

## Is L2 vocabulary learned better via context or via translation?

*Michael John Alroe, Hayo Reinders  
and Punchalee Wasanasomsithi*

### Abstract

*Various studies have shown intentional learning of L2 vocabulary to be more efficient than incidental learning from exposure to comprehensible input. Some have argued that such learning may be further enhanced by recourse to L1 translation, particularly for weaker learners. The present study aims to determine if intentional learning of new vocabulary through L1 does indeed confer an advantage over intentional learning from an L2 context. To this end, 403 Thai freshmen students were pre-tested on thirty vocabulary items set for study on their English course. They were then randomly allocated to either a translation or context group to learn those items. Time on task was controlled. A delayed post-test showed that while the translation group was better at matching the thirty English words with Thai translations, albeit marginally so, there was no benefit conferred on the translation group when it came to using the words in a contextual gap-filling exercise. This finding held for both advanced and weaker learners.*

---

### Affiliations

Michael John Alroe: Chulalongkorn University Language Institute, Bangkok, Thailand  
email: [malroe@hotmail.com](mailto:malroe@hotmail.com)

Hayo Reinders: Unitec, New Zealand  
email: [info@innovationinteaching.org](mailto:info@innovationinteaching.org)

Punchalee Wasanasomsithi: Chulalongkorn University Language Institute, Bangkok, Thailand  
email: [Punchalee.W@chula.ac.th](mailto:Punchalee.W@chula.ac.th)

KEYWORDS: VOCABULARY ACQUISITION; CODE-SWITCHING; CONTEXTUAL LEARNING; INTENTIONAL LEARNING; INCIDENTAL LEARNING

## Introduction

In the literature, it is commonly accepted that, as with ESL instruction, EFL classes should be communicative in approach and conducted essentially in English, even when the teacher and students share a common L1. Some authorities such as Macaro (2014) argue some 'code-switching' to L1 can be helpful, but it should be limited to less than 15% of a lesson. Nation (2003) agrees that L1 can have a 'small role', but says most lesson time should be devoted to meaning-focused input and output. However, those who have long experience in Asian schools and universities can attest that outside the international schools much of the foreign language instruction is given in L1. Research and official figures support this assertion. Macaro himself (2014:13) concedes that many relevant studies report percentages of L1 use in instruction to be near 50%, a level he suggests makes such classes non-communicative. Liu and colleagues (2004) collected data from thirteen Korean high schools and found that L1 was used by teachers an average of 68% of the time in English classes. Domalewska (2015), observing Thai elementary school teachers, found on average 25% of English lessons were conducted in Thai, but at times this could reach 50%. A Ministry of Education (MEXT) survey of Japanese teachers of English found the percentage who spoke mostly English in oral communication classes was under 30%, while for other English classes it was below 5% (Yamada and Hristoskova 2011). Thus, Lightbown and Spada (2013:195) conclude that despite the apparent ubiquity of the communicative approach, in 'the schools of the world, grammar translation is no doubt the most widely applied method' of instruction. Moreover, it should be pointed out that merely giving an L1 translation of an L2 word does not of itself ensure that the word will be learned.

In English-speaking countries where English is taught to foreign students, the language of instruction will normally be the target language (TL) because the students generally will not share a common L1 with each other or with the teacher. But when Duff and Polio (1990) examined foreign language classes for English-speaking students they found teacher practice that mirrored EFL classes overseas. The use of the TL in these classes varied from 100% to as low as 10%. When asked the reason they did not use the TL consistently, teachers said they feared their students would not comprehend their lessons if they taught entirely in the TL. Interestingly, 'students uniformly expressed satisfaction with the amount of English used

in their FL classes, although this could be as low as 10% (1990:162). This suggests students may cope with a much higher level of TL than teachers realise and indeed there are educational systems that mandate that all instruction in language classes be in L2. And if lessons are in the TL, then comprehensible input is likely to be maximised.

The issue is of particular relevance in an Asian EFL context because access to L2 input is often limited to the classroom. Moreover, the practice of instructing students in L1, often termed 'code-switching', further reduces the amount of comprehensible input students receive.<sup>1</sup> Lee and Macaro (2013:888) report that vocabulary instruction is the most frequent reason for switching to L1. So if it can be shown that vocabulary can be successfully learned through encountering it in context, learners may have more opportunities to be exposed to the TL.

Much research has focused on a slightly different but related issue, namely whether L2 vocabulary acquisition can be better taught by incidental means such as reading for pleasure, or alternatively by intentional learning such as gap-filling exercises, vocabulary lists and flashcards. The present consensus appears to be that both incidental and intentional learning can have a significant, beneficial impact on acquisition (Ellis 2015). Thus, Rieder (2003:250) notes that while most scholars agree that the less common English words are learned incidentally, the 'first few thousand most common words' may be learned intentionally. Elgort's (2011) investigations using priming show no discernible difference between incidentally acquired and intentionally acquired vocabulary. This is contrary to the belief that what is deliberately learned remains inaccessible to real-time interactions. If both intentional and incidental learning promote acquisition it would seem sensible to employ both. However, if instruction is given in L1, incidental learning from comprehensible input is, at least temporarily, excluded. As Webb (2007:64) noted: '... it seems likely that many aspects of vocabulary knowledge may be gained through learning in context ... (which) may provide a better chance of gaining vocabulary knowledge than decontextualised learning from translations.'<sup>2</sup>

Furthermore, the meaning of a word must ultimately be determined by its context. Polysemy dictates that an L1 or L2 dictionary can only offer a range of possible meanings for a word. A student must develop the capacity to determine the precise meaning of a word from its context, for its context will dictate which one of its possible meanings is appropriate for that context.

