

# Predicting Japanese Primary Schoolchildren's English Vocabulary Knowledge from a Sound-Meaning Recognition Test

Kohei Kanayama  
*Hokkaido University of Education*

## Abstract

The purposes of this study were to: (1) build an appropriate model for predicting primary schoolchildren's English vocabulary knowledge; (2) examine whether the developed model applies to new data; and (3) discuss how to apply the model to L2 vocabulary instruction. More specifically, the study asked third- and fourth-grade public primary school students to take a sound-meaning recognition test. All the study's target words were *katakana* English and on the first 1,000-level of word frequency according to JACET (Japan Association of College English Teachers) 8000 and SVL (Standard Vocabulary List) 12000; however, different numbers of word phonemes were included. The collected data were divided into training and testing data. To build the predictive model from the training data, the study used a generalized linear mixed model, which revealed that: (1) third-graders' scores were as high as those of fourth-graders, and (2) the more phonemes a word has, the lower the test score. From the model, the study developed a regression formula that made it possible to predict how many primary schoolchildren knew the meaning of a certain word when they listened to its pronunciation. It also found that the predicted percentages from the training data correlated moderately with the actual test scores from the testing data. Therefore, the model achieved moderately high quality for predicting new data.

**Keywords:** Predictive model, L2 vocabulary knowledge, Machine learning

## 1 Introduction

In April 2020, English education became a formal subject in Japanese primary schools. Fifth- and sixth-grade children learn English as a formal subject, and third- and fourth-grade children are introduced to it through a compulsory but not formal subject called *gaikokugo katsudo* [foreign language activity]. According to the *Course of Study* (MEXT, 2017), the Japanese primary school English curriculum, teachers in primary schools introduce 600–700 words over 4 years (third to sixth grade) to facilitate English communication. They also teach collocations and high-frequency phrases or idioms such as “get up,” “look at,” “excuse me,” and “I’m sorry.” Despite the large number of English words to be learned, children study English for only 157.5 hours over 4 years. Third- and fourth-graders learn English only once a week, while fifth- and sixth-graders learn

twice a week. A lesson lasts 45 minutes, and the academic year has 35 weeks; therefore, the 4 years of English study unfold over 210 lessons for 157.5 hours. In other words, teachers have very limited instructional time compared to the number of English words to be taught. Thus, for effective English vocabulary instruction within a limited amount of time, distinguishing between words that children know and do not know is important for teachers to concentrate on unfamiliar words. In addition, teachers should spend time turning the words already known by many students into productive words to be used in actual English communication.

However, how should teachers distinguish which words children are familiar with? One of the purposes of this study is to develop a suitable model to predict Japanese primary schoolchildren's English vocabulary knowledge. For example, how many children recognise the meaning of "apple" when they listen to its English sound (/æpl/)? The features of the word are key to this question. Previous studies have found that the shorter a word is, the easier it is to remember (Willis & Ohashi, 2012); therefore, it is predictable that "apple" is easier to remember than "strawberry." Moreover, previous studies have revealed that Japanese learners can recognise English words that exist in Japanese as *katakana* English better than those that do not (Kasahara et al., 2012; Willis & Ohashi, 2012; Yoshimura, 2009). Therefore, many primary schoolchildren know "apple," which has the feature of *katakana* English [アップル], but do not know "pear," which does not. The next section introduces some previous studies that conducted English vocabulary knowledge tests on Japanese learners of English and investigated word features that affect the ease of vocabulary learning.

## 2 Literature Review

### 2.1 Effects of Katakana English on the Ease of Remembering for Japanese Learners of English

Some studies have conducted English vocabulary knowledge tests to examine what words are easy for Japanese learners of English to learn and remember. Yoshimura (2009) conducted a sound-meaning recognition test for first- to sixth-grade students. The test asked the participants to choose an appropriate answer from four options (e.g., *chairo* [brown], *midori* [green], *murasaki* [purple], *shiranai* [I do not know]) when they heard the target word (e.g., /pɜːrpl/). Findings revealed that even among first- and second-grade children, some *katakana* English words (e.g., "house," "fruit," "milk," "eleven," "purple") had high correct rates (over 90%).

The fact that words with the features of *katakana* English are helpful for Japanese learners of English to learn is also supported by other studies (Kasahara et al., 2012; Willis & Ohashi, 2012). Willis and Ohashi (2012) examined what English words are easy for Japanese university students to learn and retain in the long term, that is, words with the features of *katakana* English, high-frequency words, and words with fewer phonemes. These researchers emphasised that *katakana* English was the most influential.

