

DEVELOPMENT AND DESIGN OF A COMPOUND VERB AR LEARNING SYSTEM EMPLOYING IMAGE SCHEMAS

Xuewang Geng¹ and Masanori Yamada²

¹Graduate School of Human-Environment Studies, Kyushu University

²Faculty of Arts and Science, Kyushu University
744 Motoooka, Nishi-ku, Fukuoka 819-0395, Japan

ABSTRACT

Recently, the number of learners of Japanese as a foreign language (JFL) has been increasing. In Japanese language acquisition, compound verbs (verbs that are composed of two verbs, e.g., *tobikomu* ‘jump into’) are frequently used in daily life; these present difficulties, including unclarity of combination and opacity of meaning. Matsuda (2001) proposed an image schema that applies core theory to Japanese compound verbs and the application of image schemas to Japanese language education. However, since image schemas are composed of simple diagrams and arrows and are highly sophisticated due to being intended for linguists, it has been suggested that they are not easy for JFL learners to understand (Tagawa and Yuizono, 2016). In this paper, we designed and developed a compound verb AR learning system with core theory and image schemas. Moreover, we discussed Japanese compound verbs acquisition, the image schema of compound verbs as well as the application of AR in learning, and also explained the design and development of the system based on problems in compound verb acquisition.

KEYWORDS

Augmented Reality, Compound Verbs, Image Schema, Mobile Learning, Japanese Acquisition

1. INTRODUCTION

In recent years, the number of learners of Japanese as a foreign language (JFL) has increased. According to the Survey Report on Japanese-Language Education Abroad by the Japan Foundation (2017), it was found that the number of Japanese learners had increased significantly from 1979 to 2015. Compound verbs are an important learning item in Japanese acquisition; they are formed from two single verbs, e.g., *tobikomu* (‘jump into’, *tobu* and *komu* are simple verbs). As they are frequently used in daily life, in many cases, it is necessary to understand their meanings.

However, it is difficult for JFL learners to master them even if they reach an advanced level (Matsuda, 2002). The difficulties include the unclarity of combination and opacity of meaning. Specifically, unclarity of combination means that the combination order of two single verbs and the existence of a compound verb are uncertain (Sano, 2005). That is to say, JFL learners do not know which two verbs can form a compound verb or the sequence of the two single verbs, and thus it is difficult for them to remember compound verbs. Moreover, opacity means that the meaning of a compound verb cannot be inferred from its two single verbs, because the meaning of a compound verb is not always the combination of the meanings of the single verbs (Chen, 2007). Therefore, learners cannot simply infer the meaning of compound verbs. Furthermore, it has also been pointed out that it is hard to distinguish the differences in meaning between single verbs and compound verbs (Matsuda, 2000). For example, the single verb *yobu* ‘call’ and the compound verb *yobikakeru* ‘call on’ are easily misinterpreted as having the same meaning, though *yobikakeru* has a distinct meaning from *yobu*.

Matsuda (2002) applied a cognitive semantic method to explain the various complex meanings of Japanese compound verbs and suggested an “image schema,” which is an image of the knowledge structure abstracted from perceptual and motor activities. While an image schema can help JFL learners to understand the meaning of a compound verb through a single image, it is not easy for them to understand, because an image schema is simply composed of abstract diagrams and arrows (Tagawa and Yuizono, 2016). Hence, it

might be concluded that it is undesirable to directly apply an image schema for JFL learners to learn compound verbs. On the other hand, employing augmented reality (AR) technology can effectively teach concepts that are difficult to understand by displaying virtual visual information in the real world as an innovative educational tool. Thus, we designed and developed a compound verb mobile learning system.

The system employed 3D animations to express the meanings of single verbs and compound verbs via augmented reality (AR), based on core theory and an image schema. In this system, learners learn the meaning of the single verbs first and then the compound verbs through combinations of verb cards (see Figure 1). In this way, the meanings of compound verbs and single verbs can be distinguished, and the system can also determine whether the combination is correct.

