Title
The effects of instructor language choice on second language vocabulary learning and listening comprehension

Permalink
https://escholarship.org/uc/item/42d560vd

Authors
Lee, JH
Levine, GS

Publication Date
2018-05-01

DOI
10.1177/1362168818770910

Peer reviewed
The effects of instructor language choice on second language vocabulary learning and listening comprehension

Jang Ho Lee
Chung-Ang University, Republic of Korea

Glenn S. Levine
University of California Irvine, USA

Abstract
The present study examined the interaction effects between teachers’ choice of language in lexical explanation and second language (L2) learners’ proficiency level on the learning of phrasal verbs and listening comprehension in a meaning-focused listening activity. Undergraduate L2 learners with two different levels of proficiency (intermediate and advanced) were assigned to three different conditions. These conditions included (1) code-switching, in which teachers briefly switched to learners’ first language for vocabulary teaching purposes, (2) L2-only, in which teachers maintained L2 for overall instruction and vocabulary teaching, and (3) control condition, in which learners were not given instruction about target vocabulary. The results revealed that teachers’ instruction about target vocabulary brought about more gains in learning of phrasal verbs than lack thereof. In addition, teachers’ use of code-switching benefited intermediate-level learners most, who were able to gain similar levels of vocabulary learning and listening comprehension to advanced learners.

Keywords
code-switching, EFL, L1 use in the L2 classroom, L2 listening comprehension, L2 vocabulary acquisition

Corresponding author:
Glenn S Levine, University of California Irvine, 243 Humanities Instructional Building, Irvine, CA 92797-3150, USA.
Email: glevine@uci.edu
I Introduction

Research on teachers’ language use in second language (L2) classrooms has increased dramatically in number over the past 20 years, with studies looking at this issue from numerous perspectives (e.g. de la Campa & Nassaji, 2009; Lee, 2016; Lee & Macaro, 2013; Levine, 2003, 2011, 2014; Macaro, 2001; Rolin-Ianziti & Brownlie, 2002; Rolin-Ianziti & Varshney, 2008; van der Meij & Zhao, 2010; Zhao & Macaro, 2016). One strand of this research has sought to measure the amount of the first language (L1) and L2 use on the part of teacher (e.g. de la Campa & Nassaji, 2009; Macaro, 2001), while another has examined the relationships between particular pedagogical functions (e.g. explaining the meaning of L2 vocabulary, giving instruction for activities) and each language code (e.g. Levine, 2003; Rolin-Ianziti & Brownlie, 2002). More recently, studies have explored students’ and teachers’ beliefs about and attitudes towards teachers’ classroom language uses (e.g. Lee, 2016; Rolin-Ianziti & Varshney, 2008). These studies have overall converged in support of the use of the L1 in L2 teaching, which stands at odds with what some national agencies and organizations have suggested (e.g. ACTFL, 2010; Curriculum Development Council in Hong Kong, 2004). That is, teachers in a wide range of contexts, regardless of students’ age or proficiency, use the L1 for pedagogical purposes, and a large number of students wishes to have, or does have, the L1 as a pedagogical resource in the process of learning the L2.

To this end, many scholars investigating classroom language use have adopted the term code-switching (CS) from the sociolinguistic literature on multilinguals and their language use to refer to the alternation of two or more languages within the same discourse, with codes usually referring to different languages (e.g. Liebscher & Dailey-O’Cain, 2003, 2004; Li Wei & Martin, 2009). Among the available models for analyzing CS, such as Li Wei and Auer’s sequential analysis approach (Auer, 1998; Li Wei, 1998) or Giles’s communication accommodation theory (Giles & Ogay, 2007), we find Myers-Scotton’s (1997) Matrix Language Frame (MLF) model most useful for analyzing the sort of classroom CS manifested in our study. According to this framework, languages involved in instances of CS can be distinguished between the matrix language (ML) which plays a dominant role, and the embedded languages (EL) which ‘also participate in CS, but with a lesser role’ (p. 3). In the present study we therefore adopt the notions of the ML and EL in defining classroom CS. In the pedagogical context of this study, the ML is the learners’ L2 (i.e. English), which is the primary language of discourse among classroom participants during class time, whereas EL is the learners’ L1 (i.e. Korean). In our context, participants would use English as an unmarked code choice but may switch to Korean both intersententially and intrasententially for a range of pedagogical and discursive purposes. Intersentential CS in the classroom context refers to the alternate uses of each language at the clausal or sentential level on the part of an English teacher and/or the students. In the case of an intrasentential CS, some elements (or constituents) in Korean are embedded in English sentences, which ‘sets the morphosyntactic frame of sentences’ as the ML (Myers-Scotton, 1997, p. 3).

Despite the broad range of studies on classroom CS and the larger goal of determining whether the L1 use helps or hinders L2 learning, however, we actually note a paucity of studies on the effects of teachers’ use of CS on L2 learning. Periodically researchers have
indeed called for more studies examining ‘the relationship between teachers’ TL [target language] and L1 use and students’ TL proficiency’ (Turnbull, 2001, p. 537) and/or ‘the effect of codeswitching on interaction [among classroom participants] and in turn on language learning’ (Macaro, 2009, p. 82); however, this call has not been answered to the extent it should be. Exceptions are Tian and Macaro (2012), Lee and Macaro (2013), and Zhao and Macaro (2016), which examined the differential effects of teacher CS and English-only instruction on English-as-a-foreign-language (EFL) learners’ vocabulary acquisition during meaning-focused activities. These studies reported the benefits of CS over L2-only in terms of vocabulary learning. While this research is interesting and important, none has examined the learners’ vocabulary learning as well as comprehension of the listening passage used in meaning-focused activities under different conditions, CS or L2-only. It would thus be an appropriate next step to explore whether the advantage resulting from teacher CS in terms of L2 vocabulary learning would foster or compromise L2 learners’ comprehension of L2 texts.

