

CLIL の効果検証

— 国内大学英語学習者の受動的語彙知識は向上するのか —

藤井 里美

Assessing the effects of CLIL on Japanese EFL university students' receptive vocabulary knowledge

FUJII Satomi

要旨

本稿では、国内大学で内容言語統合型学習（CLIL）を取り入れた英語授業を受講した日本人学習者（22 名、CEFR A2 レベル）の語彙学習過程を調査した。調査対象授業における目標語彙を語彙テスト（事前及び事後テスト）に含め、学習者に各単語の定義を記入させた。事前テストにおいて理解度の低かった単語に焦点を絞り、授業前後のテスト回答を調べたところ、授業内容や学習者の習熟度とスコアの増加・減少傾向との間には明確な関連性がなかった。単語によって学習者のパフォーマンスには異なる傾向が見られ、多くの先行研究で述べられている「CLIL は語彙習得に効果的である」という主張は、冷静に解釈する必要があると言える。CLIL の効果を正確に理解するためには、試験方法や学習者が CLIL 授業に触れる期間などを更に検討し、授業教材や授業中の発話内容なども調査する必要がある。

キーワード：receptive vocabulary, classroom contexts, EFL, CLIL, assessments

Introduction

Since the mid-1990s, a form of education called content and language integrated learning (CLIL) has spread especially in Europe and later in other continents. It is “a dual-focused educational approach in which an additional language is used for the learning and teaching of both content *and* language” (Coyle, Hood, & Marsh, 2010, p. 1). CLIL specialists have also emphasized the importance of incorporating the following in conducting lessons: the 4Cs of

CLIL (i.e., content, communication, cognition and community; Coyle *et al.*, 2010), 10 principles of CLIL (e.g., Multimodal input and Many tasks; Ikeda, 2015), 10 features that characterize CLIL (e.g., Enhancing peer communication and Supporting thinking skills; Ball, Kelly, & Clegg, 2015), to name but a few. Recently, the effects of CLIL have been mainly investigated by applied linguists (not subject teachers or their trainers, for instance), and a lot of these studies have looked into young learners studying in Europe. The present paper focuses on previous studies on CLIL learners' vocabulary learning processes in classroom contexts and investigates how older students in Japan perform in L2 vocabulary tests before and after attending CLIL lessons.

1. Previous research

Previous studies have explored various ways to measure CLIL learners' productive and/or receptive vocabulary knowledge using pre- and post-test (hereafter PreT and PostT, respectively) design. A number of studies have investigated elementary or secondary school students' performances in EU countries, such as Austria (Gierlinger & Wagner, 2016), the Czech Republic (Reynaert, 2019), Spain (Heras & Lasagabaster, 2015), and Cyprus (Xanthou, 2011a, 2011b). Recently, it is also possible to see studies conducted in other research sites, for example, an elementary school (10-12 year olds) in Taiwan (Huang, 2020) and an intensive preparation course for language minority high school graduates entering Canadian universities (Crossman, 2018). The contents or themes covered in these CLIL lessons include, for instance, History or Civics (Reynaert, 2019), Geography (Xanthou, 2011a), Science (Huang, 2020; Xanthou, 2011b), PE (Heras & Lasagabaster, 2015), Chemistry and History or Geography (Gierlinger & Wagner, 2016), urban development (Crossman, 2018; in this case, an umbrella theme of the program), whose target language was all English. The instrument(s) used in each study and its focus are listed below.

