International Journal of Innovative Computing, Information and Control
Volume **15**, Number **6**, December **2019**
ICIC International ©2019 ISSN 1349-4198
pp. 2135–2147

AUTOMATIC SELECTION AND ANALYSIS OF VERB AND ADJECTIVE SYNONYMS FROM JAPANESE SENTENCES USING MACHINE LEARNING

Masaki Murata, Kazuki Orikane and Ryota Akae

Department of Information and Electronics
Tottori University
4-101 Koyama-Minami, Tottori 680-8552, Japan
murata@ike.tottori-u.ac.jp

Received April 2019; revised August 2019

Abstract. *In this study, we investigate a method to select appropriate verb and adjective synonyms automatically for a given sentence from multiple options based on machine learning. Our experimental results demonstrate that the accuracy rates of the proposed machine learning method for selecting verb and adjective synonyms (i.e., 0.84 and 0.76, respectively) are higher than those of a baseline method that always outputs the most frequently used synonyms from the training dataset (i.e., 0.70 and 0.67, respectively). Furthermore, the machine learning performance classifies the verb and adjective synonym pairs based on whether either of the synonyms could be used in a given sentence. Additionally, we examined important machine learning features for the selected verb and adjective synonym pairs and revealed useful information for selecting accurate synonyms. The performance of our machine learning approach was then compared with that of human test subjects. Finally, we discussed the differences among nouns, adverbs, and verbs/adjectives during synonym selection.*
Keywords: Verb and adjective synonyms, Selection, Japanese, Machine learning

1. **Introduction.** Two or more words having similar definitions are considered synonyms; e.g., “see” and “watch” are synonyms. Previous studies have investigated synonyms by extracting them from corpora and dictionaries [1, 2, 3, 4, 5]. Several other recent studies have investigated the use of synonyms and word similarity for applications such as text clustering and data augmentation [6, 7, 8].

Murata et al. attempted to automatically select the appropriate notational variant for a given sentence from several alternatives using machine learning [9]. The notational variants are used in various methods for expressing a given word, such as “3”, “III”, and “three”, in English. Some examples of notational variants for the Japanese word “cherry” are 桜 (in Chinese characters) and サクラ (in Katakana characters). Murata et al. also investigated the manner in which the appropriate nominal and adverb synonyms for a given sentence are automatically selected from multiple candidates using machine learning [10, 11]. In English, “research” and “study” constitute a common nominal synonym set, while “very” and “a lot” constitute a common adverb synonym set. However, the aforementioned studies have focused only on nominal and adverb synonyms instead of verb and adjective synonyms.

This study presents a method to select appropriate verb and adjective synonyms automatically for a given sentence from various alternatives using machine learning. For example, given the candidate sets of verb and adjective synonyms (e.g., “write” and

DOI: 10.24507/ijicic.15.06.2135

"write in" and "dull" and "sluggish") for a particular sentence, we may select "write". This selection process can be referred to as the automatic synonym selection. Further, we believe that our study will ensure the selection of appropriate verb and adjective synonyms during sentence generation [12, 13]. Herein, we focus on the automatic selection of Japanese synonyms.

Although several studies related to automatic synonym extraction have been conducted using machine learning [1, 2, 3] and some studies have analyzed the synonyms based on similarity without using machine learning [14], very few studies [10, 11] have investigated the automatic selection of an appropriate synonym from candidates with very similar meanings using supervised machine learning.

Numerous researchers have studied word selection, including lexical substitution [15, 16, 17, 18]. These studies have investigated word sense disambiguation for word selection in machine translation with the objective of selecting an appropriate word from various candidates with different meanings. In contrast, we have considered a methodology for selecting an appropriate word from candidates with similar meanings. In addition, we classify the synonym pairs by considering whether they can be used interchangeably in a given sentence, which has not been widely endeavoured by previous studies, excepting [10, 11].

Our study shares similarities with Murata et al.'s studies [10, 11]. For example, our study and Murata et al.'s studies [10, 11] investigate the method by which appropriate synonyms are automatically selected for a given sentence from multiple candidates using machine learning. While Murata et al.'s studies [10, 11] handle nominal and adverb synonyms, our study focuses instead on verb and adjective synonyms and examines differences between parts of speech (POS) in synonym selection. Verbs and adjectives are as important POS as nouns and adverbs. Hence, addressing only nouns and adverbs, such as in [10, 11], is insufficient. To thoroughly examine the differences between POS in synonym selection, all the main POS (nouns, adverbs, verbs, and adjectives) must be studied. Hence, our current study aims to fill this important gap by considering verbs and adjectives. The forms of Japanese nouns and adverbs do not change, but the forms of Japanese verbs and adjectives can change. Addressing this nuance, we employ a particular method for handling verbs and adjectives, which is not explored in [10, 11].