The primary aim of the present study is, therefore, to investigate the effects of teacher CS and L2-only instruction for undergraduate EFL learners on the comprehension of listening texts, as well as the learning of target lexical items in those activities. In addition, we also considered proficiency an important factor, which was previously suggested in Lee and Macaro (2013) with adult and young EFL learners; thus, the interaction effect between teacher’s language use and learners’ proficiency was included in our investigation. The findings of the present study should contribute to building a pedagogical model of principled teacher CS in L2 lessons.

In the following section, relevant empirical studies on the effects of teacher CS on L2 learning are reviewed. The methodology of the present study is then described. Next, the results of the study are presented in response to the research questions. Finally, the results are discussed, along with implications for language pedagogy and teacher education, as well as the limitations of the study.

II Studies on the effects of teacher code-switching

As indicated in the previous section, only a small number of empirical studies on the effects of teachers’ CS use versus L2-only instruction on student learning has appeared to date. These studies are mostly concerned with the relative effects of teacher CS or L2-only instruction on L2 vocabulary learning (Hennebry et al., 2017; Lee & Macaro, 2013; Tian & Macaro, 2012; Zhao & Macaro, 2016). The studies have some common methodological grounds. First, they examine this issue in ‘a pedagogical context where vocabulary learning is not the prime objective but runs subordinate to, or parallel with, the objective of accessing the general meaning of a [listening or reading] text’ (Hennebry et al., 2017, p. 286). Second, they draw on psycholinguistics literature on learners’ mental lexicon (Jiang, 2000; Kroll & Stewart, 1994) in order to examine the mediating effect of learners’ L2 proficiency in relation to that of the teachers’ CS or L2-only instruction.

Tian and Macaro (2012) and Hennebry et al. (2017) examined the effects of teachers’ CS and L2-only instruction in the context of listening comprehension activities, respectively, with Chinese EFL undergraduates and Year 9 learners of French in the UK. Both studies included the two experimental groups (CS and L2-only), along with a control
group which did not receive instruction of the target vocabulary. Tian and Macaro called this instructional behavior on the part of the teacher lexical focus-on-form (Laufer & Girsai, 2008). These studies showed that providing instruction about target vocabulary resulted in better L2 vocabulary learning than lack thereof (control condition), and that teacher CS had an advantage over L2-only in the short term. However, in the long term the greater effectiveness of teacher CS was not sustained (Tian & Macaro, 2012). Also, the participants’ L2 proficiency was not found to be related to the effects of teachers’ vocabulary instruction in two languages.

With a similar research aim, Zhao and Macaro (2016), and Lee and Macaro (2013), conducted intervention studies on the effects of teacher CS on vocabulary learning in EFL contexts. The pedagogical context of these studies was a reading activity in which the primary goal was to comprehend the general meaning of the target text, with the learning of target lexical items embedded in the text as a secondary objective. Zhao and Macaro worked with Chinese non-English-major undergraduates, with instruction of target lexical items for the treatment groups following the reading tasks using CS or L2 only. The results showed that the treatment groups performed better than the control group on both immediate and delayed post-tests. Between the two experimental groups, the teacher’s CS showed greater effectiveness than L2-only instruction, both in the short term and long term. It was also found that this advantage was applicable to the teaching of both concrete and abstract L2 words. Lee and Macaro (2013) was the only study on the effects of teacher CS which included two different learner populations, adult undergraduate students and primary level children. The results corroborated those of Zhao and Macaro, in that CS was found to bring about more gains in L2 vocabulary learning in both the short and long terms. Another important finding of this study was that the young learners with lower L2 proficiency level and less L2 learning experience were found to benefit more from teacher CS than their adult counterparts.

A review of previous studies on the effects of teacher CS yields a consistent finding: A greater effectiveness of teacher CS over L2-only instruction. Yet the studies also point toward some important issues that remain unexplored. First, learners’ understanding of the general meaning of L2 texts was an ostensible objective of the investigated lessons, yet this has not been examined sufficiently. Provided that the aim of the pedagogical context in these studies was both understanding the general meaning of the texts and teaching target vocabulary, the effects of teachers’ language choices on the former should be given more attention. Second, as previous studies have presented consistent findings on L2 vocabulary learning, the direction of further research should be focused on the effects of teacher CS on the learning of particular categories of L2 vocabulary. The present study focused on phrasal verbs, which are considered challenging for L2 learners (Hulstijn & Marchena, 1989) and also have been examined in recent instructed L2 vocabulary studies (e.g. Nassaji & Tian, 2010). Third, an interaction effect between teachers’ language choice and learners’ L2 proficiency deserves more attention, as the findings of previous studies on the relative effects of L1 and L2 input on vocabulary learning have been rather inconclusive (Miyasako, 2002; Tian & Macaro, 2012). In addition to these issues, the present study addresses a question of whether lexical focus-on-form can be more beneficial in learning English phrasal verbs in a meaning-focused
activity than lack thereof, before examining relative effects of two types of instruction (i.e. teacher CS or L2-only instruction).

### III Research questions

Based on our review of literature, three research questions were formulated:

- **Research question 1:** Does lexical focus-on-form benefit the adult EFL learners’ learning of English phrasal verbs during a focus-on-meaning (i.e. listening) activity?

- **Research question 2:** What are the interaction effects between teachers’ language of instruction (i.e. code-switching or English-only) and proficiency (i.e. intermediate or advanced) on the EFL learners’ listening comprehension of the target listening texts?

- **Research question 3:** What are the interaction effects between teachers’ language of instruction (i.e. code-switching or English-only) and proficiency (i.e. intermediate or advanced) on the EFL learners’ acquisition and retention of English phrasal verbs?