In this paper, the authors present image schemas of Japanese compound verbs, the augmented reality for learning them, and the system design and implementation.



Figure 1. Verb card of tobu (kanji and hiragana of the verb)

2. IMAGE SCHEMAS OF JAPANESE COMPOUND VERBS

Cognitive scientists have made proposals to illustrate the process of abstraction from specific sensorimotor experiences to abstract concepts. There are many linguists and psychologists (such as Lakoff and Langacker) in the field of linguistics in cognitive science. This field is called cognitive linguistics (CL). Its purpose is to explore the model or connection of linguistics structures with human embodied experience, conceptual knowledge, and the communicative function of discourse (Gibbs, 2006). The image schema is a crucial concept in cognitive linguistics. It is a structured representation of various experiences based on our bodily orientations, movements, and interaction. For example, the image schema for the preposition *in* is a schema of a container to indicate that something contains something.

Image schemas are used pedagogically. For example, in Benjamin's (2012) study, learners of English as a Foreign Language (EFL) used their own imagination to draw image schemas of the phrasal verbs. The results of this approach show that confusion regarding phrasal verb usage was reduced as a result of drawing and collecting image schemas, suggesting that the attempt to teach phrasal verbs emphasizing conceptualization via an image schema is valuable.

There is a "core theory" developed from image schemas. Bolinger (1977) argued that if the form of a word is different, the meaning is different, and the meaning is common if the form is the same. Based on this, Tanaka and Matsumoto (1997) suggested a "core theory" that assumes a schema covering the whole ambiguous sense, and argued that an ambiguous usage can be explained by focusing on and converting core schemas. Core means a context-independent and overarching meaning. Hence, we are allowed to adjust the core via context, whereby the polysemy of vocabulary arises. Although the core theory was originally proposed to support Japanese learners of EFL for the polysemous verbs of English, Matsuda (2001) proposed the application of Tanaka's core schema to Japanese compound verbs and the use of image schemas in Japanese language education. For example, the single verb *tobu* 'skip' is represented by the movement of the arc in Figure 2, and the single verb *komu* 'enter' refers to movement to the inside, as indicated by the arrow in Figure 3. The compound verb *tobikomu* in Figure 4 overlaps the images in Figure 2 and Figure 3, and thus represents the meaning 'jump in'. In this way, the image schema of a compound verb can be

comprehensively comprehended through a single image, opening a new possibility for aiding vocabulary acquisition. However, the image schema is composed of simple diagrams and arrows and is highly sophisticated due to its use by linguists. Therefore, rather than presenting an image schema directly to a learner, it is believed that presenting visual glosses would result in a better learning effect (Sato, 2016).

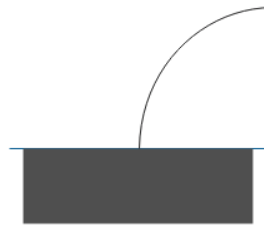


Figure 2. Image schema of 'tobu'

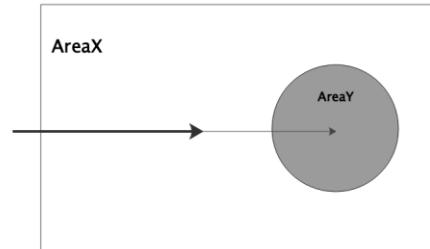


Figure 3. Image schema of 'komu' (A type)

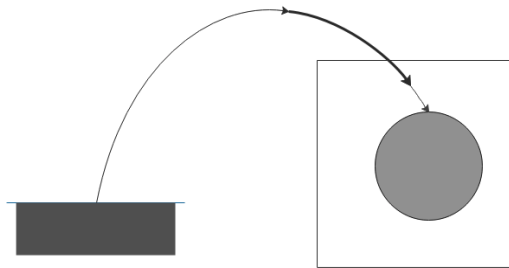


Figure 4. Image schema of 'tobikomu'

3. AUGMENTED REALITY FOR LEARNING

Recently, various learning environment designs have become possible with the spread of mobile devices such as smartphones. Augmented reality (AR) is one graphic technology for which learners need no special equipment and through which they can experience the content easily and efficiently. AR is defined as a simultaneous combination of the real world and virtual objects (Ibanez et al., 2016; Sin and Zaman, 2010). By applying AR, abstract concepts and complicated problems can be effectively taught (Walczak et al., 2006). There are several other advantages. For example, as it does not require a specialized device, its cost is cheaper than VR, and visualization, simulation, and interaction with virtual objects become possible. With these advantages, it is possible to provide new educational tools, showing that AR technology has the potential to greatly improve educational outcomes (Chiu et al., 2015).