Thus, we investigate whether EFL students are able to learn new vocabulary as effectively through studying it in an L2 context as through translation to L1. If so, the case for using L2 in EFL instruction is strengthened, for

doing so maximises comprehensible input and the opportunities for incidental learning. To this end, this study examines vocabulary acquisition via intentional learning in the course of L2 instruction in an EFL context where students and teachers share the same L1. One group of students inferred and learned the meaning of new words from the context, and the other group studied the words via translations.

## Literature review

In general, results reported in the literature have tended to favour learning via translation rather than from context. But the results may be interpreted more as showing explicit learning to be more effective than implicit learning. Some have compared learning L2 words explicitly via translation against learning vocabulary incidentally from a context where target words may be glossed but not highlighted as targets for deliberate study (Choi, Kim and Ryu 2014; Laufer and Shmueli 1997). Explicit learning through translation was generally found to be more efficient. In 1996, Prince published a paper comparing two forms of explicit learning. One group studied via translation and the other from context. The translation group achieved superior results on an immediate post-test. That paper has been widely cited (Folse 2006; Groot 2000; Huckin and Coady 1999; Hulstijn 2008; Laufer and Girsai 2008; Nation 2001; Schmitt 2010 and Webb 2009) and was seen as supportive of learning via translation to L1.

Prince also focused on whether study via translation was particularly advantageous for weaker learners. This aspect of his research was relevant to the claims often made by teachers justifying teaching through L1 to weaker students. He found that less proficient students who studied via translation did indeed do significantly better than those who studied contextually when they were tested by translation, though not when tested contextually. However, the study was problematic in that it had no delayed post-test and a later replication of the experiment found no advantage for learning vocabulary via translation (Alroe and Reinders 2015). Moreover, in his conclusion Prince himself expressed reservations about the use of translation: 'It may indeed be plausibly suggested that a highly developed ability to learn words via translation links may in some cases be detrimental to the establishment of the skills and strategies required to handle discourse' (1996:486).

A later paper by Laufer and Girsai (2008) contrasted the amount of new vocabulary learned by a translation group with that learned by a context-only group and found that the participants who studied using translation achieved vastly better scores on a post-test than did the context group.

However, the experiment was not a simple comparison of the effectiveness of approaching new vocabulary through L1 as opposed to contextual means. The context group was largely restricted to incidental learning, which occurred during the reading of a text in order to answer and correct comprehension questions. By contrast, the translation group was tasked with translating the English text into Hebrew, the results of which were corrected by the instructor. They then translated an agreed Hebrew translation back into English – both tasks that would have entailed explicit attention to the vocabulary. Time on task for the translation group was double that for the context group. Furthermore, the translation group initially encountered the target vocabulary in an L2 context and had to infer the meanings from that context to enable them to translate them into Hebrew. Thus, the study can be seen more as determining the advantages of explicit learning. It does not consider whether EFL students benefit more from learning vocabulary from context rather than translation. However, the study is notable in the way it demonstrates how the collocations of an L2 word can be made salient to students by contrasting them with the collocations of the L1 translation of that word.

Other studies did focus more precisely on comparing the results of using L1 rather than L2 during explicit instruction. Tian and Macaro (2012) conducted a large study (with 117 Chinese university students) and found weaker students who were explicitly taught vocabulary with the teacher using ‘code-switching’ initially did better than others who were taught only through L2. However, the effect size was small and the advantage disappeared by the time of a post-test two weeks later. In a similar study to the present one, Choi, Kim and Ryu (2014) compared intentional learning of new vocabulary (pseudo-words) contextually with learning from a list with translations. They found that on a five-week delayed post-test the words learned via translation were better remembered when tested explicitly. But there was no difference in implicit learning as determined by ‘reaction times and event-related potentials’ (2014:227). Moreover, there was no contextual test as learning was tested by translation to L1. Further, the learning via context was accompanied by glosses in L1 which somewhat blurred differences in treatment between the two groups.

So, while there is evidence for the greater efficiency of a pedagogy that promotes intentional over incidental learning of vocabulary, the advantage of translation to L1 over contextual learning remains open to question. To complement this emerging picture, the present study compared intentional vocabulary learning from translation with intentional learning from context in a real classroom situation with large numbers of students and

equal time on task for both modes of study. Further, we delayed the post-test to find out if the learning endured.

## Methodology

The research question was whether non-beginner language students studying new vocabulary in context are as successful as those studying the same vocabulary under the same classroom conditions but from translation. A follow-up question was whether less proficient students benefit more from translation. In this study, 403 freshmen university students were given a pre-test and then randomly assigned to learn words either from context or from translation. Prince (2012) had found that placing target words in the context of a narrative framework produced better learning than embedding words in unconnected sentences. So in our experiment the context groups encountered the thirty words spread through four coherent texts that mimicked a news report, two conversations and an advertisement. The translation groups were given the words paired with Thai translations. Both conditions had equal time on task and as the context groups were instructed to study from the texts those words that were in bold and underlined and were told they would be later tested on what they had learned, both were engaged in a task promoting intentional learning. They were given an immediate post-test for students to check what they had learned and consolidate that learning.

For further consolidation, a week later they each had a further self-corrected test that matched their assigned learning method. In all three parts of the learning phase of the experiment, the context groups' activities were all in the L2, while the translation groups' activities involved relating the L2 words to L1 translations.