### IV Methods

#### I Participants and instructors

The participants sampled for the present study were 195 undergraduate Korean EFL learners with a wide range of academic majors. The English lessons they had received during their secondary years were largely based on reading and grammar, as the college entrance exam was geared towards evaluating these aspects of the language. All the participants had taken two 3-credit-hour mandatory courses titled English Reading I, II and one three-hour mandatory course titled English Conversation I prior to the present study. The former aimed to enhance students’ English reading skills, and English was used by the instructor and students most of the time. By contrast, the latter class was designed based on a communicative approach,\(^1\) with the goal of enhancing students’ communicative competence and a pedagogical focus on speaking and listening. Along with English Conversation I, English Conversation II, in which the participants were registered by the time of the study, was operated under the ‘English maximum’ policy; thus, the instructors and students did not switch to L1 Korean frequently, and did so only for pedagogical purposes. For this course, the instructors held regular meetings to make their lessons as similar as possible, sharing their lesson plans and materials.

The participants were allocated either to ‘intermediate’ or ‘advanced’ level based on their final grade on English Conversation I. This grade was calculated based on the combined scores of the students’ performance on the following assessment criteria throughout the semester: two oral interviews with the instructor, two individual oral presentations, one computerized speaking test, their participation in class, and attendance.

The participants were assigned to one of the three instructors, two male and one female bilingual EFL instructors, including one of the authors, who all had more than
three years of teaching experience in EFL contexts. The instructors had studied abroad in English-speaking countries for more than five years prior to their teaching in Korea.

2 Group allocation

For the course English Conversation II, 195 participants were assigned to either intermediate or advanced level based on their grade in English Conversation I, and the three instructors mentioned in Section IV.1 were in charge of these classes. Participants in 12 classes divided into the two proficiency levels were than randomly allocated to English-only (EO), code-switching (CS), and control conditions (CT). Table 1 presents the allocation of the classes to each group, and instructors of these classes. Class and instructor names are all assigned an alphabetized initial for anonymity.

3 Intervention

The target learning context in the present study centered on a listening activity, a frequently used type of activity in the course. This activity deserves some description here, as it is a key element in the methodological design of the study.

In the activity, students first hear an audio recording of the text. In the first round of listening, students are asked to focus on the content only. The instructor then plays the recording of the listening text a second time, but now he or she stops the recording whenever either a target lexical item appears or a student raises a question about any lexical item. Thus, this vocabulary teaching moment can be considered lexical focus-on-form (Laufer, 2005) and can also be planned prior to the class by the instructor (Long & Robinson, 1998). After the instructor and students go through the target listening text twice in this way, the students are given either listening comprehension questions about the target text in multiple-choice or fill-in-the-blank format, or a recall task.

Table 1. Allocation of the group.

<table>
<thead>
<tr>
<th>Group</th>
<th>Class and instructor in charge</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intermediate:</td>
<td></td>
</tr>
<tr>
<td>EO</td>
<td>Class A (n = 18) / Instructor C</td>
</tr>
<tr>
<td>CS</td>
<td>Class C (n = 16) / Instructor B</td>
</tr>
<tr>
<td>CT</td>
<td>Class D (n = 14) / Instructor C</td>
</tr>
<tr>
<td>Advanced:</td>
<td></td>
</tr>
<tr>
<td>EO</td>
<td>Class G (n = 19) / Instructor A</td>
</tr>
<tr>
<td>CS</td>
<td>Class I (n = 18) / Instructor A</td>
</tr>
<tr>
<td>CT</td>
<td>Class K (n = 14), Class L (n = 13) / Instructor C</td>
</tr>
</tbody>
</table>

Notes. EO = English-only; CS = code-switching; CT = control.
The overall procedure of the listening activity was adopted in the present study, but small modifications were made to distinguish three intervention groups, while the other aspects of the task (e.g. listening to the target text twice, seating arrangements) were kept constant for all groups. For the CS and EO groups, the overall procedure was the same, but the language used to explain the target vocabulary was different and strictly controlled. That is, while English instruction was maintained throughout the lesson for the EO group, the instructor used intrasentential CS\(^2\) to provide the meaning of the target vocabulary for the CS group. The instructors’ CS was limited to embedding some constituents about the lexical meaning of a target phrasal verb either at a phrase or clause level, depending on target phrasal verbs. Thus, the syntactic structure of intrasentential CS ranged from verb or adjectival phrase to complementizer phrase in the participants’ L1, with these being a complement to ‘be’ or ‘mean’ verbs in English sentences (see Excerpt 1 below for examples). Teachers’ CS was strictly limited to giving the meaning of target vocabulary and did not include further explanation about target phrasal verbs. In the CS condition, after the instructors provided the intrasentential CS for vocabulary instruction, they switched back to English until a new target phrasal verb emerged.

The following excerpts are sentential structures of instructors’ vocabulary instruction in the CS and EO conditions, which all the instructors of the present study adopted. Note that in Excerpt 1 we do not gloss the Korean because the meaning is very close to what is said in the EO condition in Excerpt 2.

Excerpt 1. Sentential structures used for vocabulary teaching in CS condition

T: ‘Be tied up’ means ‘너무 바빠서 어디 가거나 다른 사람을 만날 시간이 없다’ in Korean.

T: The meaning of ‘be tied up’ in Korean is ‘너무 바빠서 어디 가거나 다른 사람을 만날 시간이 없다’.

Excerpt 2. Sentential structures used for vocabulary teaching in EO condition

T: ‘Be tied up’ means ‘be so busy that you are unable to see or speak to anyone else or go anywhere’.

T: The meaning of ‘be tied up’ is ‘be so busy that you are unable to see or speak to anyone else or go anywhere’.

In both conditions, the instructor wrote the target vocabulary on the board, so that the written as well as oral form of the target vocabulary item was clear. Also, the amount of time needed to explain each target phrasal verb was almost the same in both conditions. We would also note that all of the vocabulary included in the EO explanation of the target phrasal verbs were carefully selected from the list of essential vocabulary compiled for secondary-level English education in the Republic of Korea. Thus, it is unlikely that the participants would have been unfamiliar with any of these words.