The Japanese use many *katakana* English words (e.g., ピアノ (*piano*) [piano], パンダ (*panda*) [panda], オレンジ (*orenji*) [orange]). In fact, Daulton (2008) reported that about half of the 3,000 high-frequency English words are often used as *katakana* English, spread widely across newspapers, magazines, and television shows (Shibuya, 2012). Moreover, Hoshino and Shimizu (2019) analysed the textbooks *Let's Try!* (MEXT, 2018a) and *We Can!* (MEXT, 2018b), which are both widely used in English lessons (foreign language activities) in Japanese primary schools. Their findings revealed that *katakana* English accounts for about 70% and 50% of textbooks' vocabulary for reading and listening, respectively. These facts imply that many children in Japan already know some *katakana* English even before they start learning English in school.

## 2.2 Various Types of English Vocabulary Knowledge Tests for Primary Schoolchildren in Japan

A few previous studies have conducted English vocabulary knowledge tests on Japanese children. Kasahara et al. (2012) conducted three types of English vocabulary knowledge tests for fifth- and sixth-grade children: (1) a sound-meaning recognition test in which examinees chose the correct picture from four options (e.g., “wallet,” “plastic bottle,” “cashbox,” “bag”) when they heard the target word pronounced (e.g., /bæg/); (2) a sound-spelling recognition test in which examinees chose the correct spelling from four options (e.g., “giraffe,” “laugh,” “picture,” “zebra”) when they heard the target word pronounced (e.g., /dʒərəʃ/); and (3) a spelling-meaning recognition test in which examinees chose the correct picture from four options (e.g., “hat,” “scarf,” “glasses,” “dress”) when they saw the target word's spelling (e.g., “dress”). Among all three tests, the correct answer rate was significantly higher for the sound-meaning recognition test (92.8%) than for the sound-spelling recognition test (84.6%), and the spelling-meaning recognition test (84.7%).

Moreover, Koshiba (2015) conducted three types of tests for third- to sixth-graders: (1) a sound-meaning recall test in which children wrote meanings of the words after listening to their pronunciation; (2) a meaning-sound recall test in which children pronounced the word when they looked at its illustration; and (3) a spelling-meaning recall test in which children wrote the meanings of the words when they saw their spelling. Among the four grades, the sound-meaning recall score was the highest.

In general, previous studies have shown two main results: first, previous studies have indicated some word features that make it easier for Japanese learners to remember English words. Yoshimura (2009) indicated that test scores for *katakana* English were higher than those for non-*katakana* English. On the other hand, word frequency level and word length are key for the ease of vocabulary learning (Willis & Ohashi, 2012).

Second, most studies have conducted a sound-meaning test, in which students performed the best. Conceivably, this is because the *Course of Study's* main goals for primary school English education is not to help children develop a written language, but to help them develop basic communication skills and provide

many opportunities to interact with their classmates, homeroom teachers, and assistant language teachers (MEXT, 2017). Therefore, children's enhancement of sound-meaning connections is crucial for familiarising them with basic English oral communication. This is consistent with Barcroft (2020), who argued that "all language begins with input" (p. 489). This is why most of the previous studies used the sound-meaning test and found it to yield more significant results than other types of tests, such as spelling-meaning and sound-spelling tests.

### **2.3 Previous Studies' Limitations and This Study's Purposes**

The studies conducted previously however, had several limitations. First, the participants in Willis and Ohashi (2012) were not primary schoolchildren. They investigated variables that made it easy to remember words. However, their survey was conducted with Japanese university students, and their results have not yet been confirmed among children. Second, their study used vocabulary lists derived from the British National Corpus (BNC), which includes 100 million written and spoken British English words. Even though their study found high-frequency words easier to remember, the BNC is not designed for Japanese learners of English; thus, the target words used in this study were selected from JACET 8000 (The Japan Association of College English Teachers, 2016) and SVL 12000 (ALC, 2001), both developed for Japanese learners of English.

Third, previous studies did not discuss how their results could be applied to actual L2 vocabulary teaching in primary schools. Although previous studies conducted vocabulary knowledge tests for primary schoolchildren in Japan, their purpose was to compare the mean scores between the different types of tests.

Considering the limitations of the previous studies, the present study is original in that it aims to: (1) conduct a sound-meaning recognition test and build a predictive model from training data and (2) examine whether the developed model is applicable to new data. In short, this study is based on a framework for supervised machine learning. Training data or input data were collected for statistical analysis and for building a predictive model. On the other hand, testing data or new data are used to determine whether the model developed from the training data can fit into new data.

