



Effect of using texting on vocabulary instruction for English learners

Jia Li, *University of Ontario Institute of Technology*

Jim Cummins, *Ontario Institute for Studies in Education, University of Toronto*

Abstract

This article reports on a study that examined the effectiveness of an intervention using text messages to enhance the academic vocabulary acquisition of English language learners (ELLs). With a random control trial design, we compared students' learning gain of target vocabulary (direct effect) and its subsequent impact on academic vocabulary learning (transfer effect) with and without the intervention treatment. The study included 108 undergraduate ELLs in a large Canadian university in Ontario. The intervention was aligned with the lesson plans of two comparable content-based courses on English for academic purposes required for the ELLs and aimed at teaching frequently used academic words embedded within the assigned course readings. The results indicated that, with the intervention, students learned significantly more target words. However, there was no difference between the treatment and control groups on academic vocabulary post-test performance measuring the transfer effect. The pedagogical implication of the findings and suggestions for future research are discussed.

Keywords: *English Language Learners, Text Messages, Academic Vocabulary Acquisition, Intervention Study*

Language(s) Learned in This Study: *English*

APA Citation: Li, J., & Cummins, J. (2019). Effect of using texting on vocabulary instruction for English learners. *Language Learning & Technology*, 23(2), 43–64. <https://doi.org/10.125/44682>

Introduction

Teens and young adults have been at the forefront of the rapid adoption of text messages (Anderson & Rainie, 2012; Steeves, 2014). A Pew Research Center publication reported that Americans aged 18–24 texted most frequently (Lenhart, 2010). In six nation-wide surveys of Norwegians, a peak in text message usage was found between the ages of 19 and 21 (Ling, 2010). Furthermore, a recent survey revealed that about 96% of college students in the US owned a cellphone and that text messaging was the most-widely used feature, with an average of 60 texts sent by young adults per day (Smith & Page, 2015). In the past several years, there has been a shifting landscape in texting with the quick adoption of mobile technologies among non-native English speakers. Texting has become enormously popular as an important means of written communication among school-aged and college students worldwide, many of whom come from different first language (L1) backgrounds (e.g., Kasesniemi, 2003; Spagnolli & Gamberini, 2007).

Because of the unprecedented acceptance of texting within the youth population, including the avid use of text messages by non-native English speakers and English language learners (ELLs), researchers and educators have endeavored to integrate texting into language instruction and self-regulated learning interventions to help second language (L2) students learn different aspects of language (e.g., Cavus & Ibrahim, 2009; Hayati, Jalilifar, & Mashhadi, 2013; Kennedy & Levy, 2008). However, in a systematic review of the available research on intervention studies that used SMS or MMS messages to teach L2 vocabulary, we found several methodological shortcomings. These included the short duration of interventions, isolated target words, a lack of statistically reliable measures of learning outcomes, and limited or no reports on the scientific basis of the words taught. To address these issues, the present study

examined the effectiveness of an intervention that we developed to teach ELLs' frequently used academic vocabulary—target vocabulary—within a learning context using text messages. In particular, it compared ELLs' learning outcomes with and without the intervention treatment in target academic vocabulary learning (direct effect) and its subsequent impact on general academic vocabulary learning (transfer effect).

Literature Review

A growing body of empirical research has examined the effect of texting on vocabulary learning for English as a foreign language (EFL) learners at the high-school and university levels. In a thorough literature search, we located 16 experimental or quasi-experimental studies published from 2005 to 2016 in peer-reviewed journals in English that examined the effect of SMS, MMS, or email messages via mobile phones on vocabulary or idiom learning. At a minimum, these studies included either (a) a pre-test and post-test for the treatment group or (b) a post-test for both the treatment and control groups. Coincidentally, all of the 16 studies were conducted in EFL settings such as Japan, Hong Kong, Taiwan, Cyprus, China, Iran, Thailand, and Turkey (see [Appendix](#)). Except for the lab experiments of Chen, Hsieh, and Kinshuk (2008), 15 of the 16 studies were intervention studies, most with a duration of two to four weeks. With the exception of the study of Thornton and Houser (2005), who taught students vocabulary on mobile phones using short email messages that were comparable to text messages, 15 of the 16 studies directly examined the effect of SMS ($n = 12$) or MMS ($n = 3$) messages on EFL students' learning of English vocabulary ($n = 13$) or idioms ($n = 2$).

