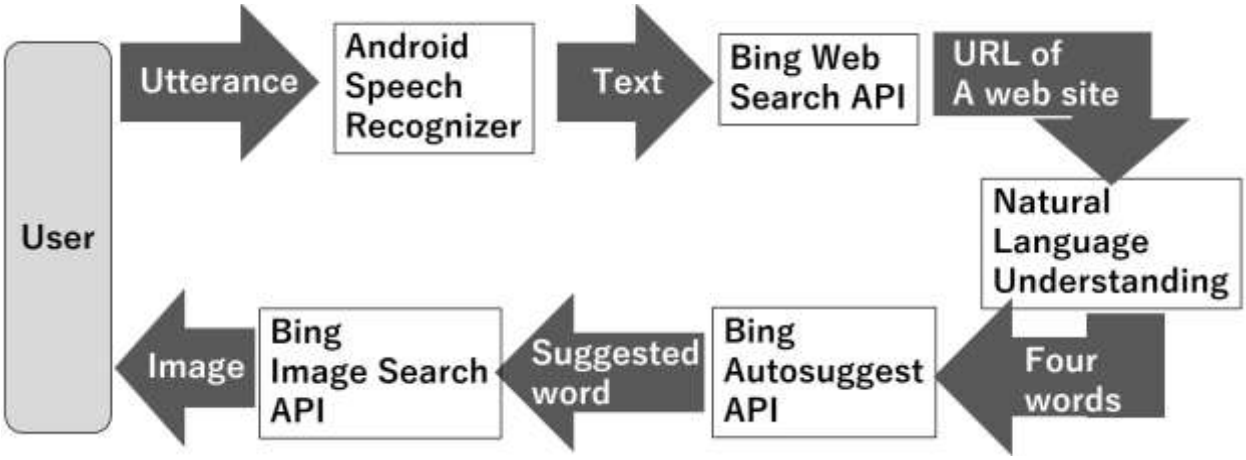

Figure 2. The system searches images with two words, which are extracted from a brainstorming and selected from candidate words with the autosuggest function.

Figure 3. Example of result by Bing Autosuggest API

Table 1. Development Environment

Classification	Specific
OS	Android7.1.1
Terminal	Nexus 5X
Integrated development environment	Android Studio
Development language	Java

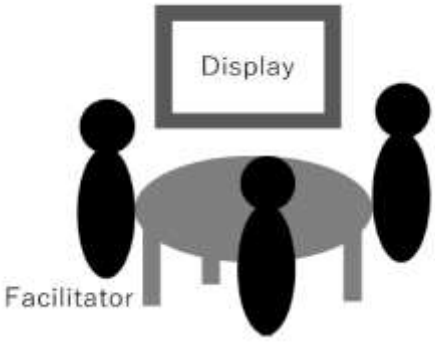

Figure 4. Position of the participants

Theme R: Recently, a lot of floods and sediment disasters have been caused by heavy rain. However, elderly people who do not have a smartphone or think that they will not be affected by these disasters may not evacuate immediately. Therefore, please suggest a tool that will facilitate elderly people's immediate evacuation.

The participants were asked to discuss the given theme with the group members for 15 minutes each. The participants did not know the differences between Condition A and B.
After the discussion of both of themes, we asked the participants about the differences between the two conditions.

4.3 Result

The participants of Group 1 pointed out that the images from Condition B gave them impressions more directly based on their conversation than that of Condition A. Similarly, the participants of Group 4 remarked that unexpected images were displayed in Condition A. Moreover, the participants of Group 3 remarked that similar images were displayed intermittingly.
Table 3 shows the main ideas generated from each group. "They" means members of a couple in a long-distance relationship in the results of Theme Q, and it refers to elderly people in the results of Theme R. We cannot say that there is difference between the ideas in Condition A and B; however, some of the ideas were entertaining. In the results of Theme Q, one participant of Group 2 came up the idea to "count a prime

