AN IMAGE RECOMMENDER SYSTEM THAT SUGGESTS APPROPRIATE IMAGES IN CREATION OF SELF-LEARNING ITEMS FOR ABSTRACT NOUNS

MOHAMMAD NEHAL HASNINEA, YUKI HIRAI, MASATOSHI ISHIKAWA, HARUKO MIYAKODA, KEICHI KANEKO, LYN PEMBERTON

Graduate School of Engineering, Tokyo University of Agriculture and Technology, Tokyo, Japan
Faculty of Business Administration, Tokyo Seitoku University, Tokyo, Japan
Department of English, Tsuda College, Tokyo, Japan
School of Computing, Engineering and Mathematics, University of Brighton, Brighton, UK
E-mail: 50015834303@st.tuat.ac.jp, yhirai@cc.tuat.ac.jp, ishikawa@tsu.ac.jp, miyakoda@tsuda.ac.jp, k1kaneko@cc.tuat.ac.jp, Lyn.Pemberton@brighton.ac.uk

Abstract- This article acquaints its readers with a research paradigm that has been designed to extract appropriate (synonymously educational) images of abstract nouns. The ongoing investigation principally purposed to assist L2 learners/educators by recommending appropriate images in the creation of vocabulary learning items for abstract nouns. Hence, a prototype version of an image recommender system that purposed to assist the users in the recommendation of appropriate images for 3-types of abstract nouns has been implemented and tested. The proposed system, in the process of learning items creation, allows the users to select their own preferred image if the system recommended topmost image is not satisfactory. Our study presumed that still images having physical or concrete existence can be addressed as appropriate learning resources in the representation of abstract nouns that firstly, represent social contexts between human, secondly, related to feeling and emotion, and thirdly, state social or religious belief. Authors agree that due to huge cultural influences and multiple behaviors of abstract nouns, this hypothesis may be a matter of debate. Therefore, to assess the images, an image evaluation experiment has been conducted with 20 participants who are actively engaged in foreign language acquisition. A post-hoc analysis of Tukey’s test revealed the significant difference (P=0.04) of our system- recommended images over Yahoo-suggested images in learners’ considerations as appropriate image resources to memorize new vocabulary.

Index Terms- Abstract Noun’s Image, Educational Technology, Image Recommender System, Learning Item Creation Tool

I. INTRODUCTION

In the area of second language acquisition (SLA), it has been revealed that vocabulary is central to language, and of critical importance to the typical language learners [1]. Knowledge of vocabulary is also considered as the backbone of learners’ competency that facilitates learning of any language tasks [2]. Lack of significant L2 vocabulary can be the reason of miscommunication or poor communication. Significant vocabulary can arguably differentiate the nervous speakers and expert speakers. In addition, deficiency in vocabulary can make a learner’s learning process slower. Nevertheless, the teaching and learning of L2 vocabulary have been undervalued in the field of SLA. One of the reasons behind is that SLA researchers and teachers have typically prioritized syntax and phonology as “more serious candidates for theorizing”, more central to linguistic theory, and more critical to language pedagogy [1].

Quick forgetting of learned vocabulary is another inevitable problem that L2 learners often experience. Research on the learning approaches [3][4][5] revealed that intentional vocabulary learning is the key for most of the EFL (English as a Foreign Language) learners’ vocabulary expansion because new words are difficult and slow to acquire without ambiguity out of context [6]. Learning new vocabulary in the classroom is considered less effective in the acquisition of intentional vocabulary. As a result, L2 learners worldwide are struggling significantly to acquire L2 vocabulary efficiently. Therefore, the popularity of self-learning systems has increased rapidly. As the demands increase, a numerous number of web-based language learning (WBLL) tools and mobile applications have been introduced.