The Effect of Texting-Based Instruction on L2 Vocabulary Acquisition

Previous intervention studies indicated an overall positive trend of vocabulary gains using texting to support EFL students' vocabulary learning. Of the 15 intervention studies, 14 reported significant learning gains after the intervention using SMS, MMS or short email messages via mobile phone to support vocabulary or idiom learning. The only exception was a 7-week study conducted by Derakhshan and Kaivanpanah (2011) with Iranian freshman that reported no significant differences in students' vocabulary gains between the treatment and control groups. Of the 14 studies reporting learning gains, 13 showed significant results—direct effects—supporting texting- over paper- or web-based instruction. The only exception was the study of Lin and Yu (2016) that compared the direct effect of messages with different text, audio, and visual content. There were two studies without control groups that showed significant gains in post-tests compared with their pre-tests (Cavus & Ibrahim, 2009; Song, 2008) and 11 studies that showed significant gains when comparing the post-test or delayed post-test scores of treatment and control groups. In these studies, the control groups often learned target words or idioms through paper materials or web-based learning activities (see [Appendix](#)). We were unable to locate any study examining transfer effects.

In their previous work, our colleagues in EFL settings have taken the initiative and have provided much insight into this increasingly important area of texting-based instruction on L2 vocabulary acquisition. Their results, however, fail to reach a consensus on the effect of texting on L2 students' vocabulary acquisition that is applicable to diverse L2 learners. In addition to the exclusive EFL contexts noted above, only eight existing intervention studies clearly indicated participants' proficiency levels, ranging from elementary to intermediate (see [Appendix](#)). The other studies made no explicit report on their participants' proficiency levels. There is an urgent need for studies to explore innovative instruction and learning support to help youth ELLs in English-speaking countries, as a large number of domestic and international learners of English with varied levels of language skills are avid text-users and encounter tremendous challenges in meeting the requirements of academic English, including academic vocabulary knowledge (e.g., Douglas, 2010).

Most of the existing research is plagued by serious methodological shortcomings. The majority involves a small and convenient sample size. Of the 15 intervention studies, only four clearly indicated the use of randomly assigned samples (Derakhshan & Kaivanpanah, 2011; Motallebzadeh & Ganjali, 2011; Motallebzadeh, Beh-Afarin, & Rad, 2011; Suwantarathip & Orawiwatnakul, 2015). The rest of the studies

either reported using convenient samples or did not report on the procedure for participant selection or group assignment. Of the 15 studies, 10 had a sample size under 50, including both treatment and control groups combined (see [Appendix](#)). Small sample sizes often decreased statistical power (i.e., the probability to correctly detect an intervention effect) and had an influence on the reliability of the results (Kline, 2004). Although all studies reported p values informing whether treatment effects existed, only one study (Hayati et al., 2013) reported Cohen's d measuring the magnitude of the standardized treatment effect (Wilkinson, 1999). Moreover, these studies lacked reliable measures to establish baseline comparability between treatment and control groups in terms of variations in students' language proficiency and other demographic characteristics (e.g., L1, age, gender, and duration of English instruction received). None of the 13 intervention studies which included a control group reported the demographic comparability between the control and treatment groups. Only four of the 10 studies which administered comparability measures on students' language proficiency levels reported internal consistency reliability tests (Motallebzadeh & Ganjali, 2011; Motallebzadeh et al., 2011; Saran, Seferoğlu, & Çağiltay, 2012; Suwantarathip & Orawiatnakul, 2015). Additionally, statistically rigorous measures to assess L2 learners' learning outcomes because of the texting-based intervention were also lacking. All of the 16 studies we reviewed employed self-developed, mostly, multiple choice questions, and only six studies reported internal consistency reliability indices for the tested items in the self-developed measures. Additionally, no studies we reviewed reported inter-rater reliability tests for data coding consistency (see [Appendix](#)).

Furthermore, previous results on the long-term effect of using texting-based instruction on word retention appeared to be inconsistent. Only six of the 13 intervention studies with control groups administered both post-test and delayed post-tests. The delayed post-tests took place two to five weeks after the post-tests. One study reported significantly greater vocabulary gains on both post-tests and delayed post-tests, supporting MMS intervention over paper- and web-based instructional materials (Saran et al., 2012). Three reported significantly greater gains in delayed post-tests rather than post-tests, supporting either SMS intervention over self-paced learning using paper materials (Alemi, Sarab, & Lari, 2012; Zhang, Song, & Burston, 2011) or enriched content presentation in text messages (Lin & Yu, 2016). One reported significantly greater gains in post-tests and not delayed post-tests, again supporting SMS intervention over paper materials (Lu, 2008). Finally, one reported no differences in vocabulary gains between the control and treatment groups in either post-test or delayed post-test scores (Derakhshan & Kaivanpanah, 2011). Better results with delayed post-tests over post-tests might partly have been due to the short duration of the interventions, low texting frequencies, or a combination of both.