In this study, we use verb and adjective synonyms obtained from the EDR dictionary because verbs and adjectives that constitute a synonym pair exhibit considerably similar meanings; our objective is to automatically select appropriate synonyms.

Because synonyms have similar meanings, it may seem like careful selection is unnecessary and that the same word can be used in all cases; however, some circumstances require the selection of the most appropriate alternative. For example, two Japanese words, *kakikomu* ("write") and *kakiireru* ("write in"), have the same meaning in the EDR dictionary. If a computer saves information on a CD or similar device, the process can be only described as *kakikomu*. However, if we write something using brushes or create a schedule, the process can be described as *kakiireru*. Thus, the selection between *kakikomu* and *kakiireru* is dependent on the sentence and necessitates the selection of the proper verb and adjective synonyms.

The primary objectives of this study are to 1) obtain high-performance automatic verb and adjective synonym selection using machine learning and 2) classify the verb and adjective synonym pairs according to whether they require proper verb and adjective synonym selection. When a verb and adjective synonym can be easily selected using machine learning, we can conclude that the synonym pair requires proper verb and adjective synonym selection. However, when this selection is difficult, we consider that either of the synonym pairs can be applicable, and hence, the pair does not require proper selection. The results of this study will contribute to the proper selection and usage of verb and adjective synonyms.

2. Task and Proposed Method.

2.1. **Task.** We are provided with two verb synonyms (or two adjective synonyms), A and B, with similar meanings; we refer to these synonyms as the target words in this study. We are also provided with a dataset of sentences that initially contained one of the target words but from which the target words have been subsequently removed. The task of this study is to process the sentences and identify the missing words. Furthermore, the system's selection is considered accurate if it selects the word originally present in the sentence.

2.2. **Proposed method.** In this study, we estimate the target word that was originally present in a given sentence from the two target words using on machine learning trained with sentences that contain either of the target words. The machine learning process assesses the target word originally present in the sentence to denote the sentence's category or class. Supervised learning was used by applying the maximum entropy method [19, 20], which is commonly used in machine learning because it performs as well as a support vector machine while outputting the importance of each feature.

One machine learning cycle was used for each pair of target words (verb and adjective synonyms); therefore, n machine learning cycles were used for n pairs of target words. In this study, we selected the most frequently used pairs of target words. Furthermore, the features of the sentences containing either of the two target words were extracted, and our machine learning method was used for selecting the most appropriate target word from each sentence.

After studying Murata et al.'s papers [9, 11], we decided to use the features of machine learning presented in Table 1. These features are described in more detail below.

The category numbers in the table are the ten-digit numbers used in a Bunrui-goi-hyou, or Japanese thesaurus [21, 22]. Words with similar meanings are represented by similar ten-digit numbers. Here, we use both the first five and first three digits of the ten-digit numbers as features for identifying the concepts embodied by each word.

Features F1 to F12 represent the information in the bunsetsu (phrase) containing the target word because the words adjacent to the target word can assist synonym selection. Meanwhile, features F13 to F48 represent useful information about the time at which the words in a bunsetsu modify or are modified by the bunsetsu containing the target word, as syntactic information about the sentence can also assist synonym selection.

While Murata et al. [9] used 62 features, we used 48 features; however, in general the features are very similar. Both feature sets use words, POS, and category numbers for a target bunsetsu, as well as a bunsetsu modified by a target bunsetsu, and a bunsetsu modifying a target bunsetsu. Furthermore, the features of Murata and Nakase [11] are the same as those we employed in our study.

We further classified the synonyms into groups based on whether the machine learning process was able to automatically select the accurate word, classifying them into "high", "medium", and "low" categories based on their recall rates. The recall rate is conceptually similar to accuracy and, in this case, represents the ratio of the number of times that the system accurately outputs a given target word to the total number of sentences in which that word was originally present.