This study is partly supported by a Grant-in-Aid for Young Scientists (B) of the Japanese Society for the Promotion of Science (JSPS) under Grant No. 15K20980

Nonetheless, the confusions remain unsolved when it is a matter of the representation of learning resources. Several representation methods of learning materials have been tested previously. Due to the inadequate IT-skills, learners from non-IT backgrounds are still struggling to cope with the generation of learning materials. In self-learning, learning items that are generated by using meaningful images are known as the best learning resources for noun acquisition. However, finding meaningful still images has been a challenge in CALL (Computer-Assisted Language Learning) research. A highlighted argument amongst L2 researchers is to find the educational images for abstract words. The main issue every researcher faces is the lack of clearly defined term ‘abstract’ [7]. Natalia [7] also argued that despite numerous approaches to distinguish between abstract and concrete nouns; linguists have so far failed to come to
the unanimous understanding of their categories. Some words are too abstract that those still remained unclear to the human mind. Accordingly, finding appropriate (synonymously educational or suitable) images for abstract nouns is extremely difficult, and also a matter of debate. Due to that, to find educational images, foreign language learners and educators around the globe are suffering significantly.

To overcome the research gaps, our study proposed an approach that arguably suggests educational images to the learners. Furthermore, an on-demand vocabulary learning system has been developed that assists learners in self-learning. The principal objective of our study is to assist foreign language learners in the acquisition of foreign language learning by self-creation of learning items.

II. LITERATURE REVIEW

Throughout the last decades, significant amounts of web-based learning and mobile learning platforms have been developed that assist learners in the self-creation of learning items. PhotoStudy[8] was developed with the concept of collaborative usage of the images that were captured by camera-phones. However, the researchers acknowledged that collaborative policing of the learning materials was victimized by email spam, and a huge volume of incorrectly marked up images was found which proved problematic for learners. UEVL[9] system was developed with the purpose to assist students in vocabulary learning by using ubiquitous technologies. This system used dynamic video clips to generate learning materials. UEVL was able to detect student’s locations through radio-frequency identification and global positioning system technologies. Yet, the study did not investigate the learning effects of UEVL system created learning materials. A study conducted by Wong et al. [10] tested their mobile assisted content creation approach on school children to acquire Chinese idioms and English prepositions. The study, however, does not suggest any guidelines on the nature of the still images that have been used.

Afterward, Orapin et al. [11] developed a system equipped with a dynamic and interactive interface that allows vocabulary learners to learn vocabulary by using a mnemonic technique. The system allows users to seamlessly browse a collection of foreign words while suggesting phonetically related words of a known language for helping the memorization of unfamiliar languages. However, no significant differences were noted when compared this approach with a traditional way (dictionary-based learning) and static visualization (images are displayed statically without any special interaction) approaches respectively.

PSI [12] is a web-based vocabulary learning system that has functions of creating English vocabulary learning materials based on short movies. Later on, an autonomous and collaborative vocabulary-learning environment called SIGMA [13][14] was developed support PSI’s operations. However, in the process of learning material creation, all the resources (including meaning, pronunciation data, and image) needed to be provided by the user. To our observation, accumulation of the resources prior to learning material generation is very problematic and time-consuming, which may lead to slow learning. Additionally, with respect to learning effectiveness, their studies [12][13][14] failed to show any significant differences over traditional methods.

In the context of L2 vocabulary acquisition, the effectiveness of still images and dynamic videos has been revealed. Researchers who view on noun acquisition believe that still images are more appropriate than dynamic videos while acquiring new nouns. Oxford [15] argued that meaningful visual images make learning more efficient because it helps learners associating to their memory. The idea of the universally accepted images has been introduced by Kaneko et al. [16]. However, their study [16] did not clarify the nature of the images that can be considered as universal images.

In our study, we have developed a platform called AIVAS [17]. The AIVAS generates a 5-second-long learning item for each noun. AIVAS reduces considerable loads on learners/educators by automatically extracting learning resources from the web engines upon user demand. Furthermore, AIVAS generates learning items for 11-languages, which gives learners to be capable of learning several languages. AIVAS generates disposable learning items; therefore, it reduces a significant amount of loads on databases. Lastly, to solve the difficulties related to finding educational images for abstract nouns, a web-based image recommender system called AIVAS-IRA has been implemented. This article emphasizes on our proposed approach that extracts appropriate images for abstract nouns.