The post-test was timed to occur three weeks after the second consolidation task. The length of the delay accorded with Schmitt's advice that 'a delayed post-test of three weeks should be indicative of learning which is stable and durable' (2010:157). This post-test first repeated the original pre-test of thirty words and then further tested the students' ability to recall fifteen of the words from contextual clues.<sup>3</sup>

## Participants and context

This experimental study used a pre-test-treatment-post-test design with random allocation of participants into treatment and control groups. Participants were recruited from existing freshmen English classes from a major Thai University. The population of the study consisted of 4,870 first-year undergraduate students who were enrolled in the Experiential English

II course in the second semester of the academic year 2014/15. This course is an integrated skills course required of almost all first-year students of the university and involves three hours of instruction per week. The textbook set for the course (*World Pass*, Upper Intermediate, Stempleski *et al.* 2006) highlights certain vocabulary that appears in reading passages and which the authors recommend studying. That vocabulary constituted the set vocabulary for students to be tested on in the university examinations. Many Thai-speaking teachers would normally teach this via code-switching, while native English speakers and some Thai teachers use context and definitions.

The study sample comprised 576 students of whom 403 (70%) participated in all stages (i.e. pre-test, treatment and post-test) of the study. They were recruited by means of convenience sampling of twenty classes enrolled in the course (three of which were lost to the experiment because of unforeseen timetabling problems that arose during the course of the study). They were taught by a total of twelve instructors, all of whom held at least a Master's degree in language teaching or a related field. In order to include a broad range of backgrounds, participants were recruited from the Faculties of Engineering, Science, Sports Science, Allied Health Science, Commerce and Accountancy, Dentistry, Education, Political Science and Economics. There was a roughly equal split between male and female students, all ranging in age from 16 to 19 years old, with the majority being 18. The proficiency level of the students varied from beginner to upper-intermediate level. Testing of a range of 139 students from a previous year's intake had been done using the Nation productive test for the 2nd thousand most common words (Cobb 2016). It gave a mean score of 11.2/18 and an SD of 3.9. A score of 15/18 was considered to show a satisfactory productive knowledge of that range of vocabulary. The 2014/15 intake including our participants was considered to be similar in ability.

### Target language

The first-year English syllabus specifies a list of vocabulary items for study in each half of each semester. We chose the vocabulary list that was set for the second half of Semester 2 and used it for conducting the experiment in the first half of that semester, so the students knew the words they learned would eventually be relevant for their course but they would not encounter those words in their normal classes until the experiment was completed.

A pilot test was done with three freshmen classes to select the thirty target words for the experiment from the list of forty-three set for study in the course. Thirty words seemed a reasonable number for students to learn

in the time class teachers were prepared to spare us for the experiment. The results of the pilot study led us to eliminate eleven words such as ‘lifetime’ and ‘note’, which over 90% of the students in the pilot knew already and a couple (e.g. ‘remedy’ known by only 11%) deemed too difficult. The three classes in the pilot study were not involved in the experiment that followed. Thai translations of the thirty words were made and repeatedly checked until four experienced Thai teachers of English agreed that the translations were appropriate for the words as used in the context of the textbook set for the course. The thirty words are given below. (All thirty were tested in matching exercises pre- and post-test. The fifteen used in the gap-fill post-test are underlined.)

physiotherapy, placebo, predict, gratitude, symptom, basis, react, proposal, graduate, antibiotic, acupuncture, worthless, internship, access, spontaneous, constantly, entitled to, thoroughly, independently, free of charge, check-up, appreciate, apply, consideration, reference, rule out, attribute, candidate, circumstances, evidence

### Manipulation of target language

For the translation group, a sheet pairing the thirty English words with their Thai equivalents was created. The Thai translations were placed in columns on the left and the English targets in matching columns to the right as shown here:

Learn the meanings of the English words in Thai			
เป็นหลัก	Basis	การขอบคุณ	Gratitude

For the context group, the thirty words were incorporated into differing text types. These texts were composed to incorporate the words in simulations of a news report, two transcripts of conversations and one job advertisement. They approximated to Laufer and Shmueli’s (1997) ‘elaborated text’, where target vocabulary is embedded in a text with some ‘lexical elaboration’ that could enable a reader to infer the meaning of a word from its context. We included no glossing but the target words were set in bold and underlined. The texts were fitted onto two sides of an A4 page and illustrated as appropriate for the genre of the text. The four texts were prefaced with the instruction: ‘Read the 4 texts below and try to learn the meanings of the words in **bold**.’ An example of one of the texts is shown in Figure 1.

**3. News report****Pop Star in Jail**

Pop star Jayden was arrested by Hong Kong Police last night after a car accident. Police say they have **evidence** to show that the young star had been drinking and on that **basis** they will charge him with dangerous driving. They have **thoroughly** checked the place where the accident happened and are sure they know exactly what happened. They have **ruled out** car problems and say Jayden caused the accident.



The young star is upset because his new 800 thousand Hong Kong Dollar BMW is now **worthless**.

Jayden's mother says police should free the boy because of the **circumstances** of the accident. She says the rain was heavy and the roads unsafe when the crash happened.

Police say the star is **entitled** to a lawyer but he has not asked for one. WIN NEWS has tried to interview the star but police would not give them **access** to him.

**Figure 1:** Context group – third text.

### Pre-testing

The participants were given a pre-test with the English target words numbered 1 to 30 and the Thai translations in a separate textbox. Students answered by writing the number of the English word next to its Thai translation. They were allowed ten minutes, which seemed adequate for most students.

### Randomisation

The students were assigned to the context or the translation group using a set of random numbers. Study and testing materials were prepared with students' names written on them so there would be no confusion over which material a student should receive.