Relative to the first aim mentioned earlier, third- and fourth-graders took a sound-meaning recognition test, in which they listened to a target word's pronunciation and chose the appropriate picture from four options. As English beginners are not prepared for demanding vocabulary tests, such as a sound-spelling recall test, and because it would be ideal if English beginners could acquire words for spoken communication, this study conducted a sound-meaning recognition test.

Furthermore, this study aimed to develop a model that primary school English teachers can use. Two main types of variables that affect test scores are: (1) learners and (2) words. In the case of learners, even though it is possible to adopt such variables as learners' grades or lengths of English learning experience, this study adopted learners' grades as the first variable. The length of English learning experience is known to be an important variable that affects scores in the sound-meaning recognition test. Ojima et al. (2011) reported a clear relationship

between second-language proficiency and total study time; in other words, the earlier children begin studying English, the longer their experience of second-language learning. However, the reality is that English teachers have a variety of students with different needs to be considered, and English learning experience cannot be used in the predictive model because it depends heavily on individual children's experiences. Therefore, this study adopted a learner's grade as the first variable.

As for the variable of words, previous studies have revealed features that ease vocabulary learning, such as word length, word frequency, and *katakana* English. In addition, other features include L2 vocabulary words' pronunciation (Willis & Ohashi, 2012), imaginability (Karakas, 2020), and familiarity (Hoshino & Shimizu, 2019; Saji & Saeki, 2013). However, this study adopted word length as its second variable. Even if placing many variables in the model makes it more accurate, it can become too complex to use. For example, to measure word familiarity, Nishide and Mizumoto (2009) asked Japanese university students to score a certain word's familiarity on a 5-point scale, with a score of 5 meaning the word is completely familiar. For example, the word "apple" received an average score of 4.91 out of 5. Therefore, if this study had constructed its predictive model by adopting such a variable, users would always be required to check each target word's familiarity value and input the value into the regression formula, which is a time-consuming process. The significance of this study lies in its construction of an accurate predictive model from a small number of variables: children's grades and word length.

On the other hand, as detailed in the second aim of the study, the constructed model's quality for predicting new data still requires confirmation. The model made it possible to calculate the predicted percentages of children correctly identifying a certain target word in the sound-meaning recognition test by placing variables in the regression formula. Therefore, this study compared the predicted percentages from training data with actual percentages from the testing data to investigate whether the newly developed model can be reliable in predicting new data. The research question is as follows: can a model that predicts Japanese primary schoolchildren's English vocabulary knowledge be applied to new data?

## 3 Study

### 3.1 Participants and Target Words

A total of 144 third- and 155 fourth-graders ( $n = 299$ ) participated in this study. The experiment was conducted in July 2019. The third-graders had been learning English for 3 months in 10 lessons since April 2019. In contrast, the fourth-graders had been learning English for 15 months in 30 lessons since April 2018. All participants learned from *Let's Try!* (MEXT, 2018a), a textbook widely used in Japan.

This study's 25 target words were all nouns with the feature of *katakana* English chosen from *Let's Try!* (MEXT, 2018a), and all within the first 1,000 words ranked by level of word frequency according to JACET 8000 and SVL 12000 (see Table 1 below). JACET 8000 contains eight levels of 1,000 words for Japanese

Table 1. Target Words Used in this Study

Words	Letters	Syllables	Phonemes	Words	Letters	Syllables	Phonemes
Apple	5	2	3	Bag	3	1	3
Ball	4	1	3	Bed	3	1	3
Book	4	1	3	Bus	3	1	3
Cake	4	1	4	Card	4	1	4
Cat	3	1	3	Dinner	6	2	5
Dog	3	1	3	Egg	3	1	2
Flower	6	2	6	House	5	1	4
Library	7	3	8	Milk	4	1	4
School	6	1	4	Science	7	2	6
Shirt	5	1	4	Star	4	1	4
Table	5	2	5	Telephone	9	3	8
Tree	4	1	3	Rain	4	1	4
Room	4	1	3				

learners of English. Similarly, SVL 12000 includes 12 levels of 1,000 words. New Word Level Checker (<https://nwlc.pythonanywhere.com/>) was used to determine the frequency level for both JACET 8000 and SVL 12000. In addition, this study used two *katakana* English dictionaries published by Sanseido Henshujo (2019) and Horiuchi (2013) to ascertain that the target words were *katakana* English. Based on Allen (2019) and Shibuya (2012), this study defines *katakana* English as: (1) Japanese words that share Japanese form and meaning with English words; (2) words that have been borrowed into Japanese; (3) words that are often used in daily life in Japan; and (4) words that are not Japanese–English words, such as *カステラ* (*castella*) [*castella cake*] and *エアコン* (*aircon*) [*air-conditioner*], which cannot be understood by English speakers.