Table 2. Condition for experiment

Group	Condition	Theme
1	B	Q
	A	R
2	A	Q
	B	R
3	B	R
	A	Q
4	A	R
	B	Q

number to calm down." Before this idea, images of prime numbers were displayed, even though nobody had said the word "prime number." In the results of Theme R, one of the participants of

Table 3. Ideas generated by the participants

Theme	Group	Condition	Ideas
Q	1	B	Each of them saves money every day. They know how much the partner saves.
			Exchange a diary with pictures every other day.
	2	A	They count prime numbers to calm down each other.
			They write their e-mail messages in English in an effort to read the mind of the partner.
	3	A	They give each other notice when they are staying home.
			They share schedules to know convenient times to call.
	4	B	They convey their hearts to each other.
			They convey where they paid money to each other.
R	1	A	They can get points toward shopping at the mall when they evacuated immediately.
			They take a bus to evacuate.
	2	B	They receive notice that this neighborhood is a disaster caution area through TV commercials.
			A system shows that this neighborhood is a disaster caution area by using the simulation of disaster.
	3	B	A system displays a simulation of a disaster.
			A system causions those around elderly people through their smartphones.
	4	A	A system rings their fixed-line phones. Because these phones do not ring usually, they get a sense of crisis.
			If telephones will be installed in the gymnasium, even elderly people who do not a smartphone will evacuate.

Group 1 came up with the idea of "a bus for evacuation." The images in Condition A induced a new idea from one of the participants.

Excerpt 1 from Group 1's discussion (see below) shows the process of coming up the idea. They were discussing what kind of method encourages elderly people to evacuate to a safe space when a disaster happens. "A" and "B" mean one of the participants and the facilitator of Group 1, respectively.

Excerpt 1

01A: People who evacuate to a safe space should achieve some benefits.
02B: Aha! How about a locally limited couponing?
03A: Haha. What kind of coupon should we distribute for the people who evacuated to the safe space? Can we catch them?
(Participant A saw the display.)
04B: Well, how about an entrance coupon of the gate ball?
05A: Haha.
06B: They can be admitted for 500 yen. How about such a coupon?
07A: Yeah.
08B: We will make some coupons. I think a point-setting system is also good. They can exchange the service points with goods of supporting companies.
09A: It is a premium treatment.
10B: When we contrast these systems,
(The images of cars were displayed)
the people become achieve some benefits.
(Participant A saw the display.)
We may go ahead in the direction.
11A: A bus for evacuation!
12B: Aha!
13A: Ha ha.
14B: Do the bus go around?
15A: Go around. Haha.
16B: Yeah. The bus visits from door to door while calling for evacuation.

Participant A expressed the opinion that the elderly people may evacuate to a safe space if they will achieve some benefits (Line 01). Participant B proposed an idea of a locally limited coupon (Line 02). While Participant B explained a structure to realize such a coupon system, six images of various kinds of cars were displayed on the display (Line 10). Then, Participant A came up with the idea of "a bus for evacuation" that goes around and visits the elderly people's homes door to door (Line 11). Before displaying the car images, no participant said the word "car." We estimated that the word

"company" (Line 08) may have triggered the system to search the images of the cars.
Moreover, we could obtain feedback from the participants, given below.

- We did not see the display when we were having a good conversation.
- I received other viewpoints from the images.
- I do not think that the voice recognition was accurate.
- Similar images were displayed continuously.
- The images were not always helpful.
- We want the images when we run out of talking points.

SWISS can display the images intermittently, even if people are talking. However, when people stop talking, SWISS does not display the images. The students want the images even when they are silent.
Excerpt 2 from Group 2's discussion in Condition A (see below), shows an example of the discussion getting bogged down. "C" and "D" are the participants and "E" is the facilitator of Group 2.