After the second round of listening, both the EO and CS groups were given a recall task, which asked them to write down the content of the text in as much detail as possible, using complete sentences. In the case of the control group, the instructor also played the recording twice but provided no explanation of the target vocabulary. This group was
then given a set of multiple-choice listening comprehension questions instead of the recall task. The control group also did not receive the recall task because a pilot study with 32 learners similar to those in the main study showed that it was difficult for this group of learners to perform the recall task at all without instruction about unfamiliar vocabulary.

4 Target listening texts and vocabulary

The texts were adapted from *English phrasal verbs in use: Intermediate* (McCarthy & O’Dell, 2004) and *English phrasal verbs in use: Advanced* (McCarthy & O’Dell, 2007), both vocabulary textbooks for English learners. They present a wide range of phrasal verbs in either a short text or dialog format. The authors chose two dialogs and one short text, respectively, presenting phrasal verbs related to ‘work’ and ‘study’. The difficulty of the target listening texts, and target phrasal verbs, were piloted with the aforementioned learners. Based on the result of the piloting, the length of the target texts was modified, and some non-target words which were judged to be difficult to comprehend for the participants in the listening mode were replaced with more commonly known words. Also, some of phrasal verbs which were known to more than half of the participants in the pilot study were replaced. The two texts used in the main study contained 239 and 243 words, respectively. The lists of the target phrasal verb used in the main study are as follows:

**On work:** be tied up, pile up, be snowed under, slave away, ease off, branch out into, fling (oneself) into, keep at, chase down, cut out for, squeeze in, wriggle out of (12 phrasal verbs)

**On study:** put through (university), fly through (test or exam), buckle down, get through (test or exam), count against, be marked down (for something), crop up, mug up (for something), fall behind, bury (oneself) in (something), pull one’s socks up (11 phrasal verbs)

These target phrasal verbs were selected based on the following criteria: they were judged to be useful for the sampled participants by the authors and instructors; they were found to be unfamiliar to most students of a similar level of English proficiency who participated in the pilot study; and as an additional means of ensuring parity between the two experimental groups, the relative amount (or length) of Korean and English explanation needed of the target lexical items was found to be comparable. Additionally, while the discursive features of the teachers’ talk in the explanations—such as length, complexity, or register—may affect learning in different ways, our intention was to keep each explanation as brief as possible. So, for example, in the case of ‘buckle down’, the explanation given in CS and EO conditions was as follows, respectively: ~에 본격적으로 하나에 집중하여 힘쓰다 (for CS), be determined to work/focus hard on something (for EO). The authors held two discussion sessions with the instructors to ensure that the English and Korean explanation of the finalized phrasal verbs would be comprehensible; both sets of explanations were revised somewhat during these sessions. The finalized texts and target phrasal verbs are given in Appendix 1.
5 Instruments

Two instruments were used to examine the effects of interaction between teachers’ language of instruction and proficiency on the participants’ listening comprehension and vocabulary learning. For listening comprehension, a recall test was used which asked the participants to write down as much as possible about the content of the listening texts. Regarding vocabulary, a receptive vocabulary test was administered through which the participants were asked to give the definition or meaning of each target item in either their Korean or English. The participants were given two language options for their answers so as not to bias either condition, following Tian and Macaro (2012). Despite this choice, about 95% of the responses were in Korean.

Both instruments were piloted along with the target listening texts and vocabulary with the same group of learners from the pilot study. It was found then that these two testing formats were neither cognitively demanding nor unfamiliar to this group of the learners.

6 Procedure

The experiment was conducted between the fourth and ninth weeks of the 16-week course. The study began three weeks into the semester so that the participants would become familiar with their instructors’ teaching styles and oral English speech style. In the fourth week, the participants were informed that the purpose of the present study was related to listening comprehension. The participants were also given an informed consent form. The three groups were then given a pre-test of target vocabulary, which consisted of both the aforementioned 23 target phrasal verbs and 17 distractor lexical items. The participants were given 10 minutes to complete this test. In the fifth week, the participants completed a trial listening comprehension activity. The purpose of this trial activity was twofold: The participants would become familiar with this type of activity, and the instructors would be given an opportunity to make adjustments to the listening activity in terms of its structure and the ways of delivering Korean and English explanations of target vocabulary in the two conditions. In the sixth week, the first target listening text on the topic ‘work’, which contained 12 phrasal verbs, was presented to the participants. This was followed by the recall task and immediate post-test of vocabulary (CS and EO groups only). The vocabulary test was administered without prior notice. The listening activity and recall task lasted for approximately 15 minutes, and the participants were given five minutes for the immediate post-test of vocabulary. In the seventh week, the second target listening text on the topic ‘study’ with another 11 of the target phrasal verbs was presented, again followed by the recall task and immediate post-test of vocabulary. Two weeks after the second immediate post-test of vocabulary, the delayed post-test of vocabulary, which consisted of all the target phrasal verbs from the two texts, was administered, again without prior notice. Following Tian and Macaro (2012), the control group was not given the immediate post-test of vocabulary, in order to avoid any priming effect. The participants were given ten minutes to complete the delayed post-test. Table 2 illustrates the sequence of events (i.e. treatment and tests) of the present study.
Regarding the scoring of the recall test, one point was given for each idea unit. This means that two sentences which expressed the same idea unit but in different wording was counted as one point. For example, ‘Two friends haven’t met for long time’ and ‘Nick didn’t have a chance to see Tim recently’ were considered the same idea unit. Also, an idea unit that did not represent the content of the target listening text was given zero points. Spelling or grammatical errors in the participants’ responses were not a criterion in scoring. For each listening text, ten point was the maximum score. By contrast, for the vocabulary test, two points were given for a completely correct definition or meaning of each phrasal verb, whereas one point was given for a semantic approximation of the correct meaning. As the total number of the target phrasal verbs was 23, the maximum score for the pre-test, and immediate post-test, and delayed post-test was all 46. For the scoring of these two types of test, one of the instructors who participated in the present study and another EFL instructor with a master’s degree in TESOL were asked to rate the participants’ responses, with 15% of the total data being rated by both raters. The interrater reliability was .87 for the recall test and .91 for the vocabulary test. The disagreement was resolved through the discussion after the rating was completed.