III. THE STUDY: IMAGE EXTRACTION OF ABSTRACT NOUNS

Appropriate images can arguably help learners in recalling their old memories and keep the newly learned words longer than usual. Because of that, the importance of appropriate images in image-based noun acquisition is significant. Previously conducted studies [7][16] pointed out some difficulties involved in finding right images for abstract nouns. The researchers agreed that the cultural influence is one of the main reasons behind that. Another common problem that researchers often face is the polysemic behavior of a single noun depending on the context. While finding the common pattern between different situations, our study considered the basic use (first symbolic meaning) of an abstract noun. That is when we talk about ‘love’: the first symbolic expression that comes to human mind is the affection between two
persons. Therefore, our research consideration was to find an image, which represents ‘love’ as a symbol of affection. Our study also considers the limitations of technology to determine the appropriateness of an image in the representation of an abstract noun. The principle objective of our study is to equip AIVAS [17] users with a hassle-free creation of learning items without relying on any other external image resources. Therefore, our study worked on a noble approach to extract appropriate (synonymously educational or suitable) images for abstract nouns by using our image re-ranking system.

This particular section talks about our proposed strategy that has been targeted for abstract nouns. Figure 1 shows our research paradigm.

A. What is an Abstract Noun? The Definition

In general definition, the term ‘abstract’ hints of something intangible. Cambridge dictionaries articulated the definition of an abstract noun as ‘a noun that refers to a thing that does not as a material object’. Along with that, many linguists defined the term as their ways. However, the ambiguities related to the definitions remain unclear.

This ongoing investigation defines abstract noun based on a noun’s abstractness-concreteness value measured by Allen et al. [19]. A study conducted by Allen et al. [19] scaled the Abstractness-Concreteness (C), Imagery (I), Meaningfulness (m), and Frequency (F) measurement of a total 925 nouns. Their research scaled 925 nouns on 7-point scales. To state a proper definition of the abstract noun, our study considered each noun’s Abstractness-Concreteness(C) values as the parameter.

Nouns with Abstractness-Concreteness(C)’s mean (M) value is less or equal than 4, have been considered as abstract nouns. A noun with C’s mean (M) is more than 4 were not treated as an abstract noun.

Abstract Nouns = C (M≤4)

The remaining sections talk about the nouns that meet our definition of abstract nouns. To clarify the confusions, some examples have been shown in Table 1.

Nouns in ‘○’ shown in Table 1 meet the definition of an abstract noun and are taken into consideration in the extraction of appropriate images. Nouns in ‘X’, even though abstract in the general sense, have not been considered as abstract nouns. However, those nouns will certainly be investigated in the future studies. Our study defined the term ‘abstract’ based on English abstract nouns’ C-value because the proposed technical platform searches the corresponding image for any language’s word into its corresponding English translation.

B. Targeted Abstract Nouns

Generalization of abstract nouns is very difficult due to the massive variations of the corresponding images for each noun. It is also very problematic to establish common patterns between different images. Additionally, due to the heavy cultural influences, individual’s view on educational images may vary. Furthermore, due to the polysemic behavior, the limitations of the image search engines have been well reflected in its searching output. Therefore, at the beginning, determining the targeted noun sets was essential. The term ‘targeted’ abstract nouns are those nouns that the hypothesis was made and the system has been tested with.