## Instructions and procedures

### First session

An experimenter visited the selected classes and was assisted by the class teacher to distribute the materials and instruct the students for this first and following sessions. For the first learning session (given one week after the pre-test) the students in the two groups were moved to different parts of the room to avoid them seeing the other group's materials. The context group was instructed to discern the meaning of the words from the texts and learn them, while the translation group was told to learn the meaning of the English words from the Thai translations. They were given twelve minutes, and careful observation by the researcher and class teacher found that the students focused conscientiously on the materials.

After the twelve minutes the students were given a test paper to check their own learning. A study by Karpicke and Roediger (2008) had shown repeated testing that pushed learners to recall vocabulary they had studied had a potent effect on consolidating long-term memory of newly learned words. So both groups were given different immediate consolidation exercises that matched their mode of learning, not to determine the amount of learning that had taken place but to help consolidate it.

The context group had to match the thirty targets with English definitions they were given. The translation group was given thirty Thai translations of the target items and had to write the matching English words beside them. The context group was given the target words in a textbox and thirty definitions with a space beside each. They were asked to write the words beside the appropriate definitions. The translation group was given only the thirty Thai translations and had to write the English words beside them from memory. Both groups were given ten minutes and were then asked to turn the sheet over to check the answers provided on the back. The sheets were collected but the results were not recorded. Many in the context group were unable to complete the self-testing in the time allocated, probably because what was required of them was more complex than what was required of the translation group.

### Second session

Again, following Karpicke and Roediger (2008), a second consolidation session was held a week after the first. The context group was given the transcripts of three short dialogues. Each dialogue had gaps for ten of the target words that were placed in textboxes beneath the dialogue. Students were asked to write the words in the gaps. Meanwhile, the translation group was given a crossword with the thirty Thai translations as the clues and the English target items as the answers. After ten minutes, students

checked their answers against those provided on the back of their sheets. They were thanked for their participation and reminded that they would not be tested on the words until the second half of the semester. As the students had different words set for the first half of the semester and the exam was imminent, it was considered unlikely they would concern themselves with thirty words not to be officially examined for another three months.

#### Post-test

The post-test was given three weeks after the second session and in the first week back from the mid-semester break. However, four of the seventeen classes had extra exams, which meant they had a four-week instead of a three-week gap. The post-test contained two parts. Part 1 repeated the pre-test where students had to match English and Thai words, but with the words in a different order. The maximum time for Part 1 was reduced to nine minutes and most students finished early. When a student indicated he or she was finished, the student was given values for the unknowns in an algebraic formula printed at the bottom of Part 1 and was asked to work out its value ('Before turning to the 2nd page, solve this maths problem. You will be given the values for a, b ... f when you finish the matching exercise:  $(a^2 + b^2 + c^3) - (d^2 + e^3 + f^4) = ?$ '). The maths problem (following Karpicke and Roediger 2008) was used to distract the students from the words they matched in the repeat of the pre-test, fifteen of which would be used in the second part of the post-test. When a student completed the algebra problem he or she turned immediately to the gap-filling exercise and was reminded not to turn back to the first part of the test. They were monitored to ensure they followed these directions.

The second part of the post-test consisted of fifteen gapped sentences and students were asked to insert the missing letters. For example:

There is some ev\_\_\_\_\_ that the drug is effective, but we need more before we can be sure.

The first letter was given for words of one syllable and the first syllable for multisyllabic words. A maximum of 18 minutes was allowed for the two parts of the test but almost all students finished early.

#### Analysis

As only one researcher was available for marking, the papers were marked (with the researcher unaware of whether the paper belonged to the translation or context group) and the marks recorded. Some days later, he again marked the papers and the second set of marks was compared with the first. Any discrepancies were resolved by re-checking the papers and any cases where the marker was in doubt were resolved by checking with one of the

other researchers. In the marking of the gapped post-test a full mark was given for phonetically correct misspellings (e.g. 'refference', 'attribut') and a half mark for reasonable approximations ('phisotherapy', 'independently').

Any students from the seventeen classes who had not completed all four parts of the experiment were deleted from the data set and statistical analysis was undertaken of the remaining 403. Scores for the pre-test and the two sections of the post-test were analysed using SPSS Statistics. First, statistical procedures were used to check that the context group and translation groups had similar knowledge of the target vocabulary items at the start of the study. Next, to determine whether participants had made any learning gains, the scores of all participants on the pre-test and the section of the post-test that repeated the pre-test were compared. Following that, the performances of the treatment and control groups were compared in the two parts of the post-test. Finally, we divided participants into advanced and weak groups, based on their scores in the pre-test, to establish whether their initial knowledge of the target vocabulary was a factor in their learning.

## Results

### Summary tables

**Table 1:** Descriptive statistics of the difference between pre- and post-test scores for the translation and context groups.

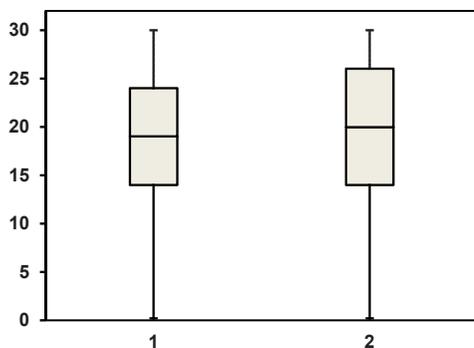
Groups	Number of participants	Pre-test – matching /30		Delayed post-test – matching /30		Delayed post-test – gap-fill /15	
		Mean	SD	Mean	SD	Mean	SD
Translation	198	19.5	7.2	26.4	5.7	11.1	3.4
Context	205	18.9	6.7	24.1	6.3	10.6	3.5

**Table 2:** Descriptive statistics comparing weak and advanced students – their groups determined by the pretest results.

