Role of Learners' Subjective Difficulty Rating Toward a System for Practicing English-Speaking

Harumi Kashiwagi, Kobe University, Japan
Min Kang, Kobe University, Japan
Kazuhiro Ohtsuki, Kobe University, Japan

Abstract

This preliminary study on how to develop a system for English-speaking practices explores the role a subjective difficulty rating should play in such a system, using a questionnaire and a level-based vocabulary list. By selecting 75 English verbs from the five difficulty levels in the list, 72 Japanese university students rated the difficulty of answering questions in English verbally on a five-point Likert scale. The results suggest that words of difficulty Levels 4 and 5 should be targeted in designing a speaking practice system for the participating students. Moreover, all the selected words in Level 3 offered some response variance in difficulty rating, with standard deviation (SD) scores of 1.0 or higher, suggesting that the choice in the selection of Level 3 words depends on the individual student. Considering the detailed results, all words with SD scores of difficulty ratings lower than 1.0 in Levels 4 and 5 were evaluated as difficult or relatively difficult by more than 80% of the students. This indicates that any speaking practice system should consider words from Levels 4 or 5 for which SD scores in difficulty ratings are lower than 1.0 as difficult words for these students. Although further studies are needed, these results indicate that the average subjective difficulty rating scores can likely provide an indication of the ideal difficulty level to target in an English-speaking practice system. Moreover, SD scores could help customize the target vocabulary for each student.

Keywords: Subjective Difficulty Rating, Speaking Practice, Language Use, Language Learning System.

Introduction

The Japanese Ministry of Education, Culture, Sports, Science and Technology (2015) created the English proficiency promotion plan to comprehensively develop four core English skills among Japanese students, especially their ability to produce English. However, research has indicated that English language learners have substantial speaking anxiety (Suleimenova, 2013); and this includes Japanese students (Woodrow, 2006). The results of our previous study (Kashiwagi, Kang, & Ohtsuki, 2018) also demonstrated that many students do not feel that they possess English-speaking proficiency. It can therefore be assumed that many Japanese students are not confident about speaking in English. The lack of opportunities for them to speak English in everyday life, outside language classes at school, may contribute to this lack of confidence. A practice environment in which learners can become familiar with speaking English is therefore necessary.

When considering how to promote English-speaking practices, we need to observe which language items are most difficult in spoken English, and identify problematic words and phrases. Thus, we have been developing a prototype system for English-speaking practices (Kashiwagi, Kang, & Ohtsuki, 2020). This system attempts to support learners by providing information on the extent to which they should attempt to use certain words and phrases. This paper describes our preliminary study of exploring how a subjective difficulty rating should play a role in designing a system of speaking practice for students. By using a level-based vocabulary list and a questionnaire on the subjective difficulty levels of English verbs, this study investigates the following research questions:

RQ1. In selecting words for English-speaking practice, what level of difficulty should be targeted?
RQ2. For which levels of word difficulty do the difficulty ratings provided by students have the greatest response variance?
RQ3. Which words do students consider most difficult?
RQ4. Could a subjective difficulty rating of English words help further develop our system of English-speaking practice in a way that aids learners' use of certain vocabulary?

We conducted this study on Japanese university students, selecting words from the vocabulary list "The New JACET List of 8000 Basic Words" (The Basic Word List Revision Special Committee of JACET, 2016; hereafter, "the New JACET8000"). The remainder of this paper describes our experiment's methodology, a discussion of our results, conclusions, the study's limitations, and recommendations for additional research.