- 1) Heras & Lasagabaster (2015): PreT, PostT, and delayed PostT to check the students' productive skills as well as comprehension. For the former, students wrote L2 words of given pictures, then completed cloze tests where words were missing. For the latter, they matched pictures with L2 terms, matched definitions with L2 terms, then chose from a box an L2 term for each picture.
- 2) Reynaert (2019): PreT and PostT using Laufer & Nation's (1999) Vocabulary Levels Test to see how learners' general productive vocabulary size increases over a year. The students completed each sentence, e.g., "I'm glad we had this opp ___ to talk" (Nation, 2001, p. 425).
- 3) Huang (2020): PreT and PostT to examine learners' science knowledge. The students matched L1 and L2 terms selecting the appropriate ones from the choices, circled pictures

- (e.g., of the sun) and/or terms (e.g., 土壤/soil) of things plants need in order to grow, produced terms either in L1 or L2 that are parts of a plant, then completed True-False questions where things plants need or don't need were provided in both languages (e.g., music [音楽]).
- 4) Crossman (2018): PreT and PostT to check the learners' vocabulary gains during a course. It used two tests, the *Gates MacGinitie Reading Test* (MacGinitie & MacGinitie, 1992) second Canadian Edition to check learners' receptive vocabulary and reading comprehension, and an in-house developed test called the *Productive Academic Word Test* to measure 114 academic lexical items. For the former, students read 48 sentences, such as "He will **spoil** it", then select the appropriate synonyms of the bolded words from the choices. For the latter, the students "read the question [e.g., The politician listed all the ____ made during his time in office, and asked the crowd for their support in the coming election.], choose the semantically appropriate word from the nine possibilities and then use its correct inflection and derivational morphology to match the syntax of the sentence. Both lexical choice and inflection must be correct" (Crossman, 2018, p. 570). In addition, the researchers asked the participants to complete writing tasks whose topic was loosely connected to the course theme to see their lexical performances.
 - 5) Xanthou's two studies (2011a, 2011b): PreT and PostT to see the impact of CLIL lessons on L2 vocabulary development. The learners provided the L1 counterparts of 100 L2 words, which were related to the content words of the course subject.
 - 6) Gierlinger & Wagner (2016): PreT and PostT to check the learners' vocabulary size before and after exposure to CLIL teaching. The students took a vocabulary size test *X-Lex The Swansea Levels Test* (Meara & Milton, 2003), on which they "rate 120 English words from several vocabulary frequency bands as either known or not known (including nonce-words as distractors)" (Gierlinger & Wagner, 2016, p. 46). It measures receptive vocabulary proficiency, and the test score reflects a learner's vocabulary breadth.

Of the studies listed above, the only study that reported negative outcomes of CLIL lessons is Gierlinger & Wagner (2016), and the rest, generally positive. In Gierlinger & Wagner (2016), both the CLIL and control groups' vocabulary growths were seen and the former outperformed the latter (i.e., the regular English class learners) "in terms of absolute test scores, [but] the relative gain of the control group [exceeded] the CLIL pupils by far" (p. 49). Furthermore, the control group outperformed the CLIL group in terms of overall receptive vocabulary growth. Gierlinger & Wagner (2016) pointed out that the CLIL instructor in this study mainly used the 1,000 most frequent words in class, and this is probably the reason why significant vocabulary

growth was observed only within this band in the CLIL group. Contrarily, Heras & Lasagabaster (2015) found that “the CLIL module had a similar positive effect on both male and female students’ learning of the technical content-related vocabulary” (p. 70). Reynaert’s (2019) findings suggested that previous experience with CLIL led to better test scores and that “significant changes in general productive vocabulary increase are observable after two years of experiencing CLIL education” (p. 158). Huang (2020) found that “CLIL helped enrich student vocabulary size, foster student science knowledge, and lead them to expect to learn other subjects in English” (p. 1). In Crossman (2018), the learners’ receptive, productive, and academic vocabulary ability improved, and “those who displayed better and more academic vocabulary tended to transition more smoothly into university and attain higher GPAs” (p. 574). In both Xanthou’s studies (2011a, 2011b), it was found that CLIL students outperformed non-CLIL control groups.

One of the aspects that lacks in these previous studies, however, is that they do not explain how the participants performed on each question, especially how those with lower overall scores did in the studies. Agustín Llach (2017), so far the only study of which I am aware that does explain it, lists CLIL learners’ actual answers, lists CLIL learners’ actual answers, but it does not look closely into their in-class learning processes. Gierlinger & Wagner (2016) conclude that “overly optimistic expectations regarding the linguistic impact of CLIL in a mixed-ability setting guided by a predominantly implicit language teaching approach need to be re-evaluated critically” (p. 37). In addition, as pointed out by some CLIL researchers, such as Agustín Llach (2017) and Crossman (2018), many of the CLIL studies have focused on young and low proficient learners, and how older and more proficient learners perform has not been sufficiently investigated.