Among 195 participants, 17 failed to complete the entire set of the tests throughout the study for personal reasons; thus, 178 was the final number of participants in our statistical analysis. For the first research question regarding the effects of lexical focus-on-form on the EFL learners’ vocabulary learning, Analysis of Covariance (ANCOVA) was conducted, with a pre-test of vocabulary as a covariate, delayed post-test of vocabulary as a dependent variable, and condition (EO, CS, and CT) as a between-group variable. It should be noted again that the control group did not take the immediate post-test of vocabulary, and thus only the delayed test was subject to this examination. For the second research question on the effect of interaction between proficiency and condition on the listening comprehension of these undergraduate EFL learners, two-way factorial ANOVA was conducted, with proficiency and condition as between-group variables, and the participants’ total score on the two recall tasks as a dependent variable. Finally, for the third research question on the effect of interaction between proficiency (intermediate
and advanced) and condition (EO and CS) on the learning of English phrasal verbs, a mixed Analysis of Variance (ANOVA) was used, with two between-group variables (proficiency and condition) and one within-group variable ‘time’, which was measured at three different testing points: Prior to the intervention; immediately after intervention; and two weeks after the intervention, respectively, measured by the pre-test, immediate post-test, and delayed post-test of vocabulary. Contrasts were also conducted in order to break down each effect of interaction between and among between-group and within-group variables (see Research question 3 in Results below). SPSS 23.0 was used for all the statistical analysis in the present study.

V Results

1 Descriptive statistics and assumptions for using parametric tests

The means and standard deviations of the participants’ scores on the three vocabulary tests at different points in time, and two recall tests are summarized in Table 3. The general pattern emerging from Table 3 is that the two pairs of experimental groups performed better overall than the control groups between the pre-test and the delayed post-test, but that the intermediate learners in the English-only group performed worse than the other three experimental groups, except for the pre-test, of course, which was given prior to the intervention. This pattern was not only shown in the vocabulary tests, but also in the text recall tests.

In order to use parametric tests such as ANCOVA, factorial ANOVA, and mixed ANOVA (for more details, see Section IV.7), the assumptions of normality and heterogeneity of variances were checked for each groups, by using Shapiro–Wilk and Levene’s tests. The examination of these assumptions showed that the data of the pre-test of vocabulary, immediate post-test of vocabulary, and recall test did not violate these assumptions. On the other hand, the normality assumption was not met for the EO and

<table>
<thead>
<tr>
<th>Group</th>
<th>Vocabulary test</th>
<th>Listening text recall</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pre-test M (SD)</td>
<td>Immediate post-test M (SD)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Text 1 M (SD)</td>
</tr>
<tr>
<td>Code-switching:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Advanced (n = 35)</td>
<td>6.34 (3.40)</td>
<td>27.31 (8.26)</td>
</tr>
<tr>
<td>Intermediate (n = 28)</td>
<td>5.21 (3.53)</td>
<td>26.46 (8.46)</td>
</tr>
<tr>
<td>English-only:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Advanced (n = 32)</td>
<td>7.03 (3.63)</td>
<td>23.56 (6.90)</td>
</tr>
<tr>
<td>Intermediate (n = 33)</td>
<td>5.39 (2.61)</td>
<td>14.42 (6.19)</td>
</tr>
<tr>
<td>Control:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Advanced (n = 25)</td>
<td>8.16 (3.18)</td>
<td>–</td>
</tr>
<tr>
<td>Intermediate (n = 25)</td>
<td>4.08 (2.60)</td>
<td>–</td>
</tr>
</tbody>
</table>

Table 3. Descriptive statistics.
Language Teaching Research 00(0)

Table 4. Pairwise comparisons among the code-switching (CS), English-only (EO), and control groups.

<table>
<thead>
<tr>
<th>Comparison</th>
<th>Mean difference</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>CS–Control</td>
<td>9.08</td>
<td>.001</td>
</tr>
<tr>
<td>EO–Control</td>
<td>6.40</td>
<td>.001</td>
</tr>
<tr>
<td>CS–EO</td>
<td>2.68</td>
<td>.022</td>
</tr>
</tbody>
</table>

Note. Adjustment for comparisons: Bonferroni.

CS group’s scores on the delayed test ($W = .927, df = 65, p = .001$ for the EO group, $W = .947, df = 63, p = .009$ for the CS group), and also homogeneity of variance among the three groups on the delayed test was found to be violated ($p = .001$). While ANOVA has been suggested to be robust to the violation of these assumptions, the findings of the present study regarding the delayed post-test should be read with some caution.

2 Research question 1

In order to examine the effects of lexical focus-on-form, ANCOVA was conducted, with the delayed post-test as a dependent variable, the pre-test as a covariate, and condition (i.e. CS, EO, control) as a between-group variable. The covariate, the participants’ pre-test score, was significantly related to the delayed post-test score, $F (1, 174) = 66.90, p < .001, \eta_p^2 = .28$. There was also a significant effect of condition on the delayed post-test score, after controlling for the pre-test score, $F (2, 174) = 37.99, p < .001, \eta_p^2 = .30$, with a large effect size. Pairwise comparisons (i.e. Bonferroni test) based on estimated marginal means among the groups all found differences at $p < .05$, with Table 4 showing the mean difference based on estimated marginal means and p value.