AR is generally categorized as comprising location-based and image-based systems. Location-based AR systems use data relating to the location of the mobile device via a GPS or Wi-Fi based positioning system. The user can also move while using the mobile device in the actual environment. The information created by the computer depends on the location of the user. Image-based AR systems focus on image recognition techniques used to determine the proper location of virtual content relative to physical objects in a real environment. Since AR has advantages such as portability and low cost, it is applied to various fields such as ubiquitous learning (Dede, 2011) and cognition (Specht et al., 2011).

Language learning environment research using the graphic technology represented by AR has also been carried out. Mashiro et al. (2011) created an English vocabulary learning support system by using AR markers corresponding to the letters of the alphabet to arrange them. In Japanese language learning, Maekawa et al. (2015) proposed a learning system for the Japanese phonograms by arranging two to three

AR markers written in hiragana. AR can be used to improve present learning methods by annotating audio, text, and 3D images to objects in real environments.

In order to solve the problem of compound verb acquisition, we designed and developed a compound verb AR learning system based on core theory and image schemas.

4. SYSTEM DESIGN AND IMPLEMENTATION

This study focuses on an image-based AR using physical object tracking in smartphones. That is, the learner scans the card with the verb in the smartphone application, and the corresponding meaning animation is displayed on the card through the screen. In the study of the AR phonogram learning system above (Maekawa et al., 2015), the learners' impression was that it was easy to imagine characters and reading by relating AR animation to characters and reading. Thus, in this study, the animation is displayed on the text; that is to say, learners can also touch the text while grasping the meaning, so it is expected that the system can also promote familiarity with verb characteristics.

4.1 The System and Compound Verb Acquisition Problems

The following design was applied to the problems of compound verb acquisition.

4.1.1 Unclarity of Verb Combinations

In order to deal with the lack of clarity regarding compound verb combination, we designed a function named the combination judgment function to present the correct and incorrect order in verb combinations. When the learner combines the two cards for single verbs V1 and V2, the system determines whether the card order is correct or not (the order of verbs, whether this compound verb exists or not). If it is incorrect, the system presents the message, "The combination of compound verbs is incorrect." On the other hand, if it is correct, an animation of the compound verb's meaning is displayed.

4.1.2 Opacity of Compound Verb Meanings

The meaning of a verb is represented by a visual gloss of 3D animation created according to the image schema (see Figure 5, Figure 6, and Figure 7) in the system. In particular, the display of compound verbs was designed based on the V1 + V2 strategy (Matsuda, 2004), which seeks to convey an understanding of the meaning of compound verbs by combining single verbs V1 and V2 after understanding the meanings of V1 and V2. In addition, as shown in Figure 8, learners first learn the meanings of the single verbs V1 and V2 and then learn the meaning of compound verbs by combining the cards for single verbs. Therefore, it is possible to understand the semantic distinction between single verbs and compound verbs via the above function. Moreover, it is possible to infer the meaning based on the context because it is based on the image schema.