Related Studies

In the field of technology and L2 speaking, some automatic speech recognition functions have been integrated into the language learning software, such as Rosetta Stone®. They enable students to carry out interactive tasks within a limited range. According to McCrocklin (2016), the introduction of this technology helps students become more autonomous in their pronunciation practice. Another research (Blake & Shiri, 2012) reported on the successes and challenges of teaching Arabic within a distance learning environment. One of their findings indicated that small-group computer-mediated communication sessions with sound and text gave students more personal attention. Furthermore, the internet-based 3-dimensional virtual environment called "Second Life" has been integrated into EFL programs. According to Wang, Song, Xia, and Yan (2009), participants perceived "Second Life" as useful and

interesting and perceived the EFL program in this virtual environment as interesting. Jehma (2020) concluded that the students developed their English communication skills through the activity with "Second Life." Thus, the use of technology has the potential to enhance students' speaking ability, motivation, and autonomy.

However, few studies have focused on students' self-assessment of their own performances (Janulevičienė & Kavaliauskienė, 2007). In some cases, self-assessment might be unreliable and not reflective of the learner's actual performance (Todd, 2002). Nevertheless, it could raise learner awareness of language use and lead to developing learner responsibility and autonomy (Janulevičienė & Kavaliauskienė, 2007). We believe that learners' subjective difficulty ratings on the detailed learning items play meaningful roles in observing their learning processes.

Our Prototype System

We have been developing a prototype system for practicing English-speaking (Kashiwagi, Kang, & Ohtsuki, 2020). This system aims to support learners by providing information on the extent to which they should attempt to use certain words and phrases. The structure of the prototype system is shown in Figure 1.
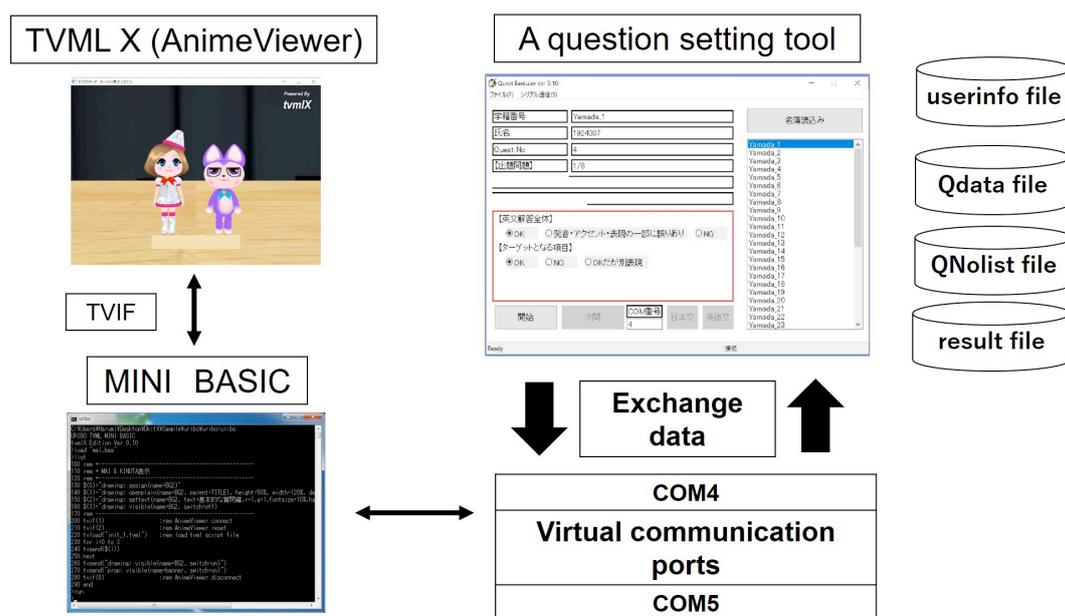

Figure 1: Prototype System Overview

Our system consists of three different types of software: a question setting tool, MINI BASIC, and AnimeViewer. The following provides an overview of how the system operates. When we start the question setting tool, question related data are sent to the software MINI BASIC, and converted into a script file for generating computer-generated (CG) content. CG characters then appear on the screen of the AnimeViewer, a viewer tool for displaying CG content. Next, CG characters give instructions and provide a question, which the student orally answers. While listening to the student's answer, the teacher checks their answer using the question setting tool. When the system finishes providing questions, and after the answers have been checked, the results are saved in the result file in the question setting tool.