Groups	Number of participants	Ability level	Pre-test – matching /30		Delayed post-test – matching /30		Delayed post-test – gap-fill /15	
			Mean	SD	Mean	SD	Mean	SD
Translation	110	Advanced	24.8	3.0	29.2	1.5	12.6	2.1
	88	Weak	12.8	4.9	22.9	7.0	9.3	3.8
Context	102	Advanced	24.4	3.3	27.9	3.2	12.3	2.4
	103	Weak	13.4	4.0	20.4	6.5	8.9	3.6

### Pre-test

The context group ( $n = 205$ ) and translation group ( $n = 198$ ) were not found to be significantly different on the pre-test in which participants had to match Thai translations with the thirty target English vocabulary, as Figure 2 illustrates.

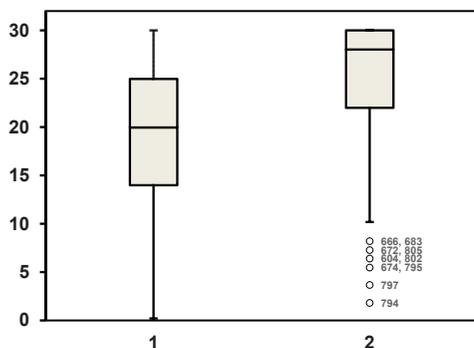


**Figure 2:** Pre-test – context (1); translation (2).

### Post-test – pre-test repeated

#### Within-participants comparison

First we determined whether the treatments had resulted in learning taking place, by comparing the pre-test scores for all students treated as one group (mean = 19.1, SD = 6.9) on the thirty matching items with the scores from the delayed post-test (mean = 25.2, SD = 6.1). Figure 3 illustrates the difference.

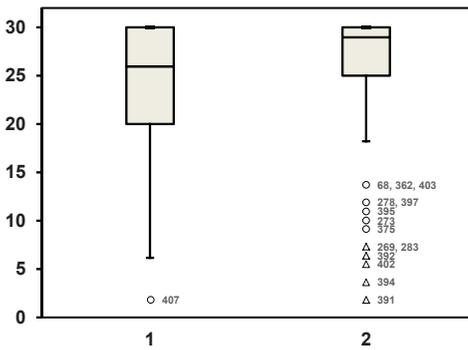


**Figure 3:** Pre-test/post-test comparison for all participants treated as one group.

A Wilcoxon signed rank test (used because examination of the histograms indicated that the distribution on the post-test might not be normal) revealed a statistically significant increase in matching L2 words with L1 translations over the course of the experiment,  $z = -16.27$ ,  $p < 0.0005$  with a large effect size ( $r = 0.57$ ). The median score on the matching test increased from 20 to 28 out of 30.

**Between-participants comparisons**

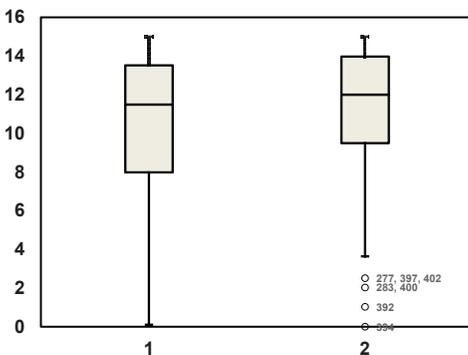
The combined results of the repeated pre-test that formed Part 1 of the post-test for the context and translation groups are shown in Figure 4.



**Figure 4:** Post-test (Part 1) scores – context (1)/translation (2) comparison.

**Post-test – gap-fill**

The results of the gap-fill test (maximum score = 15) are shown in Figure 5.



**Figure 5:** Gap-fill – context (1); translation (2).

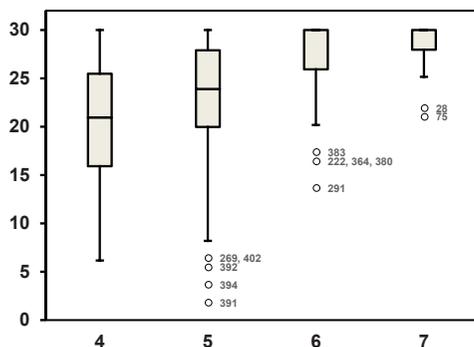
An independent-samples *t*-test was conducted to compare the gap-fill scores of the context and translation groups. The translation group mean

(11.1, SD = 3.4) was a little higher than the context group's mean (10.6, SD = 3.5), but there was no significant difference in the scores ( $t(401) = 1.55$ ,  $p = 0.53$ , two-tailed). The magnitude of the difference was very small ( $\eta^2 = 0.006$ ).

### Effect of proficiency

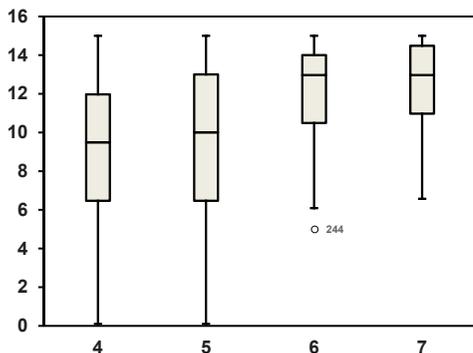
Earlier work (notably Prince 1996) suggested that weaker and more advanced students might respond differently depending on how they studied new vocabulary. In the Prince study, the weaker group achieved higher scores than even advanced students (when tested by translation) when they had studied via translation rather than via context. To investigate this possibility, the results of the pre-test were used to allocate, post hoc, the participants into weaker and more advanced groupings. This gave four groups: translation (weaker and advanced) and context (weaker and advanced).