When the lowest recall rate for either of the target words in a pair becomes at least 0.8, the words are placed in the "high" category because high recall rates for both words

Table 1. Features used in machine learning

ID	Feature explanation
F1	Nouns in the sentence
F2	Three words just before and after the target word
F3	First 1, 2, 3, 4, 5, and 7 digits of F2's category number
F4	Functional words in the bunsetsu containing the target word
F5	F4's parts of speech (POS)
F6	First 1, 2, 3, 4, 5, and 7 digits of F4's category number
F7	First functional word in the bunsetsu containing the target word
F8	F7's POS
F9	First 1, 2, 3, 4, 5, and 7 digits of F7's category number
F10	Last functional word in the bunsetsu containing the target word
F11	F10's POS
F12	First 1, 2, 3, 4, 5, and 7 digits of F10's category number
F13	Content words in the bunsetsu that modify the bunsetsu containing the target word
F14	F13's POS
F15	First 1, 2, 3, 4, 5, and 7 digits of F13's category number
F16	Functional words in the bunsetsu that modify the bunsetsu containing the target word
F17	F16's POS
F18	First 1, 2, 3, 4, 5, and 7 digits of F16's category number
F19	First content word in the bunsetsu that modifies the bunsetsu containing the target word
F20	F19's POS
F21	First 1, 2, 3, 4, 5, and 7 digits of F19's category number
F22	Last content word in the bunsetsu that modifies the bunsetsu containing the target word
F23	F22's POS
F24	First 1, 2, 3, 4, 5, and 7 digits of F22's category number
F25	First functional word in the bunsetsu that modifies the bunsetsu containing the target word
F26	F25's POS
F27	First 1, 2, 3, 4, 5, and 7 digits of F25's category number
F28	Last functional word in the bunsetsu that modifies the bunsetsu containing the target word
F29	F28's POS
F30	First 1, 2, 3, 4, 5, and 7 digits of F28's category number
F31	Content words in the bunsetsu that are modified by the bunsetsu containing the target word
F32	F31's POS
F33	First 1, 2, 3, 4, 5, and 7 digits of F31's category number
F34	Functional words in the bunsetsu that are modified by the bunsetsu containing the target word
F35	F34's POS
F36	First 1, 2, 3, 4, 5, and 7 digits of F34's category number
F37	First content word in the bunsetsu that is modified by the bunsetsu containing the target word
F38	F37's POS
F39	First 1, 2, 3, 4, 5, and 7 digits of F37's category number
F40	Last content word in the bunsetsu that is modified by the bunsetsu containing the target word
F41	F40's POS
F42	First 1, 2, 3, 4, 5, and 7 digits of F40's category number
F43	First functional word in the bunsetsu that is modified by the bunsetsu containing the target word
F44	F43's POS
F45	First 1, 2, 3, 4, 5, and 7 digits of F43's category number
F46	Last functional word in the bunsetsu that is modified by the bunsetsu containing the target word
F47	F46's POS
F48	First 1, 2, 3, 4, 5, and 7 digits of F46's category number

in a pair indicate accurate estimation. Meanwhile, the words are considered "medium" when the lowest recall rate is between 0.5 and 0.8, and when the rate is less than 0.5 the words are placed in the "low" category, indicating inaccurate estimation.

3. **Datasets Used in the Experiments.** Our experimental datasets involve both verb and adjective synonyms that have been observed in newspapers. Particularly, we used articles published in the Mainichi newspaper and extracted the verb and adjective pairs that satisfied all the following conditions.

1) Both the words in the pair have the same EDR concept identification number. Therefore, the EDR dictionary considers both words to have the same meaning.
2) Only the basic forms of both words are handled. In Japanese, the verb and adjective forms can change according to the usage scenario. For example, *ugoku* (move) becomes *ugoi ta* (moved) in the past tense. *ugoku* is the basic form; hence, we only consider *ugoku*.
3) Both words appear at least 50 times in Mainichi newspaper articles (between 1991 and 1995 and between 2011 and 2015; ten years in total) regarding verb synonym pairs. Both words appear at least 20 times in Mainichi newspaper articles (between 1991 and 1995 and between 2011 and 2015; ten years in total) regarding adjective synonym pairs.
4) The representative words given by the JUMAN morphological analyzer are different for each word, i.e., the two words are actually different and are not simply notational variants.

Condition 1 was used to extract words with considerably similar meanings.

Condition 2 was used for the following reasons: the verb forms varied in each verb; the change in verb also affects the change in surrounding words; the original verb can sometimes be easily estimated based on the change of surrounding words. This problem was solved by ensuring the usage of only the basic verb form. The same discussion is valid for adjectives as well.

Condition 3 was used to select the words most frequently used in newspaper articles.

Condition 4 was used to eliminate the notational variants because the selection of one notational variant from several notational variants has been previously addressed by another study [9].