Excerpt 2

(1 minute and 13 seconds after starting the session)
01C: Umm, in my case, I want to meet my girlfriend in a world of virtual reality (VR).
02E: Oh, I see.
03D: I see.
04E: It sounds good, the world of VR.
05C: In the world of VR.
06E: Oh, it sounds good.
07C: When I see her, well.
08E: Umm, there is your girlfriend in front of you.
09C: Yeah.
10E: You can see your moving girlfriend.
11C: Yeah, yeah, yeah.
(blank: 7.3 seconds)
12D: Uh huh.
13E: Umm.
14C: Umm.
(blank: 5.2 seconds)
15C: Let me see, is it a tool or an application?
16D: A tool or an application.
(blank: 10.9 seconds)
17E: Umm.
(blank: 7.5 seconds)
18C: When do they feel lonely?
(2 minutes and 20 seconds after starting the session.)

Participant C expressed the idea that a person can see his or her girlfriend or boyfriend in the world of VR (Line 01). Participant E tried to expand Participant C's idea by celebrating it and rewording Participant C's utterances. However, the participants could only fantasize that the girlfriend was there in front of Participant C using VR (Line 08-11). After these utterances, they repeated "Umm" which is considered "filler." Generally, it is said that the filler marks the hesitation of a speaker before starting his or her utterance [13]. However, in this case, there are no words that follow "Umm." Moreover, there is long, empty time during which SWISS could not display images.

4.4 Discussion

Most of the participants using SWISS noticed the difference between Conditions A and B (SWISS with and without the use of Bing Autosuggest API).
Although there was no apparent difference in the results between Condition A and B, some entertaining ideas were generated by the participants' discussion in Condition A.
On the other hand, it was found that we should reconsider the timing of the displayed images. Even when no participant expresses an idea, the participants want to see some images on the display.

5 IMPROVEMENT OF SWISS

5.1 Aim

SWISS is improved to display the images even while the participants are not talking.

5.2 Design

First, SWISS has to estimate the conversational activation level of the participants. Moriya estimated the degree of conversational activity by enthusiasm of the speakers during the conversation to decide the behavior of avatars