3 Research question 2

For the second research question, which was concerned with the participants’ listening comprehension, factorial ANOVA with two between-group variables (i.e. proficiency and condition) was conducted. This analysis did not find a statistical effect for condition, $F (1, 124) = 1.61, p = .21, \eta_p^2 = .01$, indicating that there was no significant difference between the CS and EO group. However, there was a significant effect for proficiency, $F (1, 124) = 6.91, p = .01, \eta_p^2 = .05$, with the advanced learners performing better than the intermediate learners in general. Finally, the interaction between condition and proficiency in terms of the listening comprehension was found to be statistically significant, $F (1, 124) = 7.53, p = .007, \eta_p^2 = .06$. Figure 1 illustrates this interaction: the advanced and intermediate learners in the CS condition showed almost no difference in their mean score, whereas under the EO condition, the advanced learners showed better performance than their intermediate counterparts. In all, CS resulted in a similar level of listening comprehension for the advanced and intermediate learner groups, whereas EO favored the advanced group more than the intermediate group.
Research question 3

For the third research question, a mixed Analysis of Variance (ANOVA) was used, with two between-group variables (proficiency and condition) and one within-group variable ‘time’. The results of mixed ANOVA revealed that there was a significant main effect of time for vocabulary learning, $F(1.87, 248) = 432.17, p < .001, \eta^2_p = .78$. Contrasts revealed that the participants’ average score on the immediate post-test was significantly higher than that on the pre-test, $F(1, 124) = 685.85, p < .001, \eta^2_p = .85$, and their score on the delayed post-test was again significantly higher than that of the pre-test, $F(1, 124) = 266.29, p < .001, \eta^2_p = .68$.

Regarding the between-group variables, there was a significant effect of condition, $F(1, 124) = 13.70, p < .001, \eta^2_p = .10$, and also of proficiency, $F(1, 124) = 11.46, p = .001, \eta^2_p = .09$. In view of the descriptive statistics presented above in Table 3, these results indicate that the advanced participants overall scored higher on the vocabulary tests than their intermediate counterparts, and that the participants in the CS condition performed better on the same tests than those in the EO condition. There was a significant interaction effect between condition and proficiency, $F(1, 124) = 8.49, p = .004, \eta^2_p = .06$, indicating that the effects of condition differed in the advanced and intermediate learners.

There was a significant interaction effect between time and condition, $F(1.87, 248) = 27.11, p < .001, \eta^2_p = .18$. This result indicates that the effects of time differed in the CS and EO group in terms of vocabulary learning patterns. Contrasts were performed to break down this interaction. The first contrast for this interaction examined at the pre-test compared to the immediate post-test, comparing the score of the CS group and EO group. This interaction was significant, $F(1, 124) = 41.43, p < .001, \eta^2_p = .25$. The second

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure1.png}
\caption{Interaction between condition and proficiency on the listening comprehension. \textit{Notes.} CS = code-switching; EO = English-only.}
\end{figure}
interaction examined at the immediate post-test and delayed post-test, and was also significant, $F(1, 124) = 26.67, p < .001, \eta_p^2 = .18$. Figure 2 illustrates this pattern.

There was also a significant interaction effect between time and proficiency, $F(1.87, 248) = 5.10, p = .008, \eta_p^2 = .04$. This result indicates that the effects of time differed in the advanced and intermediate level participants. The results of the further contrasts revealed significant interaction effects when comparing the scores of advanced and intermediate participants to the pre-test compared to the immediate post-test, $F(1, 124) = 7.79, p = .006, \eta_p^2 = .06$, and to the immediate post-test to the delayed post-test, $F(1, 124) = 5.05, p = .03, \eta_p^2 = .04$. Figure 3 illustrates the interaction between time and proficiency.

Finally, there was a significant interaction among time, condition, and proficiency, $F(1.87, 248) = 6.30, p = .003, \eta_p^2 = .05$. The first contrast for this interaction compared the pre-test and immediate post-test, when the two learning conditions CS and EO were compared, in advanced and intermediate learner groups. This contrast was found to be significant, $F(1, 124) = 9.03, p = .003, \eta_p^2 = .07$. On the other hand, the second contrast, which compared the immediate post-test and delayed post-test, when the two learning conditions CS and EO were compared, in two different proficiency-level groups, was not significant, $F(1, 124) = .61, p = .44, \eta_p^2 = .01$. Figure 4 illustrates interaction patterns among time, condition, and proficiency. The significance of the first contrast appears to result from the difference between EO intermediate and EO advanced groups, in which

---

**Figure 2.** Interaction between time and condition on the vocabulary learning.

*Notes.* CS = code-switching; EO = English-only.
Lee and Levine

15

the slope of the line between the pre-test and immediate post-test for the advanced EO group is steeper than its intermediate counterpart, showing more gains on the part of the advanced EO group through the intervention. On the other hand, between the immediate post-test and delayed post-test, there is a similar level of decline between two different proficiency groups in the case of the EO group. This drastically contrasts with the CS group, in which the advanced and intermediate learner groups showed almost identical lines which illustrate their performance at three different time points. These findings thus suggest that CS enabled the intermediate learners to perform as well as their advanced counterparts, whereas EO clearly separated the participants in two different proficiency levels of English.

VI Discussion

1 The effects of instructor’s language choice

The present study considered three related research questions that aimed to identify the effects the instructor’s language choice on the acquisition and retention of vocabulary, and on listening comprehension. First, lexical focus-on-form, in whatever language, does appear to benefit adult EFL learners’ vocabulary learning during a focus-on-meaning activity, in this case a classroom listening task. The study found significant differences between the two experimental and the control groups. This finding aligns both with previous intervention studies (Tian & Macaro, 2012; Zhao & Macaro, 2016), as
Figure 4. Interaction between time, condition, and proficiency on the vocabulary learning.
Notes. CS = code-switching; EO = English-only.
well as with the limited number of studies in which scaffolded assistance is provided to learners during task completion (e.g. Al Masaeed, 2016). In comparing CS and EO conditions, CS was more effective for learning of target vocabulary. Thus, with regard to the overall effects, these findings add to what we have learned from earlier studies (Lee & Macaro, 2013; Zhao & Macaro, 2016), suggesting that principled use of the L1 aids in the learning of new vocabulary, even when that target vocabulary is known to be challenging for L2 learners.