Condition 2 was not used in the previously performed automatic selection of noun and adverb synonyms [10, 11] because Japanese nouns and adverbs cannot change.

Along with Conditions 1, 3, and 4, the condition that both the words were assigned one conceptual identifier in the Japanese word dictionary was used in the previously performed automatic selection of noun and adverb synonyms [10, 11]. However, the majority of the verbs and adjectives in the synonym pairs in our study are polysemous. If the condition that both the words should have one conceptual identifier is satisfied, considerably fewer synonym pairs are used in the experiments. Therefore, this condition was eliminated.

In our dataset, 17 verb pairs and 9 adjective pairs satisfied all four conditions, and 40–3000 sentences were considered for each pair.

In Condition 3, we used 50 times for verbs and 20 times for adjectives, where times correspond to the number of data items. For machine learning, more times (i.e., more data items) are better. However, when we use more times, the number of obtained synonym pairs decreases. We determined that 50 times for verbs and 20 times for adjectives result in about ten synonym pairs, which is enough to investigate the synonym selection of verbs and adjectives.

4. Synonym Selection Experiments.

4.1. **Methods.** We applied our machine learning approach to the 17 verb pairs and 9 adjective pairs obtained using the aforementioned process and considering the two most frequently used synonyms in each case. Furthermore, we used our machine learning method for selecting the most appropriate synonym from each sentence.

A dataset was constructed for each pair of words by extracting sentences from the Mainichi newspaper that contained either of the words. Occasionally, the same sentence was repeated more than once in the newspaper; therefore, we eliminated redundant sentences to ensure that each sentence was considered exactly once. We then conducted ten-fold cross validation for each word pair. This process involved the division of the dataset into ten parts, considering one part as the test dataset and the remaining nine parts as the training dataset. After using the training dataset to learn the maximum entropy method, the category (class) of each item in the test dataset was estimated. The estimates were further compared with the correct categories. This process was repeated ten times by selecting a different part of the dataset for testing each time, which ensured that all the ten parts were evaluated.

We compared the proposed method with the baseline method that always outputs the most frequently used verb and adjective synonyms from the training dataset.

4.2. **Results.** The results obtained using machine learning to classify the 17 verb pairs into the three recall rate categories discussed in Section 2.2 ("high", "medium", and "low") are presented in Table 2. Meanwhile, the average accuracy rates of the proposed and baseline methods for each of the recall rate categories are presented in Table 3. The accuracies of the two methods are compared in Table 4. Here, "even" represents the number of synonym pairs for which the accuracy difference was 0.01 or less, while "proposed method (win)" denotes the number of pairs for which the accuracy rate of the proposed method was higher than that of the baseline method, and "baseline method

Table 2. Fraction of the verb synonym pairs classified into each of the three recall rate categories

Category	Ratio
High	0.41 (7/17)
Medium	0.18 (3/17)
Low	0.41 (7/17)

Table 3. Average accuracy rates of the proposed and baseline methods with respect to verb synonym pairs

	High	Medium	Low	All
Proposed method	0.95	0.75	0.77	0.84
Baseline method	0.65	0.62	0.78	0.70

Table 4. Comparison of accuracies for the proposed and baseline methods with respect to verb synonym pairs

	High	Medium	Low
Proposed method (win)	7	3	2
Baseline method (win)	0	0	3
Even	0	0	2

(win)" is the number of pairs for which the accuracy rate of the baseline method was higher.

The results obtained using machine learning to classify the 9 adjective pairs into the three recall rate categories discussed in Section 2.2 ("high", "medium", and "low") are presented in Table 5. The average accuracy rates for the proposed and baseline methods for each of the recall rate categories are presented in Table 6. Meanwhile, the accuracies of both the methods are compared in Table 7.

Table 5. Fraction of the adjective synonym pairs classified into each of the three recall rate categories

Category	Ratio
High	0.11 (1/9)
Medium	0.33 (3/9)
Low	0.56 (5/9)

Table 6. Average accuracy rates of the proposed and baseline methods with respect to adjective synonym pairs

	High	Medium	Low	All
Proposed method	0.89	0.82	0.66	0.76
Baseline method	0.54	0.69	0.68	0.67

Table 7. Comparison of accuracies for the proposed and baseline methods with respect to adjective synonym pairs

	High	Medium	Low
Proposed method (win)	1	4	1
Baseline method (win)	0	0	2
Even	0	0	1

5. **Discussion.**

5.1. **Comparison of the proposed and baseline methods.**

Tables 3 and 6 denote the accuracy rates of the verb and adjective synonyms for the proposed method, which are 0.84 and 0.76, respectively, while those of the baseline methods are 0.70 and 0.67, respectively, indicating that the proposed method was significantly more accurate than the baseline method. Similarly, Tables 4 and 7 demonstrate that the proposed method was more accurate than the baseline method for all the synonym pairs belonging to either the "high" or "medium" recall rate categories. These results indicate that the proposed method and machine learning features considered are useful for performing verb and adjective synonym selection.