This particular investigation has been limited to three types of abstract nouns. Prior to setting the target abstract nouns, the nature of abstract nouns has been

---

**Figure 1 Research Paradigm.**

**Table 1 Example of Abstract Nouns.**

<table>
<thead>
<tr>
<th>Noun</th>
<th>Advice</th>
<th>C(M=2.08)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Dream</td>
<td>C(M=3.03)</td>
</tr>
<tr>
<td></td>
<td>Love</td>
<td>C(M=1.80)</td>
</tr>
<tr>
<td></td>
<td>Offshoot</td>
<td>C(M=4.20)</td>
</tr>
<tr>
<td></td>
<td>Pollution</td>
<td>C(M=4.14)</td>
</tr>
<tr>
<td></td>
<td>Lord</td>
<td>C(M=4.18)</td>
</tr>
</tbody>
</table>

**Table 2 Targeted and Non-Targeted Abstract Nouns.**

<table>
<thead>
<tr>
<th>Type</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 (Targeted)</td>
<td>Advice, Opinion, Joke, Idea, Recognition, Strength, Answer, and Knowledge etc.</td>
</tr>
<tr>
<td>2 (Targeted)</td>
<td>Love, Hate, Victory, Freedom, Greed, Anger, Ego, Expression, and Betrayal etc.</td>
</tr>
<tr>
<td>3 (Targeted)</td>
<td>Blessing, Dream, Heaven, Hell, Death, Ghost, and Devil etc.</td>
</tr>
<tr>
<td>Others (Non-targeted)</td>
<td>Fly, Hot, Pleasure, Amount, Attitude, Citation, Code, Deduction, Poetry, Trouble, Illusion, Impact, Impulse, Array, Fact, Fault, Deed, Science, History, Profession, Capacity, and Equity etc.</td>
</tr>
</tbody>
</table>
well investigated. The three types of target abstract nouns are, 1) basic abstract nouns that represent social contexts between humans, 2) abstract nouns that are related to feeling and/or emotion, and 3) abstract nouns that represent our social and religious beliefs. Other types of abstract nouns will be investigated in the future studies. Table 2 shows some examples of the targeted nouns and non-targeted abstract nouns.

C. What are Appropriate Images?
Image search outputs given by the standard image search engines for text-based query especially for abstract nouns have been well observed. While analyzing the nature of still images, numerous types of images have been observed. In the field of image-based learning, massive variations in the images may create confusion and distraction on learners’ minds. Our observations on the nature of searching outputs given by standard image search engines are that, while searching images for abstract nouns, three types of images are often found in the topmost positions. Our study identifies those types of images as ambiguous and inappropriate learning resources for the purpose of image-based vocabulary acquisition. Our study proposed solutions to those ambiguous and inappropriate images. The proposed solutions to those three cases are as follows:
Case 1: A still image that contains textual data (texts/characters) in the image frame is supposedly inappropriate and ambiguous. Those types of images often depict non-native English speakers and learners with poor English ability. In addition, a numerous number of images contain irrelevant characters in the image frame. Suggesting those images may cause learners’ learning process slower and unattractive. Our study presumes that images having humanoid expressions or objects in the image frames can be considered as appropriate replacements over to the images with textual data.
Case 2: Still images that are visually too abstract are most likely inappropriate learning resources. These types of images are often observed in searching outputs. The term ‘too abstract’ refers to those images with neither textual data nor concrete expression in the frames. This type of images often creates confusion on the learner’s mind. The proposed alternatives to those are the images that contain some level of humanoid/object existence in the frame.
Case 3: Another type of inappropriate image is, the still image that contains multiple objects including the existence of human, object, animation, and textual data in the image frame. These types of images, to represent an abstract noun, can be a reason of confusion and distraction on the learner’s mind. Alternatively, images that are simple in nature (possibly containing minimal objects in the frame and less background noise) can be considered as meaningful appropriate images.
To establish a common pattern between images in all three cases, our study presents the definition of ‘appropriate’ images for any noun that belongs to targeted abstract noun sets. The definition of an appropriate image for the targeted abstract nouns is, an image having physical or concrete existence possibly positioned in the central position in the image frame. The proposed technological framework has been designed in the way that the outputs meet the definition of appropriate images.