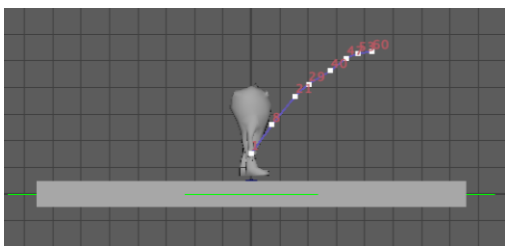


Figure 5. The motion track of the *tobu* animation

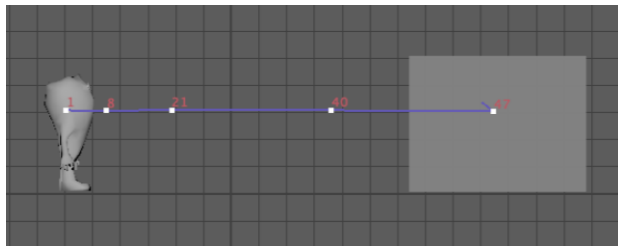


Figure 6. The motion track of the *komu* animation

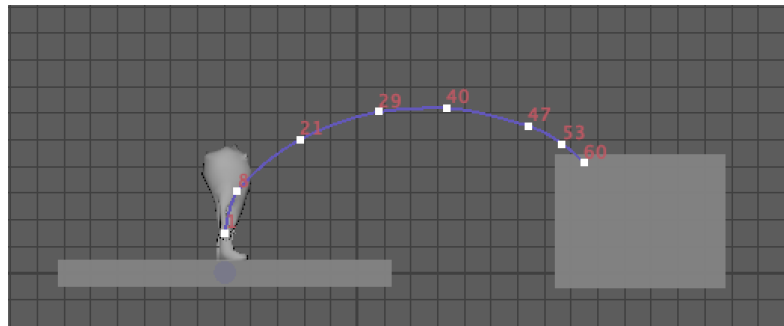


Figure 7. The motion track of the *tobikomu* animation

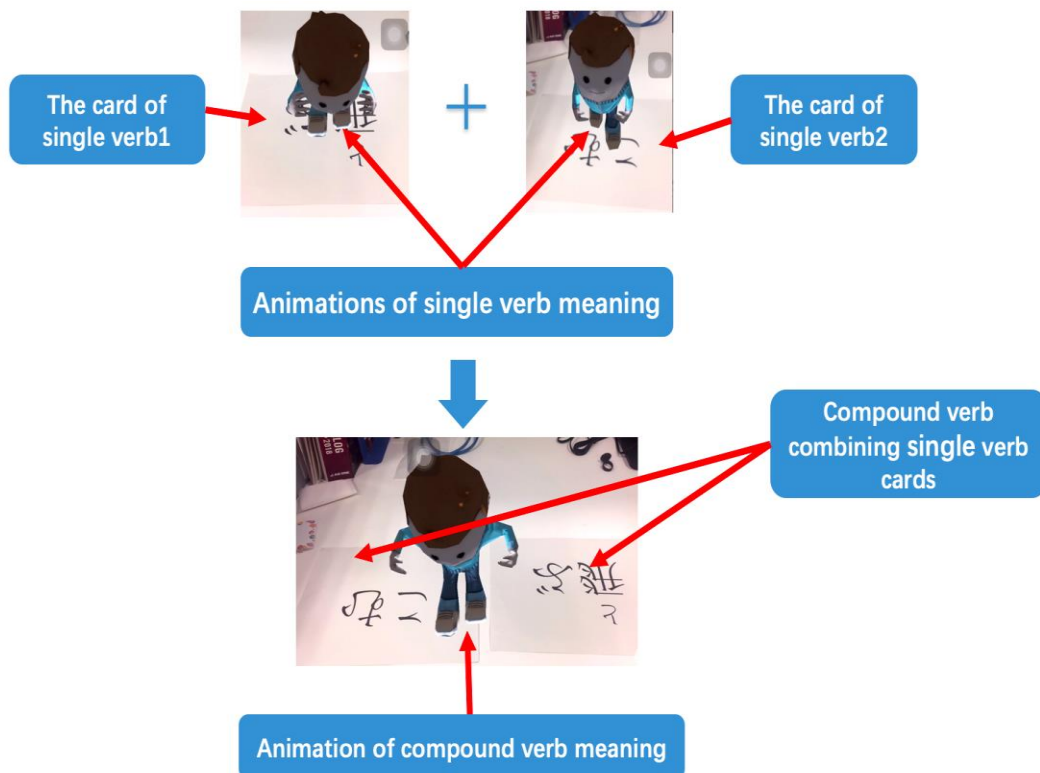


Figure 8. Screenshots of the verb combination

4.2 Outline of the System

In this study, the development language was C#. The animations of verbs were made with Maya based on the image schemas proposed by such studies as Matsuda (2004). Since iOS and Android are utilized as operating systems for smartphones commonly used in daily life, we employed a software development supporting multiplatform: Unity. Developed applications can be implemented on iOS and Android devices. Figure 9 is a configuration diagram of the system. It can be seen that the system recognizes the verb card using the Vuforia plugin. Figure 9 also shows that there are three main functions of the system, which will be described in detail below.

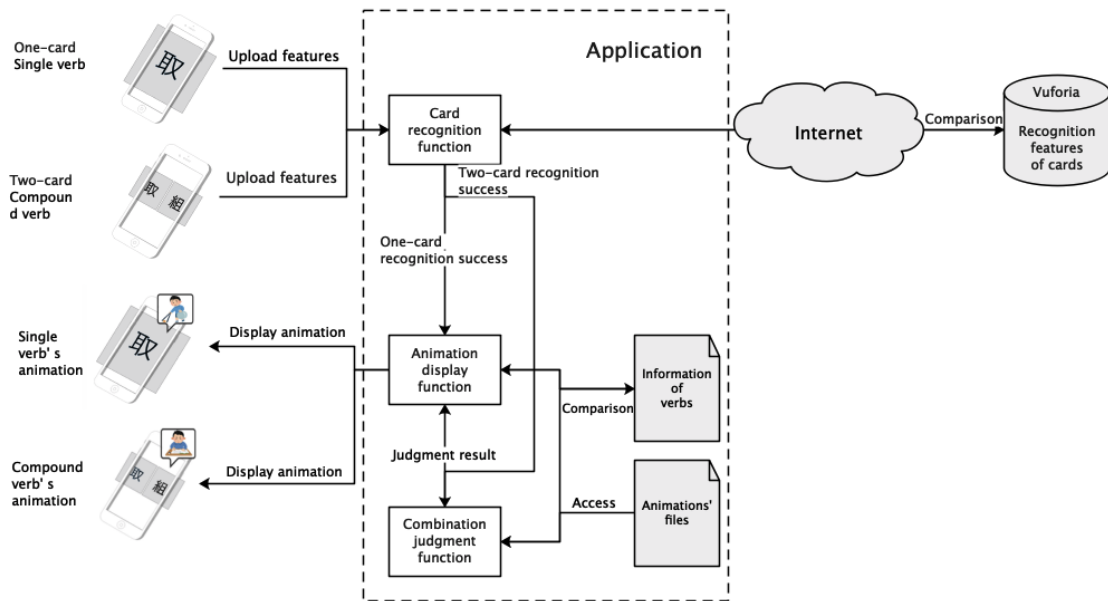


Figure 9. System configuration diagram

Figure 10 shows a flowchart of the system's function. It consists mainly of three functions.

4.2.1 Card recognition Function

The card recognition function recognizes the verb card on the camera screen and judges whether the verb card is present. The number of cards is also judged. The recognition features consist of Japanese verb characters and their readings. Recognition is made by comparing them with the recognition features uploaded to Vuforia.

4.2.2 Animation Display Function

When the card is recognized, the system is moved to the animation display function, in which the animation of the single verb or compound verb is displayed according to the number of cards. As long as the corresponding cards appear in the camera, the animation will continue to play. The animation and cards are not separated; rather, the animation displays on the cards. Moreover, the animation is not simply the animation of image schema. In the system, we use an avatar instead of a moving object and the motion trajectories of the avatar instead of arrows in the image schema (see Figure 5, Figure 6, and Figure 7). The results of one study (Sato, 2016) show that there is no significant difference between the learning effects of animated glosses and pictorial glosses in learning polysemous words. The authors believe that the learning effect of this function should be verified in future studies.