Lastly, we were unable to locate studies that investigated the transfer effect of vocabulary texting instruction on students' learning of non-target words or other aspects of English knowledge (e.g., reading comprehension skills). Previous studies measured only students' learning gains in target words or idioms. The study with Iranian EFL adult learners by Motallebzadeh and Ganjali (2011) indicated significantly greater gains on the total score of target words and reading comprehension, supporting the SMS intervention. However, separate scores on target words and reading were not reported, and there was no information on whether the reading comprehension questions directly tested the usage of target words embedded in the reading or another aspect of reading comprehension, such as a possible transfer effect. Vocabulary instruction needs to help students master a limited number of carefully selected "high leverage cross content" words (Lawrence, Crosson, Paré-Blagoev, & Snow, 2015, p. 5), so that they can become autonomous learners capable of learning new words on their own—a transfer effect. As Biemiller points out, there is a "need for planned introduction and explanation of vocabulary plus various tools to help children become more independent in dealing with new vocabulary" (2001, p. 27). This is equally applicable to learners of any age. It is critical to understand how much instruction and what type of instruction text messages can deliver to help L2 learners independently learn vocabulary beyond just target words.

Major Intervention Design Features and Texting Content

A few key design features emerged from the intervention studies examining the effect of vocabulary learning using text messages. The majority of the interventions used the push model, a one-way

communication “where teachers [controlled] the frequency and the timing” (Stockwell, 2010, p. 96) as well as the content of messages sent to learners. With three exceptions, the EFL learners sent their instructor messages with sentence construction, fill-in-the-blank, and writing exercises using target words and exchanged target word sentences with their peers (Derakhshan & Kaivanpanah, 2011; Suwantarathip & Orawiwatnakul, 2015; Tabatabaei & Goojani, 2012). Though students in some studies said they preferred two-way text messaging, interacting with their teachers and peers to create a more engaging and personalized learning experience (Cavus & Ibrahim, 2009), the push mode proved effective in providing students with a structured, practical, and convenient learning routine. This consistently enabled a small, viable amount of vocabulary knowledge input and led to significant learning gains, as was reported in the L2 reviewed intervention studies using the one-way communication push model. The push model can send well-designed “multicasting messages to a group of mobile users with a common profile, thereby improving the effectiveness and usefulness of the content delivered” (Motiwalla, 2007, p. 585). Moreover, it is easier to implement than two-way interactive texting among a large number of learners. It can also better ensure the implementation fidelity of the interventions.

Some of these studies’ intervention designs explicitly emphasized repeated exposures to target words, which conformed to a key instruction principle proven to be effective for L2 learners’ vocabulary acquisition (Hulstijn, 2001; Nation, 2001; Schmitt, 2008). For instance, Thornton and Houser (2005) sent Japanese university students three mini-lesson messages daily in multiple exposures: introducing a word, reviewing the word, and incorporating the word into a story. Cavus and Ibrahim (2009) divided 48 messages into three clusters of 16 messages and sent them to students three times over nine days.

Frequent incremental vocabulary instruction (Barcroft, 2012) at spaced intervals (Braun & Rubin, 1998) is another important cornerstone feature for some of the intervention designs. Of the 15 intervention studies, six sent student messages twice or more per day (see [Appendix](#)), often during the day time, when students were more receptive to text-based instruction. The highest texting frequency for spaced vocabulary instruction was performed by Cavus and Ibrahim (2009) who sent their participants one email message via mobile phone, at a pre-determined optimal interval, every half hour between 9:00 a.m. and 5:00 p.m. In the study by Zhang et al. (2011), university students received two messages per day, one at 12:00 p.m. and the other at 5:30 p.m. Overall, these studies demonstrated the promise of texting to provide students with repeated vocabulary exposures in structured and spaced intervals. This was more effective in supporting L2 learning than “massed practice” (Seabrook, Brown, & Solity, 2005, p. 107); however, most of the studies were missing information regarding the texting time and frequency.