IV. THE TECHNOLOGICAL FRAMEWORK
A. The AIVAS: A Web-based Learning Item Creation System
A web-based on-demand learning system called the AIVAS has been introduced, which creates learning items based on user’s demand. The AIVAS (acronyms for Appropriate Image-based Vocabulary Acquisition System) functions include, finding the meaning of a noun, image extraction, voice data extraction, and learning item creation simultaneously. A complete description of the architecture of the AIVAS’s frontend and backend implementation, and the features has been released previously [17].
A learning item (synonymously learning material or learning object) for each noun is a representation of a 5-second-long learning resource that is delivered to the learners. The components of a learning item are, the meaning of a word to be learned, a corresponding image, pronunciation data, and the spelling. The format of the learning items, sample learning items, and the representation method and necessary discussions have been published [17][18].

B. AIVAS-IRA: The Image Recommender System
AIVAS-IRA is a web-based image recommender system that has been designed to assist the users of AIVAS in the process of learning item creation. AIVAS-IRA underpinned other existing systems because the system assists L2 learners/educators by recommending an appropriate image for each noun while they create self-learning item. The system also recommends a certain number of relevant images for each noun, by which learners can select their own image preferences if the AIVAS-IRA-extracted image is not satisfactory.
The pseudo code of the IRA’s early version has been released [17]. In the article [17], a brief discussion on the algorithm implementation has also been articulated. However, several modifications have been done from the early version. In our study, sample ‘good’ images accumulation was done based on experimental data. This can be addressed as one of the main modifications of the current version of the system over the early version. Additional functions have been added to the early version. At present, the system recommends images for both basic concrete and basic abstract nouns.

C. AIVAS-IRA Implementation and Output Testing
Learners suggested sample ‘good’ images have been accumulated to determine the centroid of images. The
An Image Recommender System That Suggests Appropriate Images In Creation Of Self-Learning Items For Abstract Nouns

centroid of ‘good’ images is the parameter to determine the appropriateness of other images. An image that is closest to the centroid of ‘good’ images is the most appropriate, and vice versa. An individual’s cultural background heavily influences the images for abstract nouns. Therefore, learners suggested sample images have been prioritized over authors suggested images. An experiment with 6 participants has been conducted to accumulate sample ‘good’ images. Participants were native speakers of 6 different languages and bearing very different cultural backgrounds. All 6 participants in this experiment are engaged actively in the acquisition of foreign languages. In the accumulation of sample ‘good’ images, 14 English abstract nouns have been used as the representative nouns. All 14 English abstract nouns meet our definition of abstract nouns.

A total of 73 sample ‘good’ images for 14 abstract nouns have been accumulated. However, during image analysis, 5 images had been discarded due to the unacceptable formats of the images. Hence, a total of 68 images have been counted as sample ‘good’ images and tested the system with those. Table 3 shows the details on the participants, nominated words used, and the number of images accumulated for each word.

Table 3 Details on Sample Image Accumulation Experiment.

<table>
<thead>
<tr>
<th>L2s</th>
<th>Words Provided</th>
<th>Images Gathered</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lao</td>
<td>Anxiety</td>
<td>5</td>
</tr>
<tr>
<td>Khmer</td>
<td>Love</td>
<td>10</td>
</tr>
<tr>
<td>Pashto</td>
<td>Freedom</td>
<td>10</td>
</tr>
<tr>
<td>Japanese</td>
<td>Dream</td>
<td>5</td>
</tr>
<tr>
<td>Arabic</td>
<td>Ability</td>
<td>6</td>
</tr>
<tr>
<td>English</td>
<td>Italian</td>
<td>3</td>
</tr>
</tbody>
</table>

To demonstrate an output of the AIVAS-IRA, 8 randomly selected images for the abstract noun ‘idea’ have been downloaded. The reason for using randomly selected images is the copyright law set by image search engines. Figure 2 shows the input images (uppermost), the re-ranking environment in command prompt (in the middle), and the output images in re-ranked order (lowermost), respectively.

The image image0.jpg in the re-ranked folder is the appropriate image found from the 8 corresponding images that were taken as input for the word ‘idea’. Therefore, we argue that, for acquiring foreign vocabulary, the image0.jpg is the best image that can be used in the generation of a self-learning item for the word ‘idea’.