### 3.2 Procedure and Data Analysis

This study conducted a sound-meaning recognition test, which required third- and fourth-grade Japanese primary school students to listen to an English word's pronunciation (e.g., /æpl/) and then choose the appropriate picture from four options (see Appendix I). During the experiment, the participants took the sound-meaning recognition test, where they listened to a native English speaker's pronunciation using *GlobalvoiceEnglish3*. They were able to listen to an English sound twice per question. It took approximately 10 minutes to complete the experiment.

After the experiment, the collected data were divided into training and testing data. For training data, 70% of the collected data were chosen at random, and the remaining 30% were chosen as testing data, based on Katada and Fukuzawa (2019). Training data were used for statistical analysis, and testing data were used to calculate the mean score of the sound-meaning recognition test. Among the 144 third-graders, 100 participants' data were categorised into training data and 44 into testing data. Similarly, among the 155 fourth-graders, 109 participants' data were categorised into training data and 46 into testing data.

The study employed a generalized linear mixed model (GLMM), which combines a generalized linear model (GLM) and a linear mixed model (LMM). The GLM includes some regression models, and this study adopted the logistic regression model (LRM), which can be used when the outcome or dependent variable is binomial data. Since participants chose an appropriate picture from four options, they were either correct or incorrect, and thus produced binomial data. The GLMM can also be used to develop a regression formula that enables primary schoolteachers to predict the difficulty of students learning a certain word. Moreover, the GLMM considers fixed and random effects.

Table 2 shows how the collected data were entered into an Excel file. In the case column, there are 7,475 cases (Case 1 to Case 7,475) in total because 299 participants took the test that included 25 target words ( $299 \times 25$ ). The column labelled “Participant” reflects children’s sample number from 1 to 299. In the “Grade” column, third-graders were coded as 0, and fourth-graders were coded as 1. Similarly, in the “Data” column, training data were coded as 0, and testing data were coded as 1. In the “Letter,” “Syllable,” and “Phoneme” columns, the numbers depended directly on each target word’s features. Moreover, in the “Correct” column, 1 represented a correct answer, and 0 an incorrect answer.

For example, Case 1 means that a third-grade child (Participant No. 1) was assigned to training data and correctly chose “apple”—a target word with five letters, two syllables, and three phonemes. Similarly, Case 8 means that a fourth-grade child (Participant No. 4) was assigned to testing data but could not choose the correct picture of “egg,” which has three letters, a syllable, and two phonemes.

The GLMM was conducted to investigate whether the training and testing data were equally divided. Correct was considered a dependent variable, and grade and data as independent variables. Finally, participant was considered to be a random effect.

The GLMM was also conducted to build the most accurate predictive model from the training data. Two independent variables, grade and length of words (i.e., letter, syllable, and phoneme) represented the fixed effect, and participant the random effect. The most appropriate model was based on Akaike’s information criterion (AIC).

Indeed, the regression formula constructed from training data made it possible to predict the percentages of children who knew certain words depending

Table 2. Examples of Data Input

Case	Participant	Grade	Data	Letter	Syllable	Phoneme	Word	Correct
1	1	0	0	5	2	3	Apple	1
2	1	0	0	3	1	3	Bag	1
3	2	0	1	4	1	4	Cake	0
4	2	0	1	6	2	5	Dinner	1
5	3	1	0	5	1	4	House	0
6	3	1	0	9	3	8	Telephone	1
7	4	1	1	5	1	4	Shirt	1
8	4	1	1	3	1	2	Egg	0

on their grade and on the length of words. A correlation analysis was conducted between the predicted percentage from the training data and the actual percentage from the testing data to examine the research question.

## 4 Results

Results of the sound-meaning recognition test ( $Max = 25$ ) from the training data showed that third-graders had a mean of 21.5 (standard deviation [ $SD$ ] = 8.68), and fourth graders 21.9 ( $SD = 8.25$ ). On the other hand, results from the testing data showed that third-graders had mean of 21.0 ( $SD = 9.25$ ) and fourth-graders of 21.7 ( $SD = 8.48$ ). Table 3 shows the results of fixed effects and presents the following: estimated coefficient (*Estimated*), odds ratio (*OR*), standard error (*SE*), *z*-value, *p*-value, and a 95% confidence interval (95% *CI*). Table 3 shows no significant effect of either grade or data, indicating that the two datasets were equally divided.