5.2. **Trends of verb synonyms based on recall rates.**

We manually examined the classified verb pairs. The word pairs in the "high" category, such as *kakikomu* ("write") and *kakiireru* ("write in" or "enter"), tended to require proper synonym selection, similar to the process expressed in the following example sentences.

Example sentence 1a (synonym correctly selected by machine learning)

pasokon ga	*data wo*	*cdrom ni*	*kakikomu.*	
(personal computer)	(data)	(CD)	(write)	

(The personal computer writes the data onto the CD.)

Example sentence 1b (synonym correctly selected by machine learning)
kare ga fude de shimei wo kakiireru.
(he) (brush) (name) (write in or enter)
(He enters names with the brush.)

The word pair comprising *kakikomu* ("write") and *kakiireru* ("write in" or "enter") was classified as "high". Both words have the same EDR concept identification number, "3d06fc". In this category, machine learning achieved good recall rates (good performance), and word selection was considered easy and necessary. The machine learning results for this pair are presented in Table 8. The precision rate represents the ratio of the number of times that the system accurately outputs a given target word to the total number of sentences in which the word was judged by our method to be the word written in the corresponding line of the table. Meanwhile, the term "sentences" indicates the total number of sentences in which the word written in the corresponding line of the table was originally present. The three important features (i.e., features with high α values) are presented in Table 9. The normalized α value represents the importance associated with a corresponding feature for selecting a synonym, as learned using the maximum entropy method; further details can be observed in a previous study [20]. The "Default feature" for a category means that the category is judged to be the output category if no other feature exists. The category with "Default feature" is likely to be the output category. While *kakikomu* exhibits the "Default feature", *kakikomu* exhibits a large total number as presented in Table 8. It can be asserted that *kakikomu* is a relatively general word when compared to *kakiireru*, and *kakikomu* can be used in majority of the cases. As presented in the correct answer example 1, when a computer records data using a CD, it can only be described as *kakikomu*. However, when we write something using brushes or generate schedules, it can be described as *kakiireru*. Based on this information, it can be stated that the synonym pair requires proper verb synonym selection.

Table 8. Machine learning results for *kakikomu* and *kakiireru* ("high" category)

	Recall rate	Precision rate	Sentences
kakikomu	0.96	0.95	595
kakiireru	0.84	0.89	199

Table 9. Important machine learning features for selecting between *kakikomu* and *kakiireru* ("high" category)

kakikomu	
Feature	Normalized α value
Default feature	0.83
F1: *wo* (object case)	0.60
F1: *ran* (form)	0.52

kakiireru	
Feature	Normalized α value
F1: *fude* (brush)	0.66
F1: *yotei* (schedule)	0.65
F31: The head of the target word is a noun.	0.57

TABLE 10. Machine learning results for *misadameru* and *tsukitomeru* ("medium" category)

	Recall rate	Precision rate	Sentences
misadameru	0.56	0.63	103
tsukitomeru	0.74	0.68	131

TABLE 11. Important machine learning features for selecting between *misadameru* and *tsukitomeru* ("medium" category)

misadameru	
Feature	Normalized α value
F1: *taisetsu* (importance)	0.59
F1: *kongo* (from now on)	0.54
F1: *shourai* (future)	0.53

tsukitomeru	
Feature	Normalized α value
F1: *chosa* (research)	0.70
F1: *kenkyu* (investigation)	0.62
F1: *geiin* (reason)	0.61

Next, we discuss the word pairs classified as "medium", including *misadameru* ("determine" or "see") and *tsukitomeru* ("identify"). Again, both the words have the same EDR concept identification number, "3c4b79". The machine learning results for this pair are presented in Table 10, and the three important features (i.e., features with high α values) are presented in Table 11. With respect to "importance" and "future", as observed in Table 11, *misadameru* is used for clarifying a future situation or a situation related to importance. Regarding "research" and "reason", as observed in Table 11, *tsukitomeru* is used for clarifying the current situation. Therefore, this synonym pair tends to require proper synonym selection.