2. Research questions

Taking into account the limitations of the previous studies, the present paper aims to investigate the following two points: 1) How do CLIL lessons affect university students’ comprehension of lexical items used in class?; and 2) Are there any differences in receptive vocabulary test performances between high proficient and low proficient learners?

3. Research methods

The participants were 22 freshmen attending a private university in Tokyo. They took a placement test, i.e., the reading and listening sections of the Test of English for Academic Purposes (TEAP), immediately after their entry into the university and then were placed in an elementary-level English class. They also took the TEAP toward the end of the academic year. Their total score average was 88.25 (SD=0.89, low-high: 87.00-89.00) in the beginning, and

later 86.30 (SD = 12.89, low-high: 62.00-110.00), out of 200. This is similar to CEFR A2 level. The participants majored in either Science or Economics. They mainly interacted in their L1 when talking with their peers, but switched to the target language when working on in-class tasks (e.g., pair-works) and talking to the instructor. The instructor solely used L2 in class and incorporated the aforementioned 4Cs of CLIL, 10 principles of CLIL and 10 features that characterize CLIL.

In this one-year course, the students attended an English for Academic Purposes (EAP) course in the first semester (April to July), then a soft CLIL course in the second semester (October to January). In soft CLIL, the lesson aim is geared towards study of language rather than of a specific content or subject, and several different topics can be taught in L2 (Ikeda, 2011, p. 10). In this paper, the students' learning processes in the second semester are introduced. The class schedule is shown below (Table 1). The overall topic of the semester was *People and Technology*; hence, the subtopics dealt with the influence of technology on our daily lives. These subtopics were selected from a textbook *Contemporary Topics 1* (Solorzano & Frazier, 2009) which came with the supplementary materials (a CD and a DVD). In class, the students worked on academic listening and note-taking activities, then speaking and discussion activities about each topic. They also worked on two presentation tasks, the first one with a partner, comparing a book with its movie-version, and the second one individually, explaining the effects of technology of their own choosing (e.g., drones or 3D printers) on our daily lives.

Table 1. Class Schedule and Research Schedule

Dates	Class schedule	Research schedule
July 2015		- Consent form
October	- History (Trip to Antarctica) - Biology (GM food)	- Pre-session test (Oct. 28)
November	- Presentation 1 (Pair-work) - Media (SNS)	
December	- Technology (Robots)	- Post-session test (Dec. 4)
January 2016	- Presentation 2 (Individually)	- TEAP (Jan. 10, 16 or 23)

The research schedule is given in the far right column of Table 1. Approximately three months before the start of the second semester, the students received the instructions on the present study. Those who agreed to participate in the study (22 out of 24 students) signed the consent form. After the summer break, the students took two different vocabulary tests: a pre-session test and a post-session test (hereafter PreT and PostT, respectively), on which the students wrote the meaning of each L2 term either in English or Japanese. (Mostly, the students wrote an L1

counterpart of the L2 term on these tests.) The two tests were given unannounced. Ninety-one words, selected by the instructor as the course's target items before the onset of the semester, were tested. The details of these tests are explained in Fujii (2021). In grading the tests, one point was given to the correct answer, and zero to the incorrect one. The following were considered incorrect: blanks, wrong meanings/translations, or correct meanings but wrong parts of speech/verb tenses.

4. Results

The overall test results can be found in Fujii (2021), which showed that the students answered about 50 percent of the words correctly on each test. In this paper, the learners' test performances on 12 words are introduced. These words were selected, as each word's item facility (IF) value (i.e., the percentage of students who correctly answered each item) on PreT was below 0.15, which implies that they are suited to see the effects of the present CLIL course on the students' vocabulary comprehension. The words are shown in Table 2, along with the PreT and PostT IF values of each term.