4.2.3 Combination Judgment Function

The combination judgment function judges the correctness of the combination of two cards. After the card recognition function recognizes two cards, this function will be triggered. The system will judge the order of the card combination and whether the combined verb exists. The existence of the compound verb is judged by the verb information file preset in the application.

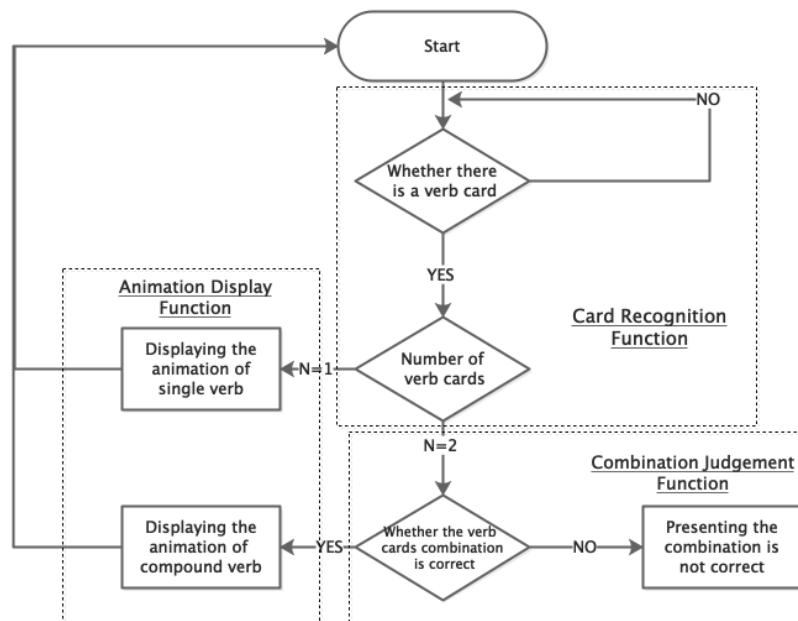


Figure 10. Flowchart of the functions

5. CONCLUSION AND FUTURE WORK

In the present study, we discussed the acquisition of Japanese compound verbs, image schemas for compound verbs, and the application of AR in learning, and explained the design and development of the system.

In future work, we will clarify the effect on learning compound verbs of this system. Furthermore, we will examine the differences in the learning effect due to literal explanations, visual glosses of the AR animation, and pictorial glosses of image schemas.

ACKNOWLEDGEMENT

This research was supported in part by the Foundation for the Fusion of Science and Technology (FOST) and JSPS KAKENHI JP 17K18659, and JP18K18657.

REFERENCES

- Benjamin, J. W. 2012. A conceptual approach to the instruction of phrasal verbs. *The Modern Language Journal*, Vol. 96, No. 1, pp. 416-438.
- Bolinger, D., 1977. *Meaning and form*. Longman Press, London, U.K.
- Chen, X. 2007. Nihongo fukugoudousi no syutokujyoukyou to shidou hei no mondaiteiki (The acquisition of Japanese compound verbs and implications for teaching: An examination of “au” and “komu” at Xi’an Foreign Language University). *Forum of International Development Studies*, Vol. 35, No. 1, pp. 93-102.
- Chiu, J. L., DeJaegher, C. J., and Chao, J. 2015. The effects of augmented virtual science laboratories on middle school students’ understanding of gas properties. *Computers & Education*, Vol. 85, No. 1, pp. 59-73.
- Dede, C., 2011. Emerging technologies, ubiquitous learning, and educational transformation. *6th European Conference of Technology Enhanced Learning*, Palermo, Italy, pp. 1-8.