A possible explanation for the aforementioned advantage of the CS for L2 vocabulary learning comes from previous psycholinguistics literature on the bilingual lexicon (e.g. Jiang, 2000, 2004; Kroll & Stewart, 1994), in which L2 lexical processing has been found to be influenced by L1 lexical entries. Based on the results of his study, Jiang (2004) suggests that ‘when the meaning of a L2 word is understood, it is the preexisting meaning [in L1 lexicon] or concept that is activated and mapped to the word form’, which is a natural way of accessing L2 words (p. 426). In other words, L2 learners may be predisposed to comprehend the meaning of an L2 word through translating it into the L1, by virtue of the link between the L1 word and its concepts (Kroll & Stewart, 1994). In a classroom setting, this means that a teacher’s CS using the learners’ L1 would facilitate this linking process, which in turn would reduce the learners’ cognitive load in noticing and registering the target lexical item. The learners could then use available cognitive resources to develop a new lexical entry for the L2 word and possibly also transfer its L1 lemma (i.e. syntactic and semantic) information to this entry (Jiang, 2000). Of course, as Jiang (2004) rightly notes, this advantage of CS over L2-only instruction should not suggest that L2 input should not be used in teaching L2 vocabulary; rather, it indicates that ‘L1 information can be additionally beneficial in recalling [L2] words’ (Hennebry et al., 2017, p. 295). Also, the results of this study and previous ones (e.g. Lee & Macaro, 2013; Tian & Macaro, 2012) have shown, the advantage of CS might be diminished in the long term, which speaks for a judicious, principled use of the L1 toward optimizing L2 use (Macaro, 2009).

The second research question aimed to uncover some of the interaction between teachers’ language of instruction and students’ proficiency level on the learners’ comprehension of the target listening texts in the classroom. We found no statistically significant effect between the CS and EO groups for condition. However, when the interaction between proficiency and condition was given close attention, it was found that intermediate learners benefited significantly more from teachers’ CS, enabling them to have a similar level of comprehension to the advanced learners. Simultaneously, the intermediate learners’ comprehension was rather disadvantaged in the EO condition, suggesting that teachers’ CS may be particularly useful to this level of learners.

The study yielded compelling and interesting answers to the third research question, considering the interaction effects between the teachers’ language of instruction—CS or EO—and proficiency level—intermediate or advanced—on the EFL learners’ acquisition and retention of English phrasal verbs. It was found that the intermediate-level learners who received assistance in L1 Korean during the listening tasks benefited more in the long term than those who received English-only assistance. More importantly, the intermediate-level learners, when provided with L1 focus-on-form assistance, were found to gain a similar level of target vocabulary knowledge as their advanced counterparts. But with
advanced learners, the language choice did not appear to matter as much, which accords with the finding of Lee and Macaro (2013) with adult and young EFL learners. This is presumably due to a stronger link between concept and L2 lexicons among advanced learners (Kroll & Stewart, 1994), who may need L1 input to a lesser extent in comprehending L2 lexical items, compared to learners with lower levels of L2 proficiency. That said, the findings from the second and third research questions converged in support of the use of CS for intermediate-level learners, not just for comprehension of target texts, but also for vocabulary learning. It would be helpful to those involved in teacher education, and of course teachers themselves, if further research were to uncover more details of such differentiation between level of proficiency and the effects of language choice, even among different age groups.

2 Implications for language pedagogy and teacher education

The arguably uncontroversial goal of classroom language learning, whether in preparation for a university entrance exam or study abroad, is some level of communicative competence, the ability to understand and speak the new language in a range of real-world settings and with people in everyday interaction. The ability to comprehend recorded language, such as that heard in audio and video media, is therefore crucial. Additionally, beyond the usual role of listening in L2 instruction for comprehension of the new language, the study points toward the usefulness of explicit vocabulary development through listening texts. The intuitive mode of teaching with listening texts of the sort used in this study is to ‘remain’ in the L2 entirely, though numerous scholars have espoused approaches that may entail the use of the learners’ L1 (e.g. Vandergrift, 2004). The present study has shown, though, that learners’ short-term acquisition and longer-term retention of new lexical items, as well as their overall comprehension of the listening texts, benefits from targeted, principled use of the L1 by the teacher. The findings thus raise a few important implications for language pedagogy and teacher education. First, language teacher education programs should include readings and materials that have future teachers learn about, reflect upon, and analyze the various uses for the L1 in their teaching practice. The fact that the CS practices in the present study were targeted and principled cannot be understated; the unprincipled use of the L1 in the L2 classroom may set students at ease, but it likely does not provide the sort of scaffolded assistance that benefits acquisition of new vocabulary and language forms.

A further implication is for the design and creation of teaching materials, as well as syllabus design. The materials created for the present study necessarily considered the ways that the teacher would engage in CS for vocabulary learning and listening comprehension. We see no reason for the creators of instructional materials, whether publishers or individual teachers for their classrooms, could not likewise attend to this facet of the teaching and learning process. In other words, while the ultimate aim of such support for comprehension is to help students move toward dealing with authentic spoken language in the world, the results of the study suggest that such pedagogical texts with planned L1 explanations have a usefulness in fostering both vocabulary learning and listening comprehension. Yet we would also suggest that even with authentic recordings, teachers could provide similar support for listening comprehension using L1
explanation, built into the design of the materials and lessons. Particularly at the inter-
mediate level, providing L1 explanations of key vocabulary items by moving from care-
fully planned texts toward authentic ones might ease the typical difficulties of
comprehension, with L1 explanations providing scaffolding that over time can help
learners make ever more effective use of contextual cues or previous knowledge toward
understanding the text.