The highest increase between PreT and PostT IF values observed was 0.64 (*BCE*), the lowest, 0.05 (*hand-eye coordination*). A decrease of 0.05 was also observed (*paraphrase*, *Czechoslovakian* and *robotics*).

Table 2. The Students' PreT and PostT IF Values on the 12 Terms

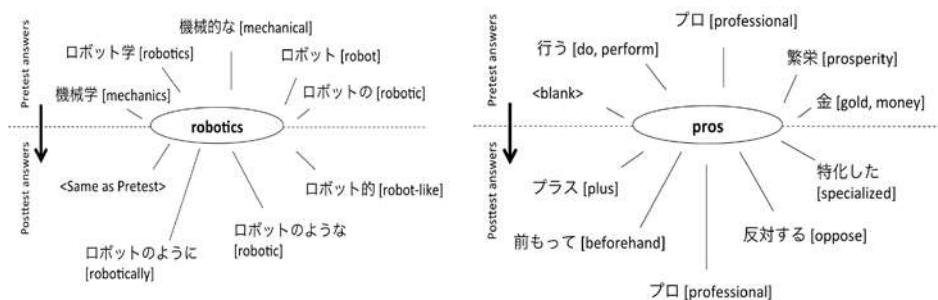
	<u>PreT</u>	<u>PostT</u>		<u>PreT</u>	<u>PostT</u>
<i>hand-eye coordination</i>	0.00	0.05	<i>Czechoslovakian</i>	0.10	0.05
<i>bored</i>	0.00	0.09	<i>dull</i>	0.10	0.45
<i>BCE</i>	0.00	0.64	<i>robotics</i>	0.14	0.09
<i>critics</i>	0.05	0.18	<i>pros</i>	0.14	0.27
<i>cons</i>	0.05	0.23	<i>attention spans</i>	0.14	0.32
<i>paraphrase</i>	0.10	0.05	<i>nuclear</i>	0.14	0.41

Note. The mean scores (standard deviations) of PreT and PostT were 0.08 (0.06) and 0.23 (0.18), respectively. A paired t-test was conducted, but no significant difference in the mean scores was observed.

The effect of the lessons was observed on *BCE*, but it was not so straightforward for the rest of the words. Firstly, the learners' performances on the words with low IF values showed various tendencies. All except one student left *BCE* blank on PreT, and more than half of the students wrote the correct L1 counterpart on PostT. On the other hand, PreT IF value of *bored*, for instance, was as low as that of *BCE*, but 16 learners wrote on PreT an L1 counterpart of *boring*

or *boredom*, which continued until the day of PostT. Many of the students seemed to have partial knowledge of the base word's and/or the suffix's meaning(s), but the lessons did not raise their awareness of its inflectional affix's precise meaning. Similarly, PreT answers for *robotics* and *nuclear* showed that 18 and 17 learners, respectively, were capable of showing their partial knowledge of the base words in L1 (e.g., ロボット [robot], 核 [nucleus]), but two and nine learners, respectively, wrote the correct answers on PostT. Furthermore, for *robotics*, there were those who answered correctly on PreT, then wrote either ロボットのように (robotically) or ロボット的な (robot-like) on PostT. Figure 1 illustrates how the participants' learning took place before and after the lessons. In class, *robotics* was used in the textbook's Technology (Robots) section as: "Scientists in the field of robotics develop robots to do many different kinds of work." (The students' answers were written in Japanese. Each answer's translation was selected by the author. The same applies to the rest of the answers introduced below.)

Figure 1. The students' PreT and PostT answers on *robotics* and *pros*



Secondly, performances on compounds did not improve. The compounds *attention spans* and *hand-eye coordination* were included in the textbook and the DVD, followed by detailed explanations by the instructor when talking about the positive and negative influences of videogames on children. The learners also discussed other influences in groups, then shared results with the class. PostT answers revealed, however, that the learners had been unsuccessful in connecting each form with its appropriate meaning. For instance, for *hand-eye coordination*, some answers were 身ぶり手ぶりで (using gestures) and 手話 (sign language). The learners' answers also seem to show that they are not connecting each compound with its concept, but instead coming up with a translation for each term found in a compound (e.g., answering only 組み合わせ [combination] for *hand-eye coordination*), which may imply that they are not perceiving these terms as set units.