We first examined the performance of the four groups in the section of the post-test that repeated the pre-test (matching L1 and L2 words, maximum score 30), as shown in Figure 6.



**Figure 6:** Repeated pre-test – context (4 weak, 6 advanced); translation (5 weak, 7 advanced).

Similarly, we analysed the results of the gap-fill section of the post-test (maximum score 15), as illustrated in Figure 7.



**Figure 7:** Gap-fill post-test – context (4 weak, 6 advanced); translation (5 weak, 7 advanced).

Statistical analysis using the Mann–Whitney  $U$  test showed that the weaker translation group performed significantly better than the weaker context group in the task involving translation (matching L1 and L2 words). The  $p$  value was 0.002 and the effect size ( $r = 0.22$ ) was small to medium. With the gap-fill exercise, a  $t$ -test showed no significant difference between the weaker translation and context groups ( $p = 0.45$ ).

A similar pattern was found with the advanced groups. The Mann–Whitney  $U$  test comparing the scores on the re-test matching showed that the translation group outperformed the context group ( $p = 0.002$ ) and the effect size was also small to medium ( $r = 0.21$ ). In the gap-fill test, a  $t$ -test revealed no significant difference between the two groups ( $p = 0.37$ ).

Thus, the level of vocabulary knowledge as measured before the treatment started seemed to have had no significant effect on the comparative performances of the translation and context groups. Proficient and less proficient students achieved similar comparative results.

## Discussion

What the context group achieved was impressive, as they had first to read and comprehend over 500 words in English, infer the meanings of the highlighted target words and commit them to memory. In the check test and consolidation exercises they had to again read over 500 words of text and link the target words either to a context or to a synonymic phrase in order to reinforce and broaden their understanding of the target vocabulary. Though their attention was drawn to the target vocabulary, it is reasonable to expect there would be opportunities for incidentally learning other aspects of the language.

Clearly, the translation group gained an advantage with respect to the test involving translation, but the advantage was marginal (as the effect size

was small to medium) and one has to consider whether it would be worth sacrificing class time where students and teachers could be communicating in L2 for the sake of that small gain which could be made up for, if it was considered necessary, by a fairly mechanical home study exercise.

Consider how, in practice, code-switching might differ from communicative teaching by considering the following examples of how an instructor in an English class in Thailand might teach the meaning of the word ‘criticise’, a word beyond the first 2,000 most frequent words in English according to the COCA (Corpus of Contemporary American English):

- 1 ‘Green groups criticise gold mining activity in Loei’. What do you think ‘criticise’ means in this headline?
- 2 ‘Criticise’ here means to say something bad or negative about a person or company.
- 3 ‘Criticise’ – in Thai we use ‘ตำหนิ’.
- 4 ‘Criticise’ ในที่นี้ หมายถึง การวิจารณ์บางสิ่งบางอย่างในทางที่ไม่ดีหรือทางลบเกี่ยวกับบุคคล.

Clearly, explanation 4 might be appreciated by students because of its clarity and simplicity, but it has less to offer in additional benefits. Vocabulary has to be comprehended, but it then has to be maintained and consolidated. Maintained because it is subject to attrition and consolidated because it is rare for a word to be entirely learned from one encounter. Explanations 1 and 2 offer most in these respects. For the results of this study indicate that students can ascertain the meaning of new words by contextual means and at the same time they can re-encounter other words that are not fully known or if known may suffer attrition. Perhaps a student who has not fully acquired ‘activity’ or ‘negative’ or other meanings for ‘green’ or ‘company’ then has the opportunity to acquire those words more fully. And words already known are re-encountered and attrition is prevented.

Further, it is not unreasonable to expect that students who are focused on learning vocabulary from context would also experience some incidental development of reading skills, including comprehension, and some widening of knowledge of collocations, levels of usage, grammatical usage and associations of the target vocabulary. These gains might be small in the short term but would accumulate over years. Nation (2014) demonstrates how a large vocabulary of 9,000 words could be accumulated through reading L2 texts.

In practice, while this experiment restricted each group to one method of learning, there is no need to ban all use of translation. It would probably be impossible anyway, for in a survey of the literature on L1 use in

communicative classes Levine concludes 'L1 was present in every class ever studied for amounts of L1 and L2 use' and most of all when students are speaking to each other (2014:335). Students learning words contextually or from L2 definitions or paraphrasing could be set homework in which they used a dictionary to find the L1 translations of words given in the context of a sentence. Following Nation's advice, they could make word cards of the target vocabulary including an L1 translation as well as a contextual sentence. However, in the classroom L2 could remain the ambient language of instruction.

A further aspect of this experiment to consider is the outcome for the less proficient learners. It is sometimes argued that teachers are justified in using more L1 to support such students. We examined our results to see if the weaker students who studied thirty vocabulary items via translation did better in comparison with similar students who studied contextually. We found the same pattern of results when we separated out those who had scored less than the average mark on the pre-test. The translation group did significantly better on the matching of target word and Thai translation (but again with a small to medium effect size); however, there was no significant difference between the two weaker groups on the gap-fill exercise. As in earlier studies, there was no sustained advantage for less proficient students from translation when they were required to use the learned words contextually. As Prince (1996) opined, the ability of a weaker student to translate a word did not seem to enhance a student's ability to use the word contextually. However, as we had no absolute beginners in our study we cannot offer any commentary on the situation with students who lack proficiency because they have not yet received instruction. We note that even a system such as the French immersion program in Canada, renowned for its strict commitment to L2-only classes, appears to be lenient with regard to L1 use in the early months of late immersion classes (Turnbull and Dailey-O'Cain, 2009).