Though our prototype system can check whether students give correct answers, it does not enable us to assess how students truly feel about the difficulty of using certain words. It is important to observe students' reflective assessment about word difficulty to observe their learning situations on an individual basis. Therefore, the role of students' subjective difficulty ratings for selecting the words used in the system must be further explored.

Methods

Word List Creation

Words for our study were chosen from the New JACET8000, an educational vocabulary list intended for Japanese university students who have studied English in middle school, high school, and/or at university in Japan. The list features 8000 words typically used in middle and high school English textbooks, entrance examinations for public high schools and universities, introductory academic books, English newspapers, and standardized English tests such as TOEIC, TOEFL, and STEP (Mochizuki et al., 2016: 75). Eight thousand words are presented in eight levels (Rank 1–1000, Rank 1001–2000, Rank 2001–3000, Rank 3001–4000, Rank 4001–5000, Rank 5001–6000, Rank 6001–7000, and Rank 7001–8000) in accordance with the frequency and educational significance of each word. We selected the words from this list because it reflects the actual situation of English learning among Japanese university students.

In this study, 15 English verbs were chosen from each of the following five levels, creating a total of 75 words: Rank 1–1000, Rank 1001–2000, Rank 2001–3000, Rank 3001–4000, and Rank 4001–5000. We selected the verbs that students would have learned thus far. Words from the highest ranks—Rank 5001–6000, Rank 6001–7000, and Rank 7001–8000—were not included because they would have been too difficult for first-year students to translate orally. Hereafter, we refer to the different ranks as Levels 1–5, where Level 1 means Rank 1–1000, Level 2 means Rank 1001–2000, and so on.

Difficulty Level	Words (and their New JACET8000 ranking)
Level 1 (Rank 1–1000)	decide (560), grow (689), reduce (698), worry (723), share (725), improve (735), apply (741), shake (751), treat (777), remove (802), notice (874), invite (952), solve (953), fix (982), burn (999)
Level 2 (Rank 1001–2000)	bury (1341), complain (1366), confuse (1373), behave (1376), float (1425), accept (1580), publish (1668), regard (1833), hide (1902), ignore (1918), observe (1930), whisper (1932), compete (1952), threaten (1968), propose (1976)
Level 3 (Rank 2001–3000)	calculate (2072), dig (2076), melt (2135), apologize (2143), spoil (2170), invest (2310), analyze (2329), consume (2332), polish (2461), spill (2475), bend (2620), sweep (2630), wipe (2635), evaluate (2838), hesitate (2918)
Level 4 (Rank 3001–4000)	swallow (3421), exhaust (3580), stimulate (3589), infect (3612), suck (3617), abolish (3624), weave (3634), conquer (3638), swell (3639), expire (3646), penetrate (3741), cultivate (3751), summarize (3754), sigh (3804), sew (3910)
Level 5 (Rank 4001–5000)	diagnose (4283), accelerate (4439), shrink (4443), violate (4447), exaggerate (4462), insure (4467), inject (4477),

furnish (4478), starve (4479), contradict (4495), refine (4506), worsen (4522), undertake (4714), resign (4937), reinforce (4963)

Table 1: The Words Used in the Self-Reflective Feedback Questionnaire

Participants

This study's participants comprised 72 first-year students at a university in Japan. They participated in the activity mentioned in the following section, and responded to the self-reflective feedback questionnaire.

Procedures

We administered warm-up activities in the language classes in which Japanese verbs (Table 1) were provided to students, who were then asked to translate them into English orally. A self-reflective feedback questionnaire was then conducted to determine students' difficulty in translating each word into English. The difficulty values were scored using a five-point Likert scale (1= easy, 2= relatively easy, 3= neutral, 4= relatively difficult, and 5= difficult).

Results and Discussion

RQ1: In selecting words for English-speaking practice, what level of difficulty should be targeted?