Models 1, 2, and 3 were developed to ascertain the ideal model. Grade and letter were adopted in Model 1, grade and syllable in Model 2, and grade and phoneme in Model 3. Table 4 shows the three models and their AIC values. The results revealed that Model 3 was the most appropriate because the AIC values of this Model were the lowest among the three models.

Table 5 shows Model 3's fixed effects and presents an estimated coefficient (*Estimated*), *OR*, *SE*, *z*-value, *p*-value, and a 95% confidential interval (95% *CI*). The Table shows no significant effect of grade (*Estimate* = 0.26, *SE* = 0.25, *z* = 1.03, *p* = 0.305), but a significant effect of phoneme (*Estimate* = -0.56, *SE* = 0.04, *z* = -12.89, *p* < 0.001). This means: (1) there is no difference in test scores between

Table 3. Results of the Sound-Meaning Recognition Test

Fixed effect	<i>Estimated</i>	<i>OR</i>	<i>SE</i>	<i>z</i> -value	<i>p</i> -value	95% <i>CI</i>
Intercept	2.23	9.34	0.15	14.42	<0.001	[1.93, 2.54]
Grade	0.18	1.20	0.19	0.97	0.33	[-0.18, 0.55]
Data	-0.13	0.88	0.20	-0.65	0.52	[-0.53, 0.27]

OR, odds ratio; SE, standard error; CI, confidence interval.

Table 4. The Value of AIC Affecting the Best Model

Variables	Model 1	Model 2	Model 3
	<i>Estimated</i> ( <i>SE</i> )	<i>Estimated</i> ( <i>SE</i> )	<i>Estimated</i> ( <i>SE</i> )
Intercept	2.82 (0.27) ***	3.94 (0.25) ***	4.88 (0.30) ***
Grade	0.14 (0.23)	0.24 (0.24)	0.26 (0.25)
Letter	-0.12 (0.04) **		
Syllable		-1.08 (0.09) ***	
Phoneme			-0.56 (0.04) ***
AIC	1855.7	1734.6	1667.7

\*\*\**p* < 0.001, \*\**p* < 0.01.

SE, standard error; AIC, Akaike's information criterion.

Table 5. Fixed Effect from GLMM (Model 3)

Fixed effect	<i>Estimated</i>	<i>OR</i>	<i>SE</i>	<i>z-value</i>	<i>p-value</i>	<i>95% CI</i>
Intercept	4.88	131.78	0.30	16.45	<0.001	[4.30, 5.46]
Grade	0.26	1.29	0.25	1.03	0.305	[-0.23, 0.75]
Phoneme	-0.56	0.57	0.04	-12.89	<0.001	[-0.65, -0.48]

OR, odds ratio; SE, standard error; CI, confidence interval.

third- and fourth-graders; and (2) since the estimated coefficient is negative, the more phonemes a word has, the lower the test score.

Since GLMM is part of LRM, the OR can illustrate a certain variable's effectiveness. The estimated coefficients in Table 5 illustrate the logarithmic ORs. When values change into ORs, their meanings become clearer. The value of the estimated coefficient of phoneme was  $-0.56$ , which is equal to the  $0.57$  OR ( $= \exp [-0.56]$ ). This means that the correct rates of words with four phonemes (e.g., "cake") were  $0.57$  ORs higher than words with three phonemes (e.g., "apple"). The same was true in the comparison of five-phoneme (e.g., "dinner") and four-phoneme words (e.g., "card"). Furthermore, correct rates of words with six phonemes (e.g., "flower") were  $0.33$  ORs higher than those of words with four phonemes ( $= \exp [-0.56 \times 2]$ ).