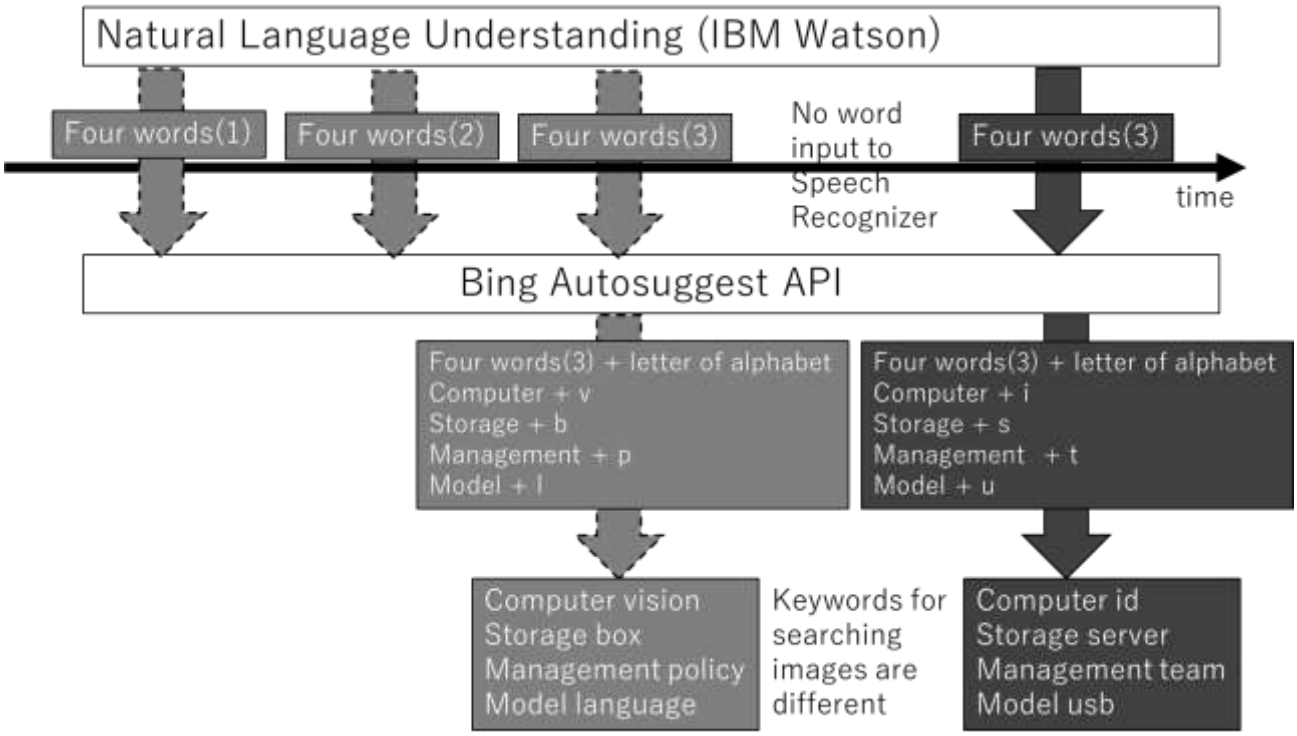

Figure 5. When no word is input to Speech Recognizer, Bing Autosuggest API uses the immediate previous four words and another letter selected at random.

(surrogates of the speakers) [14]. She analyzed 10 features that reflect the prosody of the speaker's utterances. The sound pressure and the pitch showed greater individual differences, although they were affected by factors in the system's environment (e.g., microphone sensitivity) [14]. Therefore, SWISS also uses the sound pressure and the pitch of the participants' conversation to estimate the conversational activity. When no word has been input for three seconds, SWISS displays the images based on the previous words.

Second, SWISS has to display the images even in silence. Therefore, SWISS has to search the images according to previous words. In the process of SWISS, the Watson NUL analyzes a website and extracts four words that represent its features. Then, each word is listed with a letter of the alphabet or a Hiragana character, which is selected at random, for using the autosuggest function (see Section 3). Accordingly, as shown in Figure 5, the previously used four words (Four words (3)) is re-listed with a letter, which is selected at random again (see a black part). Then, Bing Image Search API searches images with each of the four words and the new word suggested by Bing Autosuggest API.

Furthermore, if no word is input yet while SWISS is preparing to display the images or in three seconds after displaying the images, SWISS displays the image based on the other previous four words ("Four words (2)" in Figure 5).

6 DISCUSSION AND CONCLUSION

SWISS might be considered "mobile learning [15] [16]." When the university students discuss their research themes with their supervisor and laboratory associates, they try to speak with fluidity. However, they often are stuck on some problems, one of which is the inability to conceive new ideas.

In Excerpt 2, the participants uttered "umm" to each other. We believe that the participants should not give up conceiving new ideas for themes, and they want to express something for the sake of a continuous discussion. It is presumed that SWISS was available on the students' smartphones and one of the students put the smartphone in front of them during their discussion.

The results of the experiment suggested that displaying the images based on the extracted words and other words predicted from a letter of the alphabet or Hiragana via an autocomplete function (SWISS) was useful in the discussion for generating new ideas (see Section 4). Moreover, by improving the algorithm for operating SWISS, the system could display the

images even when no participant speaks for a while and no word is input to SWISS.
In the future, the participants should discuss some themes using the improved SWISS system to examine whether it is useful to the participants.

Acknowledgment

This work was supported by JSPS KAKENHI Grant Number 17H01950.