### Limitations and conclusions

This paper has limitations. There was no opportunity to conduct a long-term delayed post-test that would not have been impacted by the instruction by the regular classroom teachers. Further, the design of the study involving a list of just thirty words incurs the criticism Meara (1996) makes of similar experiments, namely that such a small number of words can be learned as a mere list of items and that it is a very small sample of the vocabulary that learners need from which to draw conclusions.

A further limitation concerns the students' attitudes to learning English, which may not have been typical of students in many EFL classes in Thailand or other countries. The participants in this experiment were enrolled in one of Thailand's more prestigious universities and had generally positive attitudes to their institution and to learning English. Thus, the students in the context groups were cooperative when asked to engage with a text to attempt to discern the meanings of English words without any assistance from Thai translations. However, many EFL students may not always be so motivated. Lin (1999) noted how in Hong Kong schools, while English is readily learned by the children of the elite who encounter the language at home and are optimistic about their career prospects, it is a language less privileged students are often alienated from. According to Lin they are bored by English lessons and irritated by local teachers who conduct English lessons in English, and they fail to learn it. However, where potent social forces impact on learning it is doubtful that tactical moves such as code-switching or increasing contextual learning can counteract them.

Despite these limitations, this experiment provides evidence for instructors who contend that students are able to acquire new vocabulary through contextual learning and suggests that one does not have to routinely resort to code-switching to facilitate learning. Doubtless, many students will prefer to check their understanding of new words in, for example, a bilingual dictionary. But there does appear to be less justification for depriving L1 learners of access to comprehensible L2 input in the classroom for fear that they cannot learn new words otherwise.

### About the authors

**Michael John Alroe** is a language teacher who has spent 14 years teaching English as a Foreign Language to school and university students in Turkey, Hong Kong and Thailand.

**Hayo Reinders** ([www.innovationinteaching.org](http://www.innovationinteaching.org)) is Professor of Education at Unitec in New Zealand and TESOL Professor and Director of the doctoral programme at Anaheim University in the USA. Hayo has published over 20 books and 100 articles in the areas of autonomy, technology, teacher education and out-of-class learning. He edits a book series for Palgrave Macmillan and is editor of the journal *Innovation in Language Learning and Teaching*.

**Punchalee Wasanasomsithi** is an associate professor at Chulalongkorn University Language Institute. After graduating with a PhD in Language Education from Indiana University, USA, she has been teaching both undergraduate and graduate English courses, as well as supervising doctoral students working in the field of English Language Teaching and English as an International Language. Her areas of interests

include language acquisition and language learning strategies. She can be reached at [punchalee.w@chula.ac.th](mailto:punchalee.w@chula.ac.th)

## Notes

- 1 Code-switching can formally be defined as ‘the practice of selecting or altering linguistic elements so as to contextualize talk in interaction’ (Nilep 2006). But it is often also used to denote, and perhaps euphemise, teaching about a foreign language in the students’ L1.  
In Korea for example, Shin (2012) investigates reasons why even teachers fluent in English fail to follow the government policy of teaching English in English. Yamada and Hristoskova (2011) and Glasgow (2012) document the resistance to the Japanese Ministry of Education’s attempts to have English taught in English.
- 2 In that study it must be conceded that Webb found no additional benefit from providing a contextual sentence in addition to the Japanese translation of the target English words. However, it may be that participants focused only on the L1 translation as the easiest route to comprehending the meaning of the target word. And the exposure to the targets involved only one period of ninety minutes.
- 3 We would have preferred to test all thirty words via gap-fill, but were constrained by the reluctance of a number of the class teachers to allow us to take up yet more of their class time.

## References

- Alroe, M. J. and Reinders, H. (2015) The role of translation in vocabulary acquisition: a replication study. *Eurasian Journal of Applied Linguistics* 1(1): 39–58.
- Choi, S., Kim, J. and Ryu, K. (2014) Effects of context on implicit and explicit lexical knowledge: an event-related potential study. *Neuropsychologia* 63: 226–34. <https://doi.org/10.1016/j.neuropsychologia.2014.09.003>
- Cobb, T. (2016) *Vocabulary Levels Test, Productive*. Version A. <http://www.lexutor.ca/tests/levels/productive>
- Domalewska, D. (2015) Classroom discourse analysis in EFL elementary lessons. *International Journal of Languages, Literature and Linguistics* 1(1): 6–9. <https://doi.org/10.7763/IJLLL.2015.V1.2>
- Duff, P. A. and Polio, C. G. (1990) How much foreign language is there in the foreign language classroom? *Modern Language Journal* 74(2): 154–66. <https://doi.org/10.1111/j.1540-4781.1990.tb02561.x>
- Elgort, I. (2011) Deliberate learning and vocabulary acquisition in a second language. *Language Learning* 61(2):367–413. <https://doi.org/10.1111/j.1467-9922.2010.00613.x>
- Ellis, R. (2015) The importance of focus on form in communicative language teaching. *Eurasian Journal of Applied Linguistics* 1(2): 1–12.
- Folse, K. S. (2006) The effect of type of written exercise on L2 vocabulary retention. *TESOL Quarterly* 4(2): 273–93. <https://doi.org/10.2307/40264523>
- Glasgow, G. P. (2012) Implementing language education policy to ‘conduct classes in English’ in Japanese senior high schools. In A. Stewart and N. Sonda (eds) *JALT2011 Conference Proceedings*. Tokyo: JALT.
- Groot, P. J. M. (2000) Computer assisted second language vocabulary acquisition. *Language Learning & Technology* 4(1): 60–81.