As to the content of text messages, the existing research seemed to be in agreement, focusing on teaching students word meanings by providing word definitions in English or the L1 and sample sentences—“discrete chunks readable on the tiny screens of mobile phones” (Thornton & Houser, 2005, p. 221). This was not only feasible, but also effective, given that 13 of the 14 intervention studies with significant learning gains included definitions or sample sentences in their text messages. It is also worth noting that some studies, to varying degrees, emphasized teaching target words in context (e.g., Song, 2008; Suwantarathip & Orawiwatnakul, 2015; Thornton & Houser, 2005). Song (2008) introduced students to new words through personal experiences in the UK on four topics (nightlife, food, weather, or travel). Each topic covered four days’ learning content, focusing on four to six target words or expressions per day. A website that aimed to increase students’ exposure to target vocabulary also provided students with content resources, such as an online dictionary and vocabulary and cultural tips. Meanwhile, target words or expressions with brief explanations in Chinese and English were sent to students through SMS. However, most interventions in these studies taught words in a decontextualized and isolated manner, only providing sample sentences and definitions in the L1 or English. This, in turn, diminished students’ repeated exposure to target words in context as well as their opportunities for further incidental learning of target and other words in authentic contexts. It is reasonable to assume that if an intervention incorporates students’ other language learning activities, directly supports classroom instruction, and teaches students words that meet their immediate needs (e.g., assigned course readings), the intervention may effectively motivate students as well as bridge their attentive vocabulary learning through explicit instruction in various contexts.

Finally, among the studies we reviewed, the intervention with the longest treatment duration (16 weeks) had the lowest texting frequency (two messages per week; Alemi et al., 2012). The intervention with the shortest duration (nine days) had the highest texting frequency (16 messages per day with one message every 30 minutes; Cavus & Ibrahim, 2009). Other intervention studies had either a relatively short duration or low frequency of text messages, so they lacked an optimal combination of duration and frequency. Examples include two messages per week for four weeks (Lin & Yu, 2016), two messages per day for two weeks (Lu, 2008), and four messages per day for 20 days (Hayati et al., 2013). According to research on learning English as an L1, a single exposure to an unknown word results in a 10% to 15% chance of learning its meaning (see Nagy, Herman, & Anderson, 1985; Swanborn & Glogger, 1999). For L2 learners, more exposure is required to acquire a word (Nation, 2001). Thus, limited treatment duration and texting frequency may be insufficient for students to achieve long-term retention of target words—something that was not measured in most of the existing studies. The transfer effect (i.e., vocabulary learning gain beyond target words or in reading comprehension) has not been investigated.

Research Questions

These studies have indicated the great potential of texting in facilitating EFL learners' vocabulary and idiom learning in a self-regulated format that can occur anytime, anywhere. However, beside the issues raised above regarding the research methods and intervention designs, previous EFL research has focused exclusively on learners with low to intermediate English proficiency levels. In the present study, we tried to overcome these limitations and examined the effect of an intervention using texting on university ELLs' learning of contextualized academic vocabulary. These ELL students had an advanced-low language proficiency level (American Council on the Teaching of Foreign Languages, 2012). The study answers the following two research questions:

1. What is the (direct) effect of the intervention using text messages on students' learning of target academic vocabulary?
2. What is the (transfer) effect of the intervention using text messages on students' learning of general (non-target) academic vocabulary?

Methods

Upon receiving ethics approval from two universities (one by the lead researcher's affiliated university and the other by the institution where the study was conducted), the students registered in six content-based classes of English for academic purposes (EAP) and their instructors at a large Canadian university in Ontario were invited to participate in the study. The students were provided with an informational invitation letter about the project and a consent form with their instructors present. It was explicitly indicated that student participation was voluntary and that they could withdraw from the study at any time and without any impact on their grades.

Participants

Participants included 108 (49% female) undergraduate students (ages 18–25) from the EAP classes. The L1s spoken by the participants were Mandarin (51%), Cantonese (14%), Russian (7%), Arabic (7%), Farsi (7%), Korean (5%), Turkish (5%), Spanish (2%), and Lithuanian (1%). The duration of their residence in Canada varied from two months to 13 years ($M = 2.56$ years, $SD = 2.90$). The years of English instruction students had received previously ranged from two to 15 years ($M = 9.27$ years, $SD = 3.57$). The academic majors of participants included economics, finance, business, accounting, computer science, biology, communication studies, psychology, actuarial science, English, law and society, and sociology. Based on the university admission requirement for English language competence, the participants met the requirement at the lower end (i.e., 80+ for the iBT, 6+ for the IELTS) and often demonstrated difficulties in reading comprehension of academic texts due to their limited academic vocabulary knowledge. It was a challenge to engage the class in meaningful discussions unless difficult academic words were explained in simpler English or in their L1s (e.g., Chinese, which a Chinese-English bilingual instructor was able to do).