The following regression formula made it possible for primary schoolteachers to predict the ease or difficulty of teaching certain words:

$$\text{Regression formula} = 4.88 + 0.26 \times \text{grade (0 or 1)} - 0.56 \times \text{phoneme}$$

To predict whether third-graders knew the *katakana* English word "apple," the number of phonemes (3) and the participant's grade (0) were included in the regression formula. Then, the value of the estimated coefficient (3.20) was calculated as equal to 24.53 odds ( $= \exp [3.20]$ ). Odds signifies the value attained when the probability of success is divided by that of unsuccess (success percentage/ unsuccess percentage). In this case, success and unsuccess imply that learners can answer correctly and incorrectly, respectively. Odds can be used as shown in the following formula:

$$\text{Successful Percentage} = \frac{\text{odds}}{1 + \text{odds}}$$

When odds of 24.53 were placed in the formula, the percentage of success was 0.961. This suggests that when third-graders listen to the word "apple" (/æpl/) being pronounced, the probability that they know its meaning is 96.1%. In this way, teachers can predict the probability of known words based on the word's number of phonemes and students' grades. Table 6 displays the predicted percentages from GLMM (Model 3) obtained from the training data and the actual percentages obtained from the testing data.

Spearman's rank method was applied to examine the correlation between the predicted percentages from the training data and the actual percentages from the testing data. The two types showed moderate correlation in third-graders

Table 6. Relationship between Predicted Percentages and Actual Percentages (%)

Data Types	Grade	Apple	Bag	Ball	Bed	Book	Bus	Cake
Training	3rd	96.1	96.1	96.1	96.1	96.1	96.1	93.3
Data	4th	97.0	97.0	97.0	97.0	97.0	97.0	94.5
Testing	3rd	100	95.5	72.7	100	86.4	100	100
Data	4th	100	88.9	88.9	100	94.4	100	100
Data Types	Grade	Card	Cat	Dinner	Dog	Egg	Flower	House
Training	3rd	93.3	96.1	88.9	96.1	97.7	82.1	93.3
Data	4th	94.5	97.0	91.2	97.0	98.2	85.6	94.5
Testing	3rd	86.4	100	72.7	100	90.9	77.3	100
Data	4th	94.4	94.4	92.9	100	100	72.2	100
Data Types	Grade	Library	Milk	Rain	Room	School	Science	Shirt
Training	3rd	59.9	93.3	93.3	96.1	93.3	82.1	93.3
Data	4th	65.9	94.5	94.5	97.0	94.5	85.6	94.5
Testing	3rd	40.9	63.6	77.3	86.4	54.5	72.7	68.2
Data	4th	46.4	83.3	89.3	89.3	50.0	67.9	72.2
Data Types	Grade	Star	Table	Telephone	Tree			
Training	3rd	93.3	88.9	59.9	96.1			
Data	4th	94.5	91.2	65.9	97.0			
Testing	3rd	95.5	95.5	72.7	90.9			
Data	4th	94.4	96.4	78.6	85.7			

Table 7. Predicted Successful Percentages of Katakana English Words with Different Phonemes

Grade	Phoneme	PSP	Examples	Grade	Phoneme	PSP	Examples
3rd	3	96.1%	Pen	3rd	6	82.1%	Orange
4th	3	97.0%	Red	4th	6	85.6%	Mountain
3rd	4	93.3%	Heart	3rd	7	72.3%	Baseball
4th	4	94.8%	Horse	4th	7	77.2%	Hospital
3rd	5	88.9%	Salad	3rd	8	59.9%	Library
4th	5	91.2%	River	4th	8	65.9%	Telephone

PSP, predicted successful percentage.

( $\rho = 0.53$  [large effect size], 95%  $CI = [0.18, 0.77]$ ,  $p = 0.006$ ) and in fourth-graders ( $\rho = 0.54$  [large effect size], 95%  $CI = [0.18, 0.77]$ ,  $p = 0.005$ ), meaning that a word predicted to be easy tends to have a higher actual score and vice versa. Thus, Model 3 was moderately higher in terms of predicting Japanese primary school-children's knowledge of English vocabulary.

Table 7 shows the predicted successful percentage (PSP) based on grade and phoneme, and some examples from the textbook, *Let's Try!* (MEXT, 2018a), respectively. The use of the PSP is discussed in the next section.

## 5 Discussion

As Table 4 reveals, Model 3 was the most appropriate for predicting new data, and the grade and phoneme variables were placed into it. Of letter, syllable,

or phoneme, the most suitable variable was phoneme, which was attributed to the type of test. The study adopted the sound-meaning recognition test, in which children listened to an English word pronounced (e.g., /æpl/) and chose the appropriate picture (of an apple) from four options. The participants did not look at the target word's spelling; in other words, the phoneme variable, related to sound information, was key to the sound-meaning recognition test.

Moreover, the syllable variable is not flexible in representing the features of target words. As Table 1 shows, the range of syllables was limited from one to three. On the other hand, word phonemes ranged from two to eight, which is why Model 3 was flexible enough to predict new data.