- Huckin, T. and Coady, J. (1999) Incidental vocabulary acquisition in a second language. In review of *Incidental L2 Vocabulary Acquisition: Theory, Current Research, and Instructional Implications*, M. Wesche and T. S. Paribakht (eds). *Studies in Second Language Acquisition* 21(2): 181–93. <https://doi.org/10.1017/S0272263199002028>
- Hulstijn, J. H. (2008) Incidental and intentional learning. In C. J. Doughty and M. H. Long (eds) *The Handbook of Second Language Acquisition* 349–81. Oxford: Blackwell.
- Karpicke, J. D. and Roediger, H. L. (2008) The critical importance of retrieval for learning. *Science* 319(5865): 966–8. <https://doi.org/10.1126/science.1152408>
- Laufer, B. and Girsai, N. (2008) Form-focused instruction in second language vocabulary learning: a case for contrastive analysis and translation. *Applied Linguistics* 29(4): 694–716. <https://doi.org/10.1093/applin/amn018>
- Laufer, B. and Shmueli, K. (1997) Memorizing new words: does teaching have anything to do with it? *RELC Journal* 28: 89–108. <https://doi.org/10.1177/003368829702800106>
- Lee, J. H. and Macaro, E. (2013) Investigating age in the use of L1 or English-only instruction: vocabulary acquisition by Korean ELF learners. *Modern Language Journal* 97(4): 887–901. <https://doi.org/10.1111/j.1540-4781.2013.12044.x>
- Levine, G. S. (2014) Principles of code choice in the foreign language classroom: a focus on grammaring. *Language Teaching* 47(3): 332–48. <https://doi.org/10.1017/S0261444811000498>
- Lightbown, P. M. and Spada, N. (2013) *How Languages are Learned* (4th edn). Oxford: Oxford University Press.
- Lin, A. (1999) Doing-English-lessons in the reproduction or transformation of social worlds? *TESOL Quarterly* 33(3): 393–412. <https://doi.org/10.2307/3587671>
- Liu, D., Ahn, G., Baek, K. and Han, N. (2004). South Korean high school teachers' code switching: questions and challenges in the drive for maximal use of English in teaching. *TESOL Quarterly* 38(4): 605–38. <https://doi.org/10.2307/3588282>
- Macaro, E. (2014) Overview: where should we be going with classroom codeswitching research? In R. Barnard and J. McLellan (eds) *Codeswitching in University English-Medium Classes*. Bristol, UK: TJ International.
- Meara, P. (1996) The classical research in L2 vocabulary acquisition. In G. Andaman and M. Rogers (eds) *Words, Words, Words* 27–40. Clevedon, UK: Multilingual Matters.
- Nation, P. (2001) *Learning Vocabulary in Another Language*. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9781139524759>
- Nation, P. (2003) The role of first language in foreign language learning. *Asian EFL Journal* 5(2): 1–8.
- Nation, P. (2014) How much input do you need to learn the most frequent 9,000 words? *Reading in a Foreign Language* 26(2): 1–16.
- Nilep, C. (2006) 'Code switching' in sociocultural linguistics. *Colorado Research in Linguistics* 19(June): 1–22.
- Prince, P. (1996) Second language vocabulary learning: the role of context versus translations as a function of proficiency. *Modern Language Journal* (80)4: 478–93. <https://doi.org/10.1111/j.1540-4781.1996.tb05468.x>
- Prince, P. (2012) Towards an instructional programme for L2 vocabulary: can a story help? *Language Learning & Technology* 16(3): 103–20.
- Rieder, A. (2003) *Implicit and Explicit Learning in Incidental Vocabulary Acquisition*. Paper presented at the EUROSLA Conference, Edinburgh.

- Schmitt, N. (2010) *Researching Vocabulary: A Vocabulary Research Manual*. Basingstoke, UK: Palgrave Macmillan. <https://doi.org/10.1057/9780230293977>
- Shin, S. (2012) 'It cannot be done alone': the socialization of novice English teachers in South Korea. *TESOL Quarterly* 46(3): 542–67. <https://doi.org/10.1002/tesq.41>
- Stempleski, S., Joannsen, K., Curtis, A. and Douglas, N. (2006) *World Pass, Upper Intermediate*. Boston, MA: Heinle & Heinle.
- Tian, L. and Macaro, E. (2012) Comparing the effect of teacher codeswitching with English-only explanations on the vocabulary acquisition of Chinese university students: a lexical focus-on-form study. *Language Teaching Research* 16(3): 367–91. <https://doi.org/10.1177/1362168812436909>
- Turnbull, M. and Dailey-O'Cain, J. (2009) *First Language Use in Second and Foreign Language Learning*. Bristol, UK: Multilingual Matters.
- Webb, S. (2007) Learning word pairs and glossed sentences: the effects of a single context on vocabulary knowledge. *Language Teaching Research* 11(1): 63–81. <https://doi.org/10.1177/1362168806072463>
- Webb, S. (2009) The effects of receptive and productive learning of word pairs on vocabulary knowledge. *RELC Journal* 40(3): 360–76. <https://doi.org/10.1177/0033688209343854>
- Yamada, H. and Hristoskova, G. (2011) Teaching and learning English in Japanese senior high schools. *Journal of Fukui-ken Eigo Kenkyu-kai* 69: 1–33.