To determine this, we first calculated the average difficulty ratings for the individual words to investigate student perceptions overall. The average difficulty ratings for the words of each level are shown in Figures 2 to 6. In these figures, the dotted red line indicates neutral difficulty values.
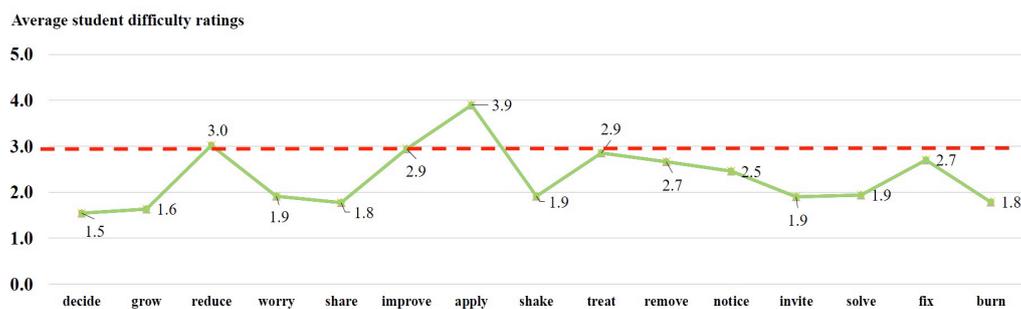

Figure 2: Average Student Difficulty Ratings of the Words in Level 1 (Rank 1–1000)
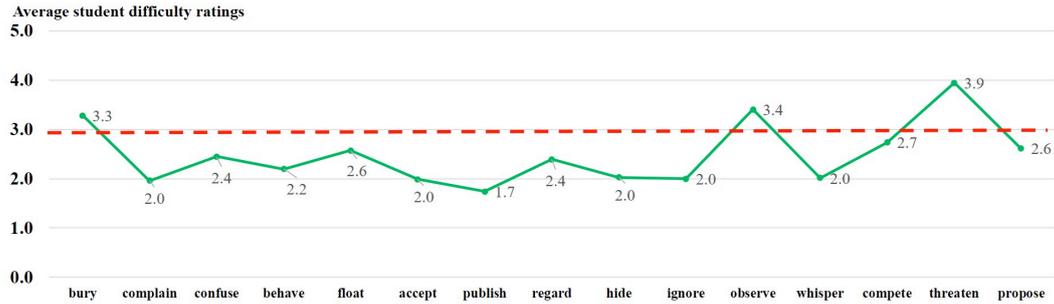
Figure 3: Average Student Difficulty Ratings of the Words in Level 2 (Rank 1001–2000)
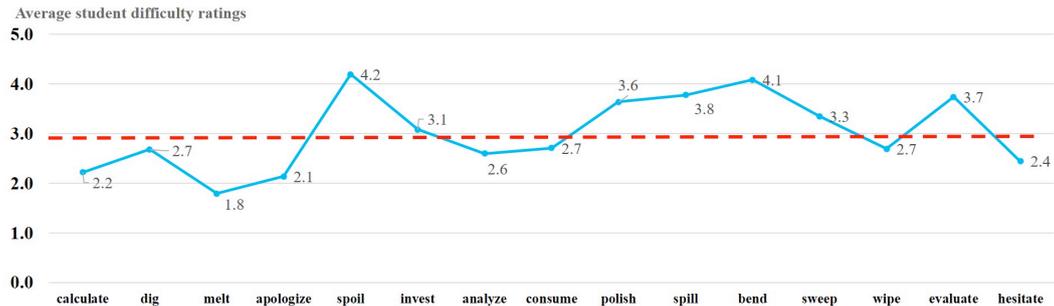