As Table 5 shows, the grade variable showed no significant difference: third-graders' scores were as high as those of fourth-graders. This result observed a ceiling effect, in that the total mean scores and SD exceeded the maximum score (Kanayama, 2020). The third-graders' total average score ( $M = 21.5$ ) and the standard deviation ( $SD = 8.68$ ) exceeded the maximum score ( $Max = 25$ ). The same was true for fourth-graders ( $M = 21.9$ ,  $SD = 8.28$ ). Indeed, the ceiling effect prevents researchers from measuring the knowledge acquired by learners (Kanayama, 2020). This likely occurs because the target words were all nouns with features of *katakana* English and because they were from the first 1,000 levels of JACET 8000 and SVL 12000. In short, participants knew most of the target words well, enabling them to achieve high scores and leading to a ceiling effect for both grades.

Answering this study's research question of whether or not a model that predicts Japanese primary schoolchildren's English vocabulary knowledge can be applied to new data, Spearman's rank method found a moderate correlation between predicted percentages from training data and actual percentages from testing data, as shown in Table 6. This revealed a general tendency of schoolchildren to know "easy" words—that is, words with fewer phonemes—but not to be as familiar with "difficult" words—for example, words with more phonemes.

One purpose of this study was to provide pedagogical implications for teaching English vocabulary in Japanese primary schools. As mentioned above, the regression formula was found to be applicable. For instance, when third- and fourth-graders listened to the pronunciation of *katakana* English words and chose their meanings, the regression formula made it possible for primary schoolteachers to predict the ease or difficulty of teaching certain words. As Table 7 shows, it is predicted that most learners knew the meaning of *katakana* English words with three to four phonemes because *PSP* was over 90% for both grades. Therefore, teachers can spend less time teaching such words. Instead, they should encourage children to use the words they already know, enabling them to be used productively. In contrast, third-graders' rates for *katakana* English words with eight phonemes were lower than 60%, indicating that teachers should spend more time introducing such words.

## 6 Conclusion

In this study, third- and fourth-graders took the sound-meaning recognition test, and GLMM from training data revealed no significant effect of grade, but a

significant effect of phoneme. From the fixed effect of GLMM, the most appropriate model was constructed to predict the scores of the sound-meaning recognition test from the training data. The model was found to be moderately reliable because of the moderate correlation between the training and testing data. Thus, the results of this study are significant because the developed model can predict primary schoolchildren's English vocabulary knowledge as long as the tested words are nouns from *katakana* English and high-frequency.

This study has some limitations. First, because participants were limited to third- and fourth-graders, whether the predictive model will apply to fifth- and sixth-graders remains unknown. Second, the sound-meaning recognition test was the only one administered; therefore, the model could not predict other types of English vocabulary knowledge. Further studies should employ, for instance, a sound-meaning recall test, a meaning-sound test, a spelling-sound test, and a spelling-meaning test. Third, the largest limitation is that target words were limited to *katakana* English nouns from the first 1,000 frequency levels of both JACET 8000 and SVL 12000; thus, the model cannot predict all words. For instance, the *katakana* English word “peach” cannot be predicted because it is not in the first 1,000 level but in the second 1,000 frequency level of both JACET 8000 and SVL 12000. Finally, a limited number of learners participated in the study. Considering that a large amount of data would make the regression formula more accurate and reliable, future studies should gather more data. Considering these limitations into account, future studies should ask various grades to take various vocabulary knowledge tests that include both non-*katakana* and *katakana* English with different word frequency levels.

## References

- ALC. (2001). *Standard vocabulary list 12000*. <https://www.alc.co.jp/eng/vocab/svl/list.html>
- Allen, D. (2019). An overview and synthesis of research on English loanwords in Japanese. *Vocabulary Learning and Instruction*, 8(2), 8–25. <https://doi.org/10.7820/vli.v08.2.Allen>
- Barcroft, J. (2020). Key issues in teaching single words. In S. Webb (Ed.), *The Routledge handbook of vocabulary studies* (pp. 479–492). Routledge.
- Daulton, F. E. (2008). *Japan's built-in lexicon of English-based loanwords*. Multilingual Matters.
- Horiuchi, K. (2013). *Gendaiyougo no kisochohshiki katakana gairaigo ryakugo jiten [Basic knowledge of modern term: Dictionary of Katakana]* (5th ed.). Jiyukokuminsha.
- Hoshino, Y., & Shimizu, H. (2019). Katakana words in English textbooks used in Japanese primary schools—A comparative study of word senses between Japanese and English. *The Japan Association of English Teaching in Elementary School*, 19, 117–129. [https://doi.org/10.20597/jesjournal.19.01\\_117](https://doi.org/10.20597/jesjournal.19.01_117)
- Kanayama, K. (2020). Is expanding spacing more effective than equal spacing for L2 vocabulary learning? *Annual Review of English Language Education in Japan*, 31, 1–16. [https://doi.org/10.20581/arele.31.0\\_1](https://doi.org/10.20581/arele.31.0_1)