Consequently, the students' inability to support their arguments with well-integrated references and paraphrases created barriers to effective academic essay writing. Overall, their performance in the EAP program tended "to be uneven" (American Council on the Teaching of Foreign Languages, 2012, p. 6) across topics and language skills.

The Intervention Design

The intervention, Word Matters, aimed to teach university undergraduate English language learners "high leverage" (Lawrence et al., 2015, p. 5) academic words in their assigned readings using text messages. It was expected that this would increase their academic vocabulary knowledge, helping them approach the reading comprehension threshold required for an academic text (Nation, 2001)—familiarity with an estimated 95% of the vocabulary (Laufer & Ravenhorst-Kalovsk, 2010)—and, in turn, assisting their learning of other academic words encountered in academic texts.

We used a randomized experimental design with pre-test and post-test data. The six classes, in which participants were registered in two comparable EAP courses and taught by three instructors, were randomly assigned to the treatment ($n = 48$, 44.4%) and control groups ($n = 60$, 55.5%), with three intact classes in each group. The intervention lasted for nine weeks.

As part of the curriculum of the courses, students in both treatment and control groups were required to read the same assigned readings and engage in class discussions of the readings with teachers' instruction. Students in the control group were asked to use online dictionaries and dictionary apps with their own devices to check the meanings of target words and sample sentences. This design aligned with the control conditions in the studies previously discussed using dictionaries (e.g., Alemi et al., 2012) or web-based materials (e.g., Saran et al., 2012; Thornton & Houser, 2005) and self-paced learning (e.g., Zhang et al., 2011). The control condition, which could be facilitated easily in class, combined instructors' explanations of words during class discussion with students learning the words by reading hard copies of articles or a novel. This condition was similar to the learning activities that often took place in the intact setting.

The participants in the treatment group received an orientation email at the beginning of the project explaining the purpose of the intervention and the detailed information they would be receiving via text messages. The intervention included sending students information about three words carefully selected from the required reading materials each day through text messages: one at 10:00 a.m., one at 4:00 p.m., and one at 8:00 p.m. Each text message included a target word, the title of the assigned reading where the word appeared, the page reference of the target word in the reading, the part of speech, the word's definition, and a sample sentence (see [Figure 1](#)). Every evening at 8:30 p.m., students also received an email message summarizing the three words sent during the day and containing a small quiz on the words learned a week before (see [Excerpt 1](#)). At the end of each week and month, the students received a downloadable summary of the words they had been sent for their own future reference. During the nine week intervention, a total of 189 target words were sent to students through text messages, with three words sent daily, or 21 words a week. Prior to the intervention, a trial text message and email message were sent to the students in the treatment group who were required to confirm their receipt of the messages by texting and email. Students reported that on average they read three text messages four days a week in the post-intervention survey, which results are reported in a separate article focusing on students' perceptions and experience of the intervention (Li, Cummins, & Deng, 2018).

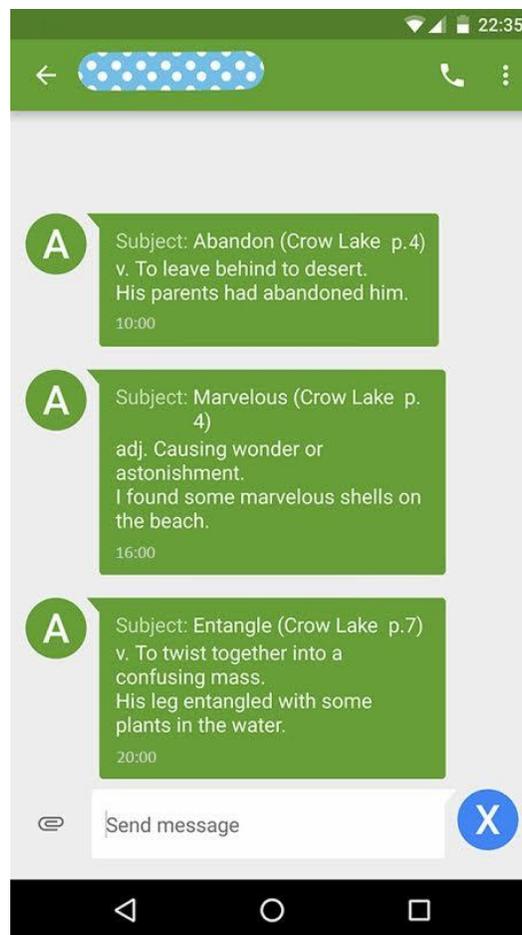


Figure 1. Sample text messages

Excerpt 1. A Sample Daily Email

Hello everyone,

How are you doing today? It's super cold in Boston these days and I believe it must be even colder in Canada.