Example sentence 2a (synonym correctly selected by machine learning)
nihon wa soren gawa no shini wo misadameru.
(Japan) (Soviet Union) (real intention) (determine or see)
(Japan determines the real intention of the Soviet Union.)

Example sentence 2b (synonym correctly selected by machine learning)
sisutemu ga hikouki no ichi wo seikakuni tsukitomeru.
(system) (airplane) (location) (accurately) (identify)
(The system identifies the location of the airplane accurately.)

The majority of word pairs classified as "low", such as *mikagiru* and *misuteru* (both mean "abandon"), do not tend to require proper synonym selection. In such cases, both words exhibit very similar meanings and either could be used in almost all the sentences.

5.3. **Trends of adjective synonyms revealed by recall rates.** We manually examined the classified adjective pairs. The word pairs classified as "high", such as *chikashii* ("close" or "near") and *mutsumajii* ("close" or "friendly"), tended to require proper synonym selection, similar to the examples presented in the following sentences.

Example sentence 3a (synonym correctly selected by machine learning)
Kare wa kare jishin to chikashii mono wo kanjita.
(he) (his own world) (close) (thing) (felt)
(He felt something close with his own world.)

Example sentence 3b (synonym correctly selected by machine learning)

watashi no	*inshou ni*	*nokotta*	*mono wa*	*karera no*	*naka*	*mutsumajii*	*sugata da.*
(I)	(impression)	(remained)	(thing)	(their)	(relationship)	(friendly)	(appearance)

(What I remained in the impression is their friendly appearance.)

The pair comprising *chikashii* ("close" or "near") and *mutsumajii* ("close" or "friendly") was classified as "high". Both the words have the same EDR concept identification number, "3cfc19". In this category, machine learning achieved good recall rates (good performance), and word selection was both easy and necessary. The machine learning results for this pair are presented in Table 12, and the three important features (i.e., features with high α values) are presented in Table 13. Based on Table 12, when the head of the target word is a noun, *chikashii* is more often used compared to *mutsumajii*. As denoted in the correct answer example 3b, *mutsumajii* is often used like *naka mutsumajii*. When *sugata* (appearance) is observed, *mutsumajii* is often used. Based on this information, this synonym pair tends to require proper synonym selection.

Table 12. Machine learning results for *chikashii* and *mutsumajii* ("high" category)

	Recall rate	Precision rate	Sentences
chikashii	0.90	0.90	53
mutsumajii	0.88	0.88	44

Table 13. Important machine learning features for selecting between *chikashii* and *mutsumajii* ("high" category)

chikashii	
Feature	Normalized α value
F31: The head of the target word is a noun.	0.83
F1: *hito* (human)	0.54
F1: *kankei* (relation)	0.52

mutsumajii	
Feature	Normalized α value
F1: *naka* (relastionship)	0.65
F1: *you* (appearance)	0.56
F1: *sugata* (appearance)	0.55

Table 14. The results of machine learning for *darui* and *kedarui* ("medium" category)

	Recall rate	Precision rate	Sentences
darui	0.96	0.91	77
kedarui	0.75	0.97	28

Next, we discuss the word pairs classified as "medium", including *darui* ("dull") and *kedarui* ("sluggish"). Again, both the words have the same EDR concept identification number, "3cea9b". The results of machine learning for this pair are presented in Table 14, and the three important features (i.e., features with high α values) are presented in Table 15. With respect to the word "body" in Table 15, we can observe that *darui* is used to express the fatigue of the body. Considering "mood" in Table 15, we can observe that *kedarui* is used when expressing tiredness. As in the correct answer 4b, *kedarui* is

TABLE 15. Important features of machine learning for selecting between *darui* and *kedarui* ("medium" category)

darui	
Feature	Normalized α value
F31: The head of the target word is a verb.	0.55
F1: *karada* (body)	0.55
F1: *shojo* (symptom)	0.53

kedarui	
Feature	Normalized α value
F31: The head of the target word is a noun.	0.60
F1: *you* (appearance)	0.53
F1: *mudo* (mood)	0.52

typically used when the head of the target word is a noun. Based on this information, it is apparent that this synonym pair requires proper synonym selection. Furthermore, *darui* and *kedarui* exhibit an interesting characteristic related to the POS. The head of *darui* is likely to be a verb, whereas the head of *kedarui* is likely to be a noun.

Example sentence 4a (synonym correctly selected by machine learning)
watashi no karada wa darui noni atama wa saeru.
(my) (body) (dull) (head) (clear)
(My body is dull, but my head is clear.)