Implications for teaching practice also emerge from the present study. The study
entailed using several listening texts and activities in a controlled fashion over the course
of a semester, but in the ‘normal’ progression of a language class the teacher regularly
uses listening passages of various sorts. The findings of our study point toward the provi-
sion of similar CS assistance for vocabulary acquisition and listening comprehension,
returning repeatedly to target lexical items. While our study showed a notable drop in
performance from the immediate to the delayed post-test, such spiraling that would
include ongoing, principled CS may help learners to retain target vocabulary better over
the long term, and to cumulatively improve their listening comprehension.

3 Limitations and further research

As indicated earlier, the results of this study should be viewed with some caution, in part
because tests of assumptions of normality and heterogeneity of variances were found to
be violated for the delayed post-test. But of course, other factors also speak for not over-
stating the benefits of teacher CS for vocabulary acquisition and listening comprehen-
sion identified in this study, while also pointing toward interesting directions for further
research. First, only two proficiency groups/levels were included. It would be helpful to
include introductory-level learners, though with EFL in Korea as well as in many other
locations this would mean school-age learners. In addition, this study involved learners
who share a common L1. For future research, it would be interesting to conduct a similar
study in classes in which the students may share the dominant language of education but
come from different language backgrounds. Finally, the present study necessarily ana-
lyzed the means of learners’ performance, without regard to individual differences within
groups. It would be fruitful to consider the individual differences, to explore the variation
within each group through qualitative means and triangulate these with the quantitative
analysis.

VII Conclusion

The challenge to empirically investigate the roles of the learners’ L1 in L2 learning has
been taken up by numerous colleagues, in large part because changes in educational
policy, curriculum design, teacher education, and classroom practice with regard to CS
or L1 use should be based on sound theoretical foundations as well as valid empirical
evidence. The present study aimed to contribute with evidence of one such role, the ways
the teacher’s use of the students L1 can foster acquisition of new vocabulary and facili-
tate general listening comprehension. The implications for language pedagogy and
teacher training were highlighted, but we conclude with the important observation that
this role for the L1 is but one small piece of a much larger puzzle. More work is needed
in the range of semiotic modes at work in the instructed setting to explore the roles the learner’s L1 plays in the process of L2 learning.

Declaration of conflicting interest
The authors declare that there is no conflict of interest.

Funding
This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Notes
1. By communicative approach, we refer to the ‘weak’ version of communicative language teaching (CLT) described by Howatt (1984).
2. Recall that ML in this context was English (i.e. a target language), and EL was the participants’ L1 (i.e. Korean). In this intrasentential CS, English sets the overall structure of the sentence, and meaning of a target phrasal verb in Korean is embedded in the complete English sentence.

References


**Appendix 1. Target texts.**

**The material for instructional session #1**

1. Being Busy

   **Nick:** Hello Tim. I haven’t seen you for ages. Have you *been* a bit *tied up*¹ at work?
   **Tim:** Yes, Nick. I’ve been. I was away for a conference and the work just *piled up*² while I was away. I’ve *been snowed under*³ ever since. Although I *slave away*⁴ till seven at night and even work at weekends, I still can’t seem to catch up with everything!
   **Nick:** Oh! That sounds terrible. But if you don’t *ease off*⁵ at work, you will make yourself ill.

2. Worked hard

   **James:** How are things going for you at work now, Rachel?
   **Rachel:** Well, James. We’re *branching out into*⁶ a new product line – children’s clothes – so I’m pretty busy. At the moment I’m working on some new designs for leisurewear. We have *flung ourselves into*⁷ having a complete range for children of all ages. It’s quite difficult but I’m sure I’ll develop some good designs if I *keep at*⁸ it.
   **James:** Well, Rachel, you always were good at sticking at things.
   **Rachel:** Well, I’d certainly rather do a project myself than have to *chase down*⁹ other people to make sure they’re doing what they promised to do. But sometimes I am not sure I am *cut out for*¹⁰ this job.
   **James:** Do you think you could *squeeze in*¹¹ lunch with me sometime?
   **Rachel:** Mm, maybe, but not until next week.
   **James:** That’s fine. But don’t try to *wriggle out of*¹² it at the last minute!

---

*Short text extract not exceeding 400 words adapted from: Michael McCarthy, Felicity O’Dell, English Phrasal Verbs in Use Intermediate © Cambridge University Press 2004*
The material for instructional session #2

Robin and Chris are twin brothers. They have a rich grandfather who offered to put them through university. Robin flew through his exams at school, got good grades, and easily got into a good university. Chris wanted to go to the same university, but it was harder for him to get in. However, in his last few months at school, he buckled down to his studies and managed to get through all the necessary exams. Both brothers wanted to major in chemistry. Robin had worked in the science lab before and this experience was counted towards his degree. Chris spent his holidays playing sports. Robin continued to study hard and soon left other students behind. At the end of six months the professor moved Robin up to the higher level class. Robin was in this group but Chris was not. The professor felt Chris was more interested in soccer than chemistry and this counted against Chris. He was marked down for careless mistakes in a number of his essays and for the fact that his point about the main subject crops up in so many places in one essay. The professor also thought that Chris just spent time trying to mug up for exams. By the end of the year he had fallen behind the rest of his year. Now, to get higher grades next semester, he is now burying himself in academic textbooks. Obviously, he needs to pull his socks up now.

Short text extract not exceeding 400 words adapted from: Michael McCarthy, Felicity O’Dell, English Phrasal Verbs in Use: Advanced © Cambridge University Press 2007