Keep warm, friends!

Let's see what we have for today:

principle (from Crow Lake, p. 23)

n. A rule that teaches you what is right and wrong and that influences your actions.

It's against my principles to cheat.

glean (from Crow Lake, p. 23)

v. To collect bit by bit.

At present we're gleaning information from all sources.

devotion (from Crow Lake, p. 23)

n. A feeling of strong love or loyalty.

She has cared for the poor with selfless devotion.

~~~~quiz~~~~word game~~~~quiz~~~~word game~~~~quiz~~~~word game~~~~

Do you still remember these words?

Complete the following sentences with the words you learned last week. Write your answers on a piece of paper and check tomorrow's email for the answers!

1. She claims that she can \_\_\_\_\_ future events.
2. If war broke out, it would be \_\_\_\_\_ for the whole world.
3. I would like to offer a historical \_\_\_\_\_.

Answer for yesterday's quiz: 1. abandon; 2. marvelous; 3. entangle

Hope you enjoy learning these words today!

All the best,

Mei

### Data Collection Procedure

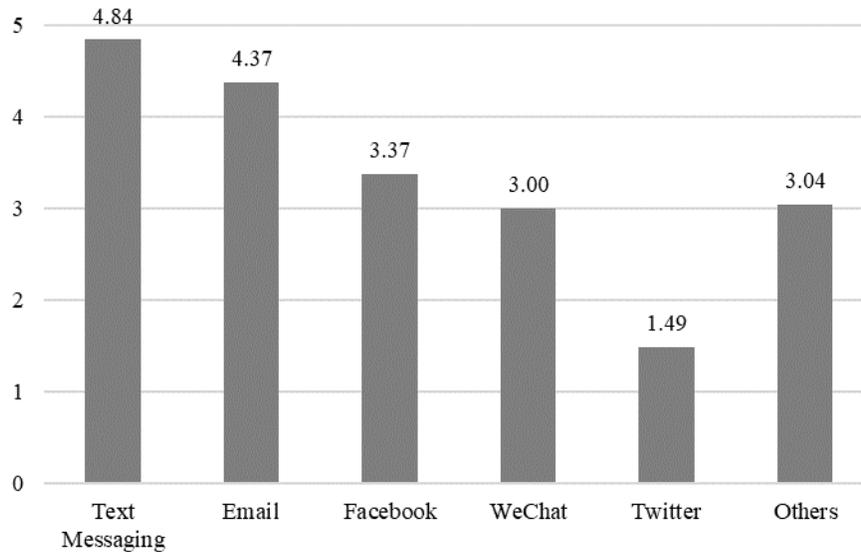
All the participants in both the treatment and control groups were informed about the research procedures, including learning conditions, number of tests, test times, and length. One week before and after the intervention, both the treatment and control groups received a pre-test and post-test. To ensure accurate comparable difficulty levels between the pre- and post-measures, the same vocabulary tests were administered. These included a 60-item target vocabulary test developed by the research team based on a selection of academic words from assigned readings that were texted to students during the intervention, and a 30-item (non-target) general academic vocabulary test adapted from the vocabulary levels test (academic level) by Schmitt, Schmitt, and Clapham (2001). All non-target words had appeared in assigned readings. The participants were administered a pre-intervention survey that focused on their demographics and technology use. A post-intervention survey and interviews were also conducted to collect information on students' learning behaviors during the intervention and their feedback on the intervention; those results were reported in a separate article (Li et al., 2018). The pre- and post-vocabulary tests and the pre-intervention survey were administered to the treatment and control groups in the classroom by the lead researcher, a research assistant, or course instructors. Students were given 30–40 minutes to complete the pre-tests or post-tests. It took about 15–20 minutes for students to complete the surveys.

### Instrument Development

#### *Pre-Intervention Survey*

In order to develop a feasible and effective intervention, we first developed a survey to identify technology applications that students often used and that were aligned with student interests or preferences. The survey consisted of two parts: Part 1 asked for demographic information (i.e., major, L1, years in Canada, country of origin, years of English instruction) and Part 2 asked participants to report the frequency of their technology use (i.e., email, text messaging, Facebook, Twitter, and other social media). The research team developed the survey in consultation with seven ELL students registered in the courses in the previous and current school terms and three of their instructors. Three rounds of revisions were made in accordance with their suggestions and comments.

The survey results helped the team design a user-friendly intervention. Participants reported their frequency of using technology on a 5-point Likert scale (i.e., 1 = never, 2 = once a month, 3 = once a week, 4 = once a day, and 5 = more than once a day; see [Figure 2](#)). They reported that the average frequency of using text messages was approaching more than once a day. They used text messages ( $M = 4.84$ ,  $SD = 0.43$ ) more often than email ( $M = 4.37$ ,  $SD = 0.69$ ), followed by Facebook ( $M = 3.37$ ,  $SD = 1.68$ ), WeChat ( $M = 3.00$ ,  $SD = 1.89$ ), and Twitter ( $M = 1.49$ ,  $SD = 1.02$ ).