V. IMAGE EVALUATION EXPERIMENT

D. Method and Material

To assess the AIVAS-IRA-recommended images for abstract nouns, at first, six English abstract nouns (sadness, pain, recognition, idea, hell, and angel) have been selected. Secondly, 16-top ranked images (.jpg formatted images only) for each noun have been downloaded and re-ranked them by using AIVAS-IRA. Google Image Search APIs have been used as the source of the images that have been downloaded. Finally, a survey questionnaire has been prepared by comparing AIVAS-IRA versus Google versus Yahoo UK suggested top-ranked image for each noun. The top-ranked images (from both Google and Yahoo UK) for the selected nouns have been gathered on the same day the outputs have been tested. It can be mentioned that top-ranked images in both Google and Yahoo can be changed, periodically.

During the survey, 20 participants were asked to rate each image on a 5-point scale based on their assessment of two factors: firstly (Q1), appropriateness (how appropriate) of the image to the corresponding noun, and secondly (Q2), as a learning resource to memorize the word. That is, a maximum of 10-points (5-points for each factor) was assigned to each word. There was no time restriction during the survey. That is, the participants could spend as much time as they wanted to return the feedback form.
Participants could select from a paper form and an online form to submit the feedback for the survey questionnaires. It has been observed that an average of 5 to 7 minutes of time required getting each participant’s feedback.

E. Result

Tukey’s test was used for comparison of the multiple measurements that have been made during the image evaluation experiment. Statistical analysis revealed a significant difference between the AIVAS-IRA and Yahoo in the area of Q2 (P=0.04; post-hoc analysis Tukey’s test). However, no difference is present between the AIVAS-IRA versus Google in the area of Q2 (P=0.35). Besides, no significant differences were noticed between the AIVAS-IRA versus Google, and AIVAS-IRA versus Yahoo in the area of Q1 (P=0.92 and P=0.24, respectively).

VI. DISCUSSION

This article reports on three research questions, firstly, what kind of still images can be considered as appropriate learning resources for acquiring abstract nouns? Secondly, is a technological platform able to recommend appropriate images for abstract nouns? Thirdly, if so, whether or not those images will possibly fulfill learners’ expectations? For that reason, this article presents with i) the definition of abstract nouns, and ii) the definition of appropriate images for three-types of abstract nouns. To extract the appropriate images for abstract nouns, an image recommender system called AIVAS-IRA has been implemented. The AIVAS-IRA assists AIVAS users by suggesting the best possible image for each noun to memorize. Additionally, the image recommender system is able to recommend some relevant images to the learners. By that way, the learners will be helped significantly to understand the nouns that have multiple meanings. For example, the Japanese word 𨱂 (kin) can be used to express gold, money, metal, cash and currency. An evaluation experiment with 20 participants of 9 nationalities has been conducted to assess AIVAS-IRA suggested top-ranked images for abstract nouns. Result revealed a significant difference between the AIVAS-IRA and Yahoo in the area of Q2 (P=0.04; post-hoc analysis Tukey’s test). Hence, this article yield conclusion that, AIVAS-IRA is able to recommend educational images for abstract nouns that will help foreign language learners in quick memorization and long-term memory retention.

VII. LIMITATIONS & FUTURE DIRECTIONS

Few improvements of the ongoing investigation are under consideration. The assessment of AIVAS-IRA-recommended images with respect to both short-term and long-term memory retention will be investigated in future studies. Our study limits to three particular types of abstract nouns. The proposed definition of ‘appropriate’ images may be a matter of debate and may not be a suitable definition for other types of abstract nouns. We consider that as one of the limitations of the current work. At this moment, the prototype version of both AIVAS and AIVAS-IRA may not be visually attractive and user-friendly. Learners with inadequate IT-skills may find the systems’ operations rather complex. User-friendly interfaces will be designed before releasing the software to the marketplace.

REFERENCES

An Image Recommender System That Suggests Appropriate Images In Creation Of Self-Learning Items For Abstract Nouns