Thirdly, words included in an in-class activity name did not have strong impact on the

learners' comprehension. PostT answers on *pros* and *cons* were not so successful even though the instructor had written an activity name "Debate (Pros and cons)" on the board, and all the students had participated in a debate as members of either a pros or cons group. However, one improvement observed was that, for *pros*, some of PreT answers that have no connection with the term (i.e., 行う [do, perform] and 金 [gold, money]) disappeared from the students' PostT answers. (See Figure 1.)

The students' actual answers are shown in Appendix. Due to space restrictions, 11 out of 22 students' answers are listed. Student 1 was the top scorer of the TEAP in January, and Student 20, the bottom (110 and 62, respectively). Students 1 to 5 are the first to the fifth high-achievers and 15 to 20, low-achievers. (Two students out of 22 either could not take the test or did not complete the test.) Looking at the two groups' performances, more precise and/or correct answers can be evidenced in the former. This is especially so for the words *pros* and *cons*. Contrarily, the latter members opted to leave more answers blank than the former and seem unsuccessful in retrieving information related to the lessons (especially for *dull*, *attention spans*, *critics*, *cons* and *hand-eye coordination*).

5. Discussion and pedagogical implications

The results showed that neither classroom activities nor learners' language proficiencies were strong factors that affected the CLIL learners' receptive knowledge of the target words (with low PreT IF values, in the present study). Depending on the tested terms, patterns observed in the learners' performances differed. Why were the performances on *BCE* exceptionally successful? Some of the possible reasons are positive learning experiences in class and/or the term's being in sync with the students' needs. As to other words with small improvements, perhaps the students a) did not know the appropriate L1 counterparts of the tested terms (e.g., 注意持続時間 for *attention spans*, and 視覚と手の協調 for *hand-eye coordination*) as they had had insufficient exposure to the L1 terms and/or concepts in their L1 environments, and b) paid more attention to partial aspects of the terms' base words, not the affixes (e.g., writing the meaning of *boring* for the tested word *bored*), to be able to keep up with the CLIL lessons' higher cognitive loads. In Gablasova (2014), for instance, when asked to orally define subject-specific vocabulary in their L1, CLIL learners "appeared unaware of the correct disciplinary terms and chose more general words for expressing their knowledge" (p. 161). She believes this is caused by "lexical development which [includes, for instance,] the deepening knowledge of words' meanings" (p. 161). To learn affixes, Thorndike (1941) recommends that learners should go through a list of words that include the same affix. The introduction to such learning strategies is recommended when learners reach a low-intermediate level (Nation, 2001, p. 267). If so, the poor outcome

of the *boring-bored* distinction in this study, for instance, is likely to be due to an insufficient amount of its input/encounters and/or the learners' developmental stages at that time. Hence, one possible way to overcome this problem is to let these learners take CLIL lessons for a longer period. Additionally, the non-contextualized test format used in the study may not have been appropriate to elicit different word knowledge aspects these learners had acquired in class. Furthermore, some test takers wrote a term's antonym (反対する [oppose] for *pros*) or opted to write definitions unrelated to the class content for polysemous words (プロ [professional] for *pros*). Thus, giving partial credit is favorable if the teaching goal is to create comfortable learning environments, especially for those who do not like CLIL classes.