*Figure 2.* Frequency of technology use by participants before intervention

### **Target Academic Vocabulary Selection**

Prior to the study, the lead researcher and author had taught the EAP courses mentioned above for two years and observed the challenges that ELL students with an advanced-low competence level encountered when processing assigned readings.<sup>1</sup> The challenges were often due to students' difficulties in comprehending the meanings of academic and low-frequency words (Coxhead, 2000; Nation 2001). These words also appeared in their assigned reading book, *Crow Lake* (Lawson, 2003), corresponding to a ninth-grade reading level for native-English speakers. Our observation was later confirmed by the results of student pre-vocabulary tests.

The research team used Cobb's (2016) [VocabProfilers](#) to identify academic and low-frequency words and determine the number of occurrences in the assigned readings. In order to optimize students' vocabulary learning and support their reading comprehension, the intervention focused on academic words and difficult (low-frequency) words that met three selection criteria: (a) they were critical for the comprehension of the assigned readings (e.g., *inhabitant*, *disguise*, *segregate*), (b) there were multiple occurrences of the words in the assigned readings, and (c) they were likely to be encountered frequently by students in a variety of academic domains (e.g., *predict*, *perspective*, *demonstrate*), as recommended by research experts in the area of vocabulary, language, and literacy development (e.g., Baumann & Graves, 2010; Nation, 2001; Snow & Uccelli, 2009; Stahl & Nagy, 2006).

A total of 200 vocabulary items were selected from the required readings on Canadian culture, language, and social issues for the courses in which participants were enrolled. These included the book *Crow Lake* (84 target words), a book chapter titled "Aboriginal People" (44 target words; Steckley, 1997), a journal article titled "Managing Homeless Youth in Toronto" (46 target words; Gaetz, Tarasuk, Dachner, & Kirkpatrick, 2006), and a newspaper article titled "Why Can't We Talk" (26 target words; Erasmus, 2002). Of these words, 117 were on the academic word list (AWL; Coxhead, 2000), and 81 were low-frequency words (e.g., *devastate*, *fidget*, *marvelous*). Of the 200 total words, 189 were used as target words for the 9-week intervention. Definitions of the vocabulary and example sentences were adapted from the contexts in which they occurred in the readings or dictionaries. The research team carefully selected and modified the definitions and wrote example sentences, ensuring their relevance to the students' academic and daily lives.

### **Target Academic Vocabulary Test**

The research team developed a 60-item target vocabulary test with multiple-choice questions through random selection from a target pool of 189 words. The test was administered to students both as a pre-test

and as a post-test. Due to a timing conflict with students' final exams, we were unable to administer delayed post-tests. To examine the reliability of the scale, we applied Cronbach's alpha to test the internal consistency of the scale and an intra-class correlation test to examine the inter-rater reliability of the two raters. The internal reliability index for the 60 items was .94, indicating excellent internal consistency of the scale in measuring the students' target vocabulary performance. We made three attempts to achieve high quality coding for all the vocabulary tests. In the end, the inter-rater reliability was .99 with 15% of randomly selected test scores, indicating an exceptionally high rater agreement.