Example sentence 4b (synonym correctly selected by machine learning)
kedarui benchi no kuuki ga ippenshita.
(sluggish) (bench) (air) (has changed completely)
(The air of the sluggish bench has changed completely.)

The majority of word pairs classified as "low", such as *migurushii* and *mittomonai* (both meaning "degrading"), do not tend to require proper synonym selection. In this case, the two words have considerably similar meanings and either could be used in almost all the instances.

5.4. **Comparison with human verb and adjective selection performance.** We investigated human performance for each of the three recall categories by randomly selecting two synonym pairs from each of the "low", "medium", and "high" categories[1] for a total of six pairs. For the words in each pair (A and B), we randomly extracted five sentences containing A and five sentences containing B from the Mainichi newspaper articles. These ten sentences were provided to three human subjects, who were asked to judge which word (A or B) is most suitable for each sentence. The accuracy rates were calculated for each category using 60 data points (all possible combinations of three subjects, two synonym pairs, two words, and five sentences). These experiments were conducted for selecting both verbs and adjectives, the results of which are respectively presented in Tables 16 and 17.

These results demonstrate that the accuracy rates were high for the "high" and "medium" categories, while the accuracy rates were low for the "low" category. This comparison suggests that our machine learning method can approximately estimate the difficulty of synonym selection faced by human subjects.

[1]In case of the "high" category for adjectives, we used only one synonym pair. This is because we obtained only one case for the "high" category. Therefore, 30 data points in total were used.

Table 16. Accuracy rates of verb synonyms during human selection

Recall rate category	High	Medium	Low
Accuracy	0.75 (45/60)	0.83 (50/60)	0.51 (31/60)

Table 17. Accuracy rates of adjective synonyms during human selection

Recall rate category	High	Medium	Low
Accuracy	0.93 (28/30)	0.91 (55/60)	0.63 (38/60)

6. **Difference among the Selections of Nouns, Adverbs, and Adjectives/Verb Synonyms.** In this section, we discuss the differences among the selections of nouns, adverbs, and adjectives/verb synonyms. This discussion is useful for analyses investigating differences between POS in synonym selection.

While selecting the noun synonyms [10], words were considered to be the important features.

While selecting the adverb synonyms [11], words were the important features in various cases; however, in some cases, parts of speech were considered to be the important features. The feature that the head of a target word is an adjective as well as the feature that the head of a target word is a verb were both important at times.

While selecting verb/adjective synonyms, words were the important features in many cases; however, in some cases, parts of speech were the important features. The features that the head of a target is a verb and that the head of a target is a noun were sometimes important. The feature that the head of a target word was the noun was initially discovered in the current study.

7. **Conclusion.** In this study, we have investigated a method for automatically selecting appropriate verb and adjective synonyms for a given sentence from multiple alternatives using machine learning. Our experiments confirmed that the accuracy rates of the proposed machine learning method for verb and adjective synonyms (0.84 and 0.76, respectively) were higher than those of the baseline method that always outputs the most frequent synonym from the training data set (0.70 and 0.67, respectively). Based on the machine learning performance, we classified the synonym pairs in terms of whether they require proper selection for correct usage. For example, the pair comprising *kakikomu* ("write") and *kakiireru* ("write in" or "enter") was identified as needing proper selection. We further examined significant features for some synonym pairs to reveal what information is useful for performing verb and adjective synonym selection. Furthermore, we compared the performance of our machine learning approach with that of human test subjects. Through our experiments, we confirmed that our machine learning method can approximately estimate the difficulty of synonym selection faced by human subjects.

Finally, we discussed the differences among nouns, adverbs, and verbs/adjectives during synonym selection. The following findings were obtained. In noun synonyms, words were important features. In adverb synonyms and verb/adjective synonyms, POS were also important features. In particular, for verb/adjective synonyms, it was useful when the head of a target word was the noun, which is a feature initially discovered in the current study.

Noun synonyms and adverb synonyms have been explored in previous studies, while verb/adjective synonyms are addressed in the current study. In future studies, we will investigate synonyms in other POS, including onomatopoeia such as *zaa zaa* and *para para*, which are Japanese onomatopoeia that describe the state of rain.

REFERENCES

[1] D. Lin, Automatic retrieval and clustering of similar words, *Proc. of COLING-ACL*, pp.768-774, 1998.

[2] R. Barzilay and K. R. McKeown, Extracting paraphrases from a parallel corpusy, *Proc. of the 39th Annual Meeting on Association for Computational Linguistics (ACL 2001)*, pp.50-57, 2001.