REFERENCES

1. Fujita, H. and Nishimoto, K.: Lovelet: A Heartwarming Communication Tool for Intimate People by Constantly Conveying Situation Data. CHI2004 (latebreaking results: poster), CD2, pp.1553. (2004).
2. Osborn, A.F.: Applied Imagination. 2nd ed., New York:Scribner, (1957).
3. Ohiwa, H., Kawai, K., and Koyama, M.: Idea processor and the KJ method. J. of Information Processing, vol. 13, no. 1, pp. 44—48. (1990).
4. Amitani, S. and Hori, K.: Knowledge Nebula Crystallizer for Knowledge Liquidization & Crystallization-from a Theory to a Methodology of Knowledge Management. In Proc. of Expertise In Design, Design Thinking Research Symposium, vol. 6. (2003).
5. Miura, M., Sugihara, T., and Kunifuji, S.: GKJ: Group KJ method support system utilizing digital pens. IEICE trans. on information and systems, vol. 94, no. 3, pp. 456—464. (2011).
6. Clayphan, A., Collins, A., Ackad, C., Kummerfeld, B., and Kay, J.: Firestorm: a brainstorming application for collaborative group work at tabletops. In Proc. of the ACM international conference on interactive tabletops and surfaces, pp. 162--171, November. (2011).
7. Tausch, S. Hausen D., Kosan, I., Raltchev, A. and Hussmann, H.: Groupgarden: supporting brainstorming through a metaphorical group mirror on table or wall. In Proc. of the 8th Nordic Conference on Human-Computer Interaction: Fun, Fast, Foundational, ACM, pp. 541--550, October. (2014).
8. Montag-Smit, T. and Maertz C. P. Jr: Searching outside the box in creative problem solving: The role of creative thinking skills and domain knowledge. J. of Business Research, vol. 81, pp. 1—10. (2017).
9. Nijstad, B. A., Stroebe, W., and Lodewijkx, H. F.: Cognitive stimulation and interference in groups: Exposure effects in an idea generation task. J. of experimental social psychology, vol. 38, no. 6, pp. 535--544. (2002).
10. Wang, H. C., Fussell, S. R., and Cosley, D.: From diversity to creativity: Stimulating group brainstorming with cultural differences and conversationally-searched pictures. In Proc. of the 2011 ACM conference on Computer supported cooperative work, pp. 265--274, March. (2011).
11. Wang, H. C., Cosley, D., and Fussell, S. R.: Idea expander: supporting group brainstorming with conversationally triggered visual thinking stimuli. In Proc. of the 2010 ACM conference on Computer supported cooperative work, pp. 103--106, February. (2010).
12. Shibata, D., Yamaguchi, Y., Oshima, C., and Nakayama, K.: Continuous Display of Images Searched by Keywords Extracted from a Brainstorming Session and Suggested by an Autosuggest Function. Proc. of ICESS2018, p.76--80. (2018).
13. Corley, M., and Stewart, O. W.: Hesitation disfluencies in spontaneous speech: The meaning of um. Language and Linguistics Compass, vol. 2(4), pp. 589--602. (2008).
14. Moriya, Y., Tanaka, T., Miyajima, T., and Fujita, K.: Estimation of conversational activation level during video chat using turn-taking information. In Proc. of the 10th asia pacific conference on Computer human interaction, ACM, pp. 289--298. (2012)
15. Al.Ragawi, N. and Zahary, A.: Obstacles and Opportunities for Yemeni Students to Use Mobile Learning. The International Journal of E-Learning and Educational Technologies in the Digital Media, vol. 3(4), pp.161--173, The Society of Digital Information and Wireless Communications (SDIWC) (2017).
16. Samsudin, M. R., Guan, T. T., Yusof, A. M., Rahim, N. N., and Mamat, S. A.: Conceptual Framework in Teaching and Learning Using Mobile Application for Special Students. The International Journal of E-Learning and Educational Technologies in the Digital Media, vol. 4(1), pp.1--10, The Society of Digital Information and Wireless Communications (SDIWC) (2017).