### **General Academic Vocabulary Test**

To assess student participants' general academic vocabulary knowledge, a 30-item academic vocabulary test was adapted from the vocabulary levels test by Schmitt et al. (2001). The original Cronbach's alpha reliability index of the vocabulary levels test was .96. The internal consistency reliability for the present study using Cronbach's alpha was .91 for the scale with 30 items, indicating excellent internal consistency. The inter-rater reliability was .99 using intra-class correlation with 15% of randomly selected test scores, indicating an exceptionally high rater agreement.

### **Data Analysis Procedure**

The analysis involved three steps using the SPSS statistical package, Version 21. Step 1 was to test the comparability of the control and treatment groups' existing target academic and general academic vocabulary knowledge prior to the intervention (i.e., condition, instructor, and course). One-way analyses of variance (ANOVAs) were performed on the target academic and general academic vocabulary pre-test scores to assess any statistically significant group differences. Comparability was established as the results revealed no significant group differences for students' target academic ( $F_{(1, 92)} = 0.02, p = .90$ ) or general academic ( $F_{(1, 78)} = 0.08, p = .77$ ) vocabulary knowledge. The results of further analyses found no significant differences in target academic vocabulary for condition ( $F_{(1, 90)} = 0.23, p = .63$ ), instructor ( $F_{(2, 90)} = 0.75, p = .47$ ), or course ( $F_{(5, 90)} = 1.32, p = .26$ ). Similarly, there were no significant differences in general academic vocabulary for condition ( $F_{(1, 78)} = 0.08, p = .77$ ), instructor ( $F_{(2, 90)} = 0.37, p = .69$ ), or course ( $F_{(5, 78)} = 0.17, p = .97$ ). These findings showed the comparability for pre-test results between control and treatment groups, between different instructors, and between different courses. So, the one-way ANOVA was used for post-test data, and the pre-test scores were not used as a covariate.

To test the comparability of student background characteristics yielded from the pre-survey between treatment and control groups, we used Chi-square tests for percentage data and univariate ANOVA tests to compare variable means. The Chi-square tests revealed that the control and treatment groups did not differ significantly in terms of age ( $p = .34$ ), gender ( $p = .37$ ), L1 ( $p = .87$ ), or home country ( $p = .65$ ) distributions. ANOVA tests revealed that the control and treatment groups did not differ significantly regarding years in Canada ( $p = .64$ ) or years of English instruction ( $p = .61$ ) or regarding frequency of using text messages ( $p = .93$ ), email ( $p = .30$ ), WeChat ( $p = .08$ ), Facebook ( $p = .25$ ), Twitter ( $p = .06$ ), or other social media ( $p = .55$ ). Therefore, none of these variables was included as a covariate in ANOVA tests for Research Questions 1 and 2. Next, two separate one-way ANOVA were conducted for the pre-tests and post-tests on target academic and general academic vocabulary scores to examine the intervention effect on students' learning of target academic and general academic words with and without the intervention. Effect sizes were calculated using Cohen's  $d$ .

## **Results**

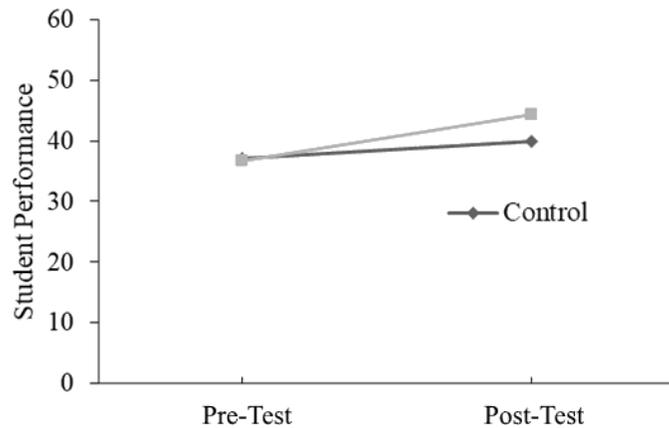
To answer Research Question 1, the ANOVA revealed that the treatment group performed significantly better in their post-tests ( $M = 44.41, SD = 11.77$ ) than in their pre-tests on target academic words ( $M = 36.76, SD = 14.64, F_{(1, 80)} = 6.47, p = .01, \text{partial } \eta^2 = .08$ ). The effect size calculation using Cohen's  $d$  revealed a medium intervention effect ( $d = .58$ ). Participants in the control group, however, performed essentially the same in their post-tests ( $M = 39.85, SD = 12.39$ ) and pre-tests of target academic words ( $M = 37.14, SD = 14.39, F_{(1, 107)} = 1.10, p = .30, \text{partial } \eta^2 = .01$ ; see [Table 1](#