[3] T. Wang and G. Hirst, Exploring patterns in dictionary definitions for synonym extraction, *Natural Language Engineering*, vol.18, no.3, pp.313-342, 2012.

[4] F. Hu, Z. Shao and T. Ruan, Self-supervised synonym extraction from the web, *Journal of Information Science and Engineering*, vol.31, no.3, pp.1133-1148, 2015.

[5] X. Ma, X. Luo, S. Huang and Y. Guo, Multi-distribution characteristics based Chinese entity synonym extraction from the web, *International Journal of Intelligent Information Technologies*, vol.15, no.3, pp.42-63, 2019.

[6] M. Wang, L. Wang, N. Tian and B. Li, Automatic interpretation of Chinese noun compounds based on word similarity, *ICIC Express Letters, Part B: Applications*, vol.7, no.6, pp.1215-1221, 2016.

[7] F. S. Al-Anzi, D. AbuZeina and S. Hasan, Utilizing standard deviation in text classification weighting schemes, *International Journal of Innovative Computing, Information and Control*, vol.13, no.4, pp.1385-1398, 2017.

[8] X. Wang, Y. Sheng, H. Deng and Z. Zhao, CHARCNN-SVM for Chinese text datasets sentiment classification with data augmentation, *International Journal of Innovative Computing, Information and Control*, vol.15, no.1, pp.227-246, 2019.

[9] M. Murata, M. Kojima, T. Minamiguchi and Y. Watanabe, Automatic selection and analysis of Japanese notational variants on the basis of machine learning, *International Journal of Innovative Computing, Information and Control*, vol.9, no.10, pp.4231-4246, 2013.

[10] M. Murata, Y. Goda and M. Tokuhisan, Automatic selection and analysis of synonyms in Japanese sentences using machine learning, *The 2014 International Conference on Artificial Intelligence*, pp.1-4, 2014.

[11] M. Murata and M. Nakase, Automatic selection and analysis of adverb synonyms in Japanese sentences using machine learning, *International Journal of Advanced Intelligence*, vol.9, no.3, pp.343-353, 2017.

[12] K. Knight and V. Hatzivassiloglou, Two-level, many-paths generation, *Proce. of the 33rd Annual Meeting on Association for Compu-Tational Linguistics (ACL 1995)*, pp.252-260, 1995.

[13] K. Uchimoto, S. Sekine and H. Isahara, Text generation from keyword, *COLING 2002*, pp.1037-1043, 2002.

[14] P. Edmonds and G. Hirst, Near-synonymy and lexical choice, *Computational Linguistics*, vol.28, no.2, pp.105-144, 2002.

[15] K. Y. Kim, S. Y. Park and S. J. Lee, Automatic sense disambiguation for target word selection, *Proc. of the 10th Pacific Asia Conference on Language, Information and Computation*, pp.109-114, 1995.

[16] K. Uchimoto, S. Sekine, M. Murata and H. Isahara, Word translation based on machine learning models using translation memory and corpora, *Senseval 2*, pp.155-158, 2001.

[17] W. Aziz and L. Specia, USP*wlv* and WLV*usp*: Combining dictionaries and contextual information for cross-lingual lexical substitution, *Proc. of the 5th International Workshop on Semantic Evaluation, ACL 2010*, pp.117-122, 2010.

[18] J. Matsuno and T. Ishida, Constraint optimization approach to context basedword selection, *Proc. of the 22nd International Joint Conference on Artificial Intelligence*, pp.1846-1852, 2011.

[19] A. L. Berger, S. A. D. Pietra and V. J. D. Pietra, A maximum entropy approach to natural language processing, *Computational Linguistics*, vol.22, no.1, pp.39-71, 1996.

[20] M. Murata, K. Uchimoto, M. Utiyama, Q. Ma, R. Nishimura, Y. Watanabe, K. Doi and K. Torisawa, Using the maximum entropy method for natural language processing: Category estimation, feature extraction, and error correction, *Cognitive Computation*, vol.2, no.4, pp.272-279, 2010.

[21] M. Murata, K. Kanzaki, K. Uchimoto, Q. Ma and H. Isahara, Meaning sort – Three examples: Dictionary construction, tagged corpus construction, and information presentation system, *The 2nd International Conference on Computational Linguistics and Intelligent Text Processing, CICLing 2001*, Mexico City, pp.305-318, 2001.

[22] NLRI, *Bunrui Goi Hyou*, Shuuei Publishing, 1964.