In many of the previous studies, CLIL students' vocabulary test scores tended to improve after the lessons; yet, it was not simply so in the present study in regard to how well the participants comprehended each word used in class. Compared to language lessons that concentrate on one or two specific skills (e.g., speaking and listening course), it is likely that CLIL lessons give language learners a more holistic language learning environment; yet, it is still unclear whether this approach is suited to effectively develop learners' knowledge of target lexical items in classroom contexts. This may cause teachers who are inexperienced in conducting CLIL lessons and/or without sufficient background knowledge of CLIL to feel discontented in how things turn out in class on a daily basis. Likewise, learners may feel in a similar way, as there are cases where their skills do not seem to improve, which is something that especially low-achievers have claimed in other studies (e.g., Yang, 2015). One participant in this study wrote complaints on the end-of-the-semester questionnaire administered by the institution saying s/he was disappointed about the course because it did not cover TOEIC materials. Furthermore, Student 19, a science major, for instance, seemed dissatisfied with the surface-level scientific information covered in the course. The gap between the institutional objectives and the students' own goals may have caused, for instance, some participants' demotivated attitudes in class, low vocabulary test scores and poor performances on the TEAP in January.

6. Limitations and future research

In future studies, the following should be considered: Will the learners be able to a) provide more precise meanings if the terms appeared in context on PreT and PostT?, b) understand the terms' precise meanings in class if they encountered more input with the target words included?, and c) express their knowledge more successfully if they answered vocabulary test questions using a different language? A meaning recall test was used in this study as it is said to assess learners' receptive vocabulary knowledge more accurately than other measurement formats (e.g., Nakata, Tamura, & Aubrey, 2020); however, it is worth a trial to see how Japanese CLIL

learners perform using other formats. In addition, numbers of target word occurrences in the present course should be checked, along with how the instructor and students produced each term in their interactions. In terms of Point c above, Gablasova (2014) explains there are four ways for bilingually educated learners to express subject knowledge on tests: using L2 (the language of instruction), using L1, translanguaging (a mixture of both languages) and testing in both of the bilingual's languages. If L2 is used, learners will only be tested what they have been taught. Using L1, as in the present study, can be troublesome for learners as they may feel some difficulty "retrieving the knowledge in their non-instructional language or [lack] essential subject-specific vocabulary" (Gablasova, 2014, p. 152). The third is to allow them to use a mixture of both languages so that the test takers can show their content knowledge. The last is to have learners take two exams, one in L1 and the other in L2, testing the same material. In the present study, the first three cases were observed (answers that used L2, L1 and/or translanguaging); hence, the last option is worth consideration.

Using different research methods, it is also worth looking into the following: a) learners' L2 grammatical knowledge before the intervention, b) effects of CLIL on high-proficient (above A2 level) learners' vocabulary knowledge, and c) effects of CLIL on students' learning processes of base words and other lexical aspects.

Lastly, it is important to note that the present study solely looked at the learners' performances on the target words with low PreT IF values. Their performances on the terms with higher PreT IF values should also be investigated.

7. Conclusion

In this study, Japanese EFL university students ($n = 22$, CEFR A2 level) attending a semester-long CLIL course were recruited, and their L2 vocabulary learning processes were observed using a pre- and post-test design. It focused on their receptive knowledge of the target lexical items with low PreT IF values to see the effects of the intervention. The results showed that the learners' performances varied depending on the tested terms, whose patterns were not clearly affected by CLIL class activities and the students' language proficiencies. To further understand the relationship between these test results and the lessons, it is necessary to look into transcriptions and lesson materials to find out, for instance, spoken and written frequency of each term in the lessons. Many of the previous studies on CLIL learners' vocabulary learning processes in classroom contexts have reported positive effects of this educational approach; however, upon closer investigation on each target item, it is recommended that practitioners and researchers cautiously interpret these previous findings. To observe the effects of this approach, it is also crucial to think through effective ways to assess students' vocabulary learning processes (especially

for low proficient learners) and the duration of CLIL lesson provisions to these learners.

Acknowledgement and funding

I am grateful to the anonymous reviewers for the invaluable advice on an earlier draft. This work was supported by JSPS KAKENHI Grant Number JP21K13059.