- Gibbs, R. W., 2006. *Embodiment and cognitive science*. Cambridge University Press, USA.
- Ibanez, M. B., Serio, A. D., Villaran, D., and Delgado-Kloos, C., 2016. The acceptance of learning Augmented Reality environments: A case study. *2016 IEEE 16th International Conference on Advanced Learning Technologies (ICALT)*, Austin, USA, pp. 307-311.
- Japan Foundation, 2017. *Survey report on Japanese-language education abroad*. Retrieved Jan 22, 2019 from https://www.jpff.go.jp/j/project/japanese/survey/result/dl/survey_2015/Report_all_e.pdf
- Maekawa, S., Nakatani, Y., and Yonezawa, T. 2015. Sutorisei wo fukasita AR kyarakuta insutarakusyon ni yoru hyoonmoji gakusyu sisutemu (Phonogram learning System on AR animal with narrative interaction). *The Institute of Electronics, Information and Communication Engineers Transactions*, Vol. J98-D, No. 1, pp. 71-82.
- Mashiro, Y., Yamanouti, T., Seki, Y., Hatyutori, A., and Hayami, H. 2011. Maka wo mochiita kakucyou genjitsukan no eitango gakusyusien sisutemu no sisaku (Makers for augmented reality development of English words learning system). *Multimedia, Distributed, Cooperative and Mobile Symposium 2011*, Vol. 1, No. 1, pp. 1209-1214
- Matsuda, F. 2000. Fukuguodoushi no imi rikai houryaku no jittai to syutoku konnanten (Acquisition of compound verbs in Japanese by advanced learners of Japanese: Learners' strategies of comprehending the meanings of compound verbs). *Japanese Language Education*, Vol. 20, No. 1, pp. 52-65.
- Matsuda, F. 2001. Koa zushiki wo mochiita fukugoudoushi kougou "Komu" no ninchiimironteki setsumei (A cognitive semantics account of polysemous -komu in compound verbs based on a core schema). *Japanese Language Teaching*, Vol. 111, No. 1, pp. 16-25.
- Matsuda, F. 2002. Nihongogakusyusya ni yoru fukugoudoushi "komu" no syutoku (Acquisition of compound verb "komu" by Japanese learners). *Japanese Language Education in the World*, Vol. 12, No. 1, pp. 43-62.
- Matsuda, F. 2004. *Nihongo fukugoudoushi no syutoku kenkyu* (Research of Japanese compound verb acquisition). HITSUJI Syobou Press, Tokyo, Japan.
- Sano, K. 2004. JFL Gakusyusya no fukugoudoushi no jittai to konnanten: Annke-to cyosahoukoku (The actual state and difficulty of acquiring compound verbs by JFL learners: A questionnaire survey report). *Language Culture and Japanese Language Education*, Vol. 29, No. 1, pp. 88-91.
- Sin, A. K., and Zaman, H. B., 2010. Live Solar System (LSS): Evaluation of an augmented reality book-based educational tool. *2010 International Symposium on Information Technology*, Kuala Lumpur, Malaysia, pp. 1-6.
- Specht, M., Ternier, S., and Greller, W., 2011. Mobile augmented reality for learning: A case study. *Research Center for Educational Technology*, Vol. 7, No. 1, pp. 117-127.
- Tagawa, T. and Yuizono, T., 2016. Eigokudoushi no gokan gakushu o shien suru taburetto tanmatsu kyouzai no kaihatsu to hyouka (Development and evaluation of image-based learning system to acquire English phrasal verbs on tablet terminals). *The Special Interest Group Technical Reports of IPSJ*, Vol. 26, No. 1, pp. 1-7.
- Sato, T. 2016. Applicability of technology-enhanced visual glosses for explicit L2 vocabulary learning: The enhancement of metaphoric competence through the learning of English polysemous words. *Ampersand*, Vol. 3, No. 1, pp. 43-51.
- Tanaka, S. and Matsumoto, Y., 1997. *Nieigo Hikaku Sensyo 6* (Comparison of Japanese and English, Selection 6). KENKYUSHA Press, Tokyo, Japan.
- Walczak, K., Wojciechowski, R., and Cellary, W., 2006. Dynamic interactive VR network services for education. *2006 Proceedings of ACM symposium on virtual reality software and technology (VRST 2006)*, Limassol, Cyprus, pp. 277-286.