Figure 4: Average Student Difficulty Ratings of the Words in Level 3 (Rank 2001–3000)
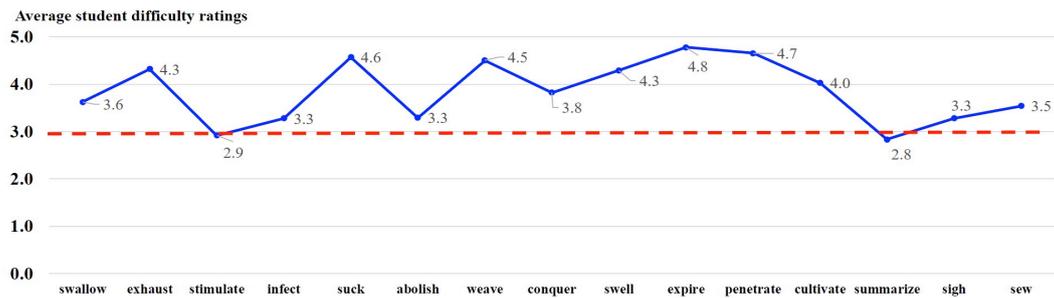

Figure 5: Average Student Difficulty Ratings of the Words in Level 4 (Rank 3001–4000)
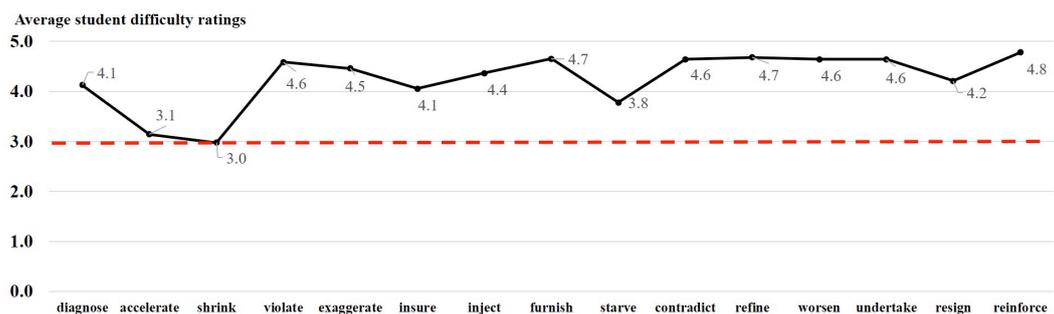
Figure 6: Average Student Difficulty Ratings of the Words in Level 5 (Rank 4001–5000)

The ratings tended to increase roughly in conjunction with how the words ranked on the New JACET8000. Specifically, the largest increase in the perceived difficulty occurred between Levels 3 and 4 (Figures 4 and 5, respectively). There was also some increase between Levels 2 and 3 (Figures 3 and 4, respectively).

Next, to determine which words should be targeted in English-speaking practice, we calculated the number of words in each level with an average difficulty rating higher than 3.

As shown in Table 2, the number of words in Levels 1 and 2 were two and three, respectively. There was a marked increase in the number of words in Level 3, which was seven. Furthermore, the number of words increased again to 13 and 14 in Levels 4 and 5, respectively. This means that the most marked increase in the number of words with a difficulty ranking higher than 3 was between Levels 3 and 4, which corresponds with the average ratings in the respective levels detailed above.

Level	Number of words with an average difficulty rating higher than 3
Level 1	2
Level 2	3
Level 3	7
Level 4	13
Level 5	14

Table 2: Number of Words Perceived as Most Difficult Per Level (15 Words Per Level)

Taken together, these results indicate that most students consider it difficult to translate the words from Levels 4 and 5 into English. Some students also considered it difficult to translate Level 3 words into English. These results suggest that English-speaking practices for the students in this study should target words from Levels 4 and 5. The words in Level 3 received more varied difficulty ratings from students, a phenomenon which is analyzed in more detail in the results for RQ2.

RQ2: For which levels of word difficulty do the difficulty ratings provided by students have the greatest response variance?

To answer this question, we calculated the standard deviation (SD) scores of the difficulty ratings in the words of each difficulty level. The number of words for which the SD scores were 1.0 or higher are shown in Table 3.

Level	Number of words with an SD score of 1.0 or higher
Level 1	12
Level 2	13
Level 3	15
Level 4	11
Level 5	6

Table 3: Response Variance in Difficulty Ratings

These results demonstrate a relatively large number of words for which the SD scores were 1.0 or higher in Levels 1 and 2. Even more pronounced was the response variance in difficulty ratings for Level 3 words, all of which had SD scores of 1.0 or higher. This variance is less pronounced in Levels 4 and 5. Based on these results, we recommend determining whether Level 3 words should be targeted based on the needs of the individual student.

RQ3: Which words do students consider most difficult?

Based on the findings for RQ1, we investigated which Level 4- and Level 5-words were most frequently considered difficult. To provide a rough indication, we determined which words were evaluated as being either "difficult" or "relatively difficult" by more than 80% of students as an indicator. We assessed these words further by calculating their average difficulty and the SD scores of these ratings (Table 4). As Table 4 shows, seven words in Level 4 and ten words in Level 5 were evaluated as difficult or relatively difficult by more than 80% of students. This criterion was met by all the words for which the difficulty ratings had SD scores lower than 1.0. These results suggest that words for which the SD scores were lower than 1.0 should be considered difficult words. They also suggest that an SD score of 1.0 in difficulty scores could be an indicator for determining what words are difficult for students.

Level	Word	The percentage of students who evaluated it as difficult or relatively difficult	The average value of the difficulty ratings	The SD score of the difficulty ratings
Level 4	exhaust	84.7	4.3	1.0
	suck	90.3	4.6	0.9
	weave	90.3	4.5	0.8
	swell	83.3	4.3	1.1
	expire	95.8	4.8	0.7
	penetrate	91.7	4.7	0.8
	cultivate	80.6	4.0	1.2
Level 5	violate	91.7	4.6	0.8
	exaggerate	88.9	4.5	0.8
	inject	86.1	4.4	1.3
	furnish	88.9	4.7	0.9
	contradict	94.4	4.6	0.8
	refine	95.8	4.7	0.8
	worsen	97.2	4.6	0.6
	undertake	94.4	4.6	0.7
	resign	87.5	4.2	0.99
	reinforce	98.6	4.8	0.5

Table 4: Level 4- and Level 5-Words Considered Difficult or Relatively Difficult

RQ4: Could a subjective difficulty rating of English words help further develop our system of English-speaking practice in a way that aids learners' use of certain vocabulary?

This study used a subjective difficulty rating with a five-point Likert scale to analyze the difficulty levels of certain words. The results suggest that average subjective difficulty ratings can be used as a rough index of the ideal vocabulary level for learners to target. Moreover, the SD scores of the difficulty ratings could help determine whether a specific word should be selected for all students or only some students. Though a subjective difficulty rating is not an objective vocabulary test, it allows us to observe how students truly feel about the difficulty of using certain words. This differs from a paper-based vocabulary test, in which it is difficult to assess students' feelings about word difficulty when they give correct answers. Although future studies are needed, subjective difficulty ratings can determine the best

vocabulary words for learners to practice on an individual basis, thus improving their learning situations. Our system aims to support learners by providing them with information to the extent to which they can use words and phrases. Students' subjective difficulty ratings should play a role in selecting the words used in English-speaking practices, based on the individual learner.

Limitations and Recommendations

The current study has certain limitations. It was conducted with a small group of students, and the feedback questionnaire investigated only 75 English verbs. More studies are needed to target a larger number of students and words. Additionally, it is possible that the perceived difficulty of using some words could change after students practice them, while other words may never become easier to use, even after practice. By conducting pre- and post-practice questionnaires, it is possible that the detailed changes in the perceived difficulty of each word after practice could be observed. We hope to investigate this in subsequent studies.

Conclusions

This preliminary study for developing a system for English-speaking practices investigated the role of a subjective difficulty rating through a questionnaire, using the vocabulary list "the New JACET8000." We conducted a self-reflective feedback questionnaire among 72 Japanese university students on the difficulty level of 75 English verbs, drawn from five different difficulty levels.

The results suggest that: (1) the words in Levels 4 and higher should be targeted in student practice; (2) Level 3 words should be selected for practice based upon the individual student; (3) all Level 4- and Level 5-words for which the SD scores in response variance were lower than 1.0 should be considered difficult words; and (4) individual students' subjective difficulty ratings should play a role when selecting their target words in the English-speaking practice system.

This is a continuous study, for which we aim to target more students and words in the future, to investigate these issues further. We also hope to assess how the detailed differences in the difficulty levels of individual words could change by conducting a questionnaire both before and after students practice using these words, and analyzing the results of the two questionnaires.

References

Blake, R. J. & Shiri, S. (2012). Online Arabic language learning: What happens after? *L2 Journal*, *4*(2), 230–246.

Japan Ministry of Education, Culture, Sports, Science and Technology. (2015). *Seito no eigoryoku kojyo suishin plan* [The English proficiency promotion plan]. Retrieved from https://www.mext.go.jp/a_menu/kokusai/gaikokugo/__icsFiles/afieldfile/2015/07/21/1358906_01_1.pdf

Janulevičienė, V., & Kavaliauskienė, G. (2007). Self-assessment of vocabulary and relevant language skills for evaluation purposes. *Santalka*, *15*(4), 10–15. http://www.cpe.vgtu.lt/index.php/cpe/issue/view/13.html

Jehma, H. (2020). Second Life: A three-dimensional virtual world for developing Thai EFL learners' English communication skills. *Turkish Online Journal of Distance Education*, *21*(Special Issue-IODL), 49–60. https://dergipark.org.tr/en/pub/tojde/issue/56079

Kashiwagi, H., Kang, M., & Ohtsuki, K. (2018, March). *A basic study on the conformity of Japanese university students in language communication activities*. Paper presented at The Asian Conference on Education & International Development 2018, Kobe, Japan. https://papers.iafor.org/submission40071/

Kashiwagi, H., Kang, M., & Ohtsuki, K. (2020, September). *Role of a subjective difficulty rating in using a system for practicing English speaking*. Paper presented at the Barcelona Conference on Education 2020, Online from Barcelona, Spain. https://papers.iafor.org/submission58730/

McCrocklin, S. M. (2016). Pronunciation learner autonomy: The potential of Automatic Speech Recognition. *System*, *57*, 25–42.

Mochizuki, M., Ishikawa, S., Sugimori, N., Murata, M., Uemura, T., Shimizu, S., & Kanamaru, T. (2016). The New JACET8000: On Compiling a Vocabulary List for English Language Learners at Japanese Universities. In The Basic Word List Revision Special Committee of JACET (Eds.), *The New JACET List of 8000 Basic Words* (70–90). Tokyo: Kirihara Shoten.

Suleimenova, Z. (2013). Speaking anxiety in a foreign language classroom in Kazakhstan. *Procedia - Social and Behavioral Sciences*, *93*, 1860–1868.

The Basic Word List Revision Special Committee of JACET (The Japan Association of College English Teachers). (2016). *The New JACET List of 8000 Basic Words*. Tokyo, Japan: Kirihara Shoten.

Todd, R. W. (2002). Using self-assessment for evaluation, *English Teaching Forum*, *40*(1), 16–19.

Wang, C., Song, H., Xia, F., & Yan, Q. (2009). Integrating Second Life into an EFL program: Students' perspectives. *Journal of Educational Technology Development and Exchange*, *2*(1), 1-16.

Woodrow, L. (2006). Anxiety and speaking English as a second language. *The RELC Journal*, *37*(3), 308–328.

Contact email: kasiwagi@kobe-u.ac.jp