References

- Agustín Llach, M. P. (2017). The effects of the CLIL approach in young foreign language learners' lexical profiles. *International Journal of Bilingual Education and Bilingualism*, 20(5), 557-573.
- Fujii, S. (2021). Vocabulary learning processes of Japanese EFL university students in a content and language integrated learning (CLIL) course. *The Journal of the Institute for Language and Education Research*, 38, 167-182.
- Ball, P., Kelly, K., & Clegg, J. (2015). Putting CLIL into practice. Oxford: Oxford University Press.
- Coyle, D., Hood, P., & Marsh, D. (2010). CLIL: Content and language integrated learning. Cambridge: Cambridge University Press.
- Crossman, K. (2018). Immersed in academic English: Vocabulary and academic outcomes of a CLIL university preparation course. *International Journal of Bilingual Education and Bilingualism*, 21(5), 564-577.
- Gablasova, D. (2014). Issues in the assessment of bilingually educated students: Expressing subject knowledge through L1 and L2. *The Language Learning Journal*, 42(2), 151-164.
- Gierlinger, E., & Wagner, T. (2016). The more the merrier – Revisiting CLIL-based vocabulary growth in secondary education. *LACLIL*, 9(1), 37-63.
- Heras, A., & Lasagabaster, D. (2015). The impact of CLIL on affective factors and vocabulary learning. *Language Teaching Research*, 19(1), 70-88.
- Huang, Y. (2020). The effects of elementary students' science learning in CLIL. *English Language Teaching*, 13(2), 1-15.
- Ikeda, M. (2011). Basic principles of CLIL. In Y. Watanabe, M. Ikeda & S. Izumi (Eds.), *CLIL (Content and language integrated learning): New challenges in foreign language education at Sophia University, Volume 1: Principles and methodologies* (pp. 1-13). Tokyo: Sophia University Press.
- Ikeda, M. (2015). グローバル・リーダーの素養を伸ばす CLIL 型授業 [CLIL lessons that develop global leaders' knowledge]. *英語展望*, 122, 22-28.
- Laufer, B., & Nation, P. (1999). A vocabulary-size test of controlled productive ability. *Language Testing*, 16(1), 33-51.
- MacGinitie, W., & MacGinitie, R. (1992). *Gates MacGinitie Reading Tests*. 2nd ed. Toronto, ON: Nelson Education.
- Meara, P., & Milton, J. (2003). *X-Lex, the Swansea Levels Test*. Newbury: Express.
- Nakata, T., Tamura, Y., & Aubrey, S. (2020). Examining the validity of the LexTALE test for Japanese college students. *The Journal of Asia TEFL*, 17(2), 335-348.
- Nation, I. S. P. (2001). *Learning vocabulary in another language*. Cambridge: Cambridge University Press.

- Reynaert, B. (2019). CLIL and productive vocabulary acquisition in the Czech context. *Advances in Language and Literacy Studies*, 10(4), 153-160.
- Solorzano, H., & Frazier, L. (2009). *Contemporary topics 1: Academic listening and note-taking skill*. New York: Pearson Education.
- Thorndike, E. L. (1941). *The teaching of English suffixes*. Columbia University: Teachers College.
- Xanthou, M. (2011a). Current trends in L2 vocabulary learning and instruction: Is CLIL the right approach? *Advances in Research on Language Acquisition and Teaching: Selected Papers*, Athens, 459–471.
- Xanthou, M. (2011b). The impact of CLIL on L2 vocabulary development and content knowledge. *English Teaching: Practice and Critique*, 10(4), 116–126.
- Yang, W. (2015). Content and language integrated learning next in Asia: Evidence of learners' achievement in CLIL education from a Taiwan tertiary degree programme. *International Journal of Bilingual Education and Bilingualism*, 18(4), 361-382.