- Karakas, A. (2020). Exploring psycholinguistic difficulties in learning foreign language vocabulary in light of the English language. *i-manager's Journal of Educational Psychology*, 13(3), 36–45. <https://doi.org/10.26634/jpsy.13.3.16546>
- Kasahara, K., Machida, N., Osada, E., Takashi, T., & Yoshizawa, S. (2012). Vocabulary knowledge of 5th and 6th graders at elementary school: Connections between sound, meaning, and spelling. *The Japan Association of English Teaching in Elementary School*, 12, 90–101. [https://doi.org/10.20597/jesjournal.12.0\\_90](https://doi.org/10.20597/jesjournal.12.0_90)
- Katada, Y., & Fukuzawa, S. (2019). *Chokkan de wakaruru Excel de kikaigakushu [Machine learning using Excel by intuition]*. Impress.
- Koshihara, W. (2015). Comparison of the vocabulary knowledge of third and sixth grade students—Through meaning, phonology, and orthography. *The Japan Association for the Study of Teaching English to Children*, 34, 39–52.
- MEXT (2017). *Shogakkou gakushu shido yoryo kaisetsu gaikokugo katsudo gaikokugo hen* [Study of course guideline for primary school English version]. Kairyudo.
- MEXT (2018a). *Let's try!* Tokyoshoseki.
- MEXT (2018b). *We can!* Tokyoshoseki.
- Nishide, K., & Mizumoto, A. (2009). An attempt to measure the familiarity level of 8000 English words. *Tsuru University Graduate School Review*, 13, 57–92. <http://trail.tsuru.ac.jp/dspace/bitstream/trair/180/1/Y013057.pdf>
- Ojima, S., Matsuba-Kurita, H., Nakamura, N., Hoshino, T., & Hagiwara, H. (2011). Age and amount of exposure to a foreign language during childhood: Behavioral and ERP data on the semantic comprehension of spoken English by Japanese children. *Neuroscience Research*, 70, 197–205. <https://doi.org/10.1016/j.neures.2011.01.018>
- Saji, R., & Saeki, Y. (2013). Word familiarity and English pronunciation ability of Japanese 4<sup>th</sup> graders. *The Japan Association of English Teaching in Elementary School*, 13, 150–162.
- Sanseido Henshujo [Sanseido Editing Center]. (Ed.). (2019). *Miyasui katakana shingo jisho [Easy-to-read dictionary of katakana]* (3rd ed.) Sanseido.
- Shibuya, T. (2012). *Shougakkou yonensei no katakana eigo no imi no rikai: Eigo bogo washa no hatsuon suru eitango no rikai [Recognition of English loanwords by fourth year pupils: Understanding English words pronounced by native speakers]*. *The Japan Association of English Teaching in Elementary School*, 12, 44–56. [https://doi.org/10.20597/jesjournal.12.0\\_44](https://doi.org/10.20597/jesjournal.12.0_44)
- The Japan Association of College English Teachers. (2016). *The new JACET 8000 basic words*. Kirihara.
- Willis, M., & Ohashi, Y. (2012). A model of L2 vocabulary learning and retention. *The Language Learning Journal*, 40(1), 125–137. <https://doi.org/10.1080/09571736.2012.658232>
- Yoshimura, H. (2009). Koritsu shogakko jido no goichisiki ni kansuru chousa no jissenhoukoku [Practical report on examining vocabulary knowledge of children in public primary school]. *The Japan Association of English Teaching in Elementary School*, 9, 23–30. [https://doi.org/10.20597/jesjnl.9.0\\_23](https://doi.org/10.20597/jesjnl.9.0_23)

## Appendix I: Sample of Sound-Meaning Recognition Test

Listen carefully and circle the picture of the words you hear.

Question 1 (/æpl/)



Question 2 (/bæg/)



Question 3 (/mlk/)



*Note.* All of the original instructions were written in Japanese.