Appendix The Students' PreT and PostT Answers on the Target Words - Words with Low PreT IF Values -

	BCE		dull	robotics		Czechoslovakian		bored	pros	
	PreT	PostT		PreT	PostT	PreT	PostT		PreT	PostT
1	-	-	面倒な [troublesome]	たいくつ [boredom]	機械的な [mechanical]	ロボットの	チェコスロ ベニア [*Czechoslo venia]	たいくつな [boring]	-	賛成する [agree]
2	-	-	疑い [n. doubt]	頭の鈍い [dull-witted], 単調な	機械的な	機械的な	-	たいくつな	-	賛成する
3	-	-	退屈な	退屈な	ロボットの [adj. robotic]	ロボット工学	-	退屈な	-	賛成 [approval]
4	-	-	-	やる気のない [not motivated]	機械学 [mechanics]	機械学	シゼコスロ バキアン [*Cizechoslo vakian]	つまらない [boring]	プロの [adj. professional]	前もって [beforehand]
5	-	-	やつ [guy]	単純な [simple]	ロボット [n. robot]	-	チェコスロ バキア [Czechoslov akia]	あきた	プロ [n. professional]	賛成
15	-	-	-	-	機械的	機械的な	-	退屈させる [v. bore]	～に飽きる [get tired of...]	参成 (*unable to translate)
16	-	-	-	-	機械的	ロボット工学	-	退屈にされる [seem bored at...]	退屈する [be bored]	-
17	-	-	人形 [doll]	要旨 [the main points]	ロボット	ロボット	-	じっとみる [v. stare]	プロ	プラス [n/v. plus]
18	-	-	-	-	ロボット	ロボット	スロバキア人 [n. Slovak]	たいくつな [boring]	プロアエツ ショナルな [adj. professional]	-
19	-	-	-	たいくつな	ロボット	ロボット	-	たいくつ	-	-
20	役員 [n. official]	-	汚い [dirty]	ロボット	機械的	きかいてきな [mechanical]	チェコスロ バキアの [n. Slovak]	乗り上げる [run onto]	行う [do]	過程 [n. process]

	paraphrase		attention spans		critics		cons		nuclear		hand-eye coordination	
	PreT	PostT	PreT	PostT	PreT	PostT	PreT	PostT	PreT	PostT	PreT	PostT
1	段落 [paragraph]	-	-	意識のある期間	根本的な [basic]	危機 [crisis]	-	反対	核 [nucleus]	核	-	-
2	言い替え表現 in other words	言い替え表現、特別警戒 期間	[full-alert period]	-	非難 [criticize]	批難する	-	反対、 object	核の	核の	communicate with someone by the hand	-
3	段落	段落	-	集中している時間	-	評論家	カン [n. can]	反対	核	核	-	目と手の 相互反応
4	パラフレーズ	段落	出席頻度 [attendance rate]	出席状況 [attendance records]	-	殺傷する [kill and wound]	-	ともに [with]	核	核	-	身ぶり 手ぶりで [with gesture]
5	-	-	注意する間	注意期間	批判	批判	-	-	核	核の	-	-
15	-	-	-	-	-	-	反対	核	原子力 [nuclear power]	核	-	-
16	文 [sentence]	口葉 [*unable to translate]	期間 [term, period]	期間に注意 [attention to the period]	-	-	-	-	粒子 [particle, grain]	粒子	手話 [sign language]	-
17	段落	段落	出席期間 [attendance period]	提出期間 [submission period]	市民 [citizen]	-	コーン、 とうもろこし [corn]	とうもろこし	核	核	-	-
18	-	重要な フレーズ [important phrase]	滞在期間 [length of stay]	-	ひ評家	批判家	反対の	-	核	核の	-	どう体しりよく [dynamic/ kinetic vision]
19	-	題 [n. title]	-	-	-	-	-	-	核	核	原子力	-
20	図 [figure, diagram]	言葉 [language]	注意期間	?? の期間 [period of ??]	いじめる [v. bully]	慣習 [custom, usage]	コーン	穀物 [grain]	核の	核の	-	目で見える [visible]

Note. The numbers (1 to 20) indicate the student numbers (Student 1, the top TEAP scorer, Student 20, the lowest). The correct answers are underlined. A dash (-) means a blank answer. The translations of the answers are shown in brackets. If a translation is not provided, see the earlier translated one. The ones with * could not be translated. The translations 賛成 and 反対 for *pros* and *cons*, respectively, are marked correct in this study. For clarity, parts of speech are provided for some answers: n. = noun, v. = verb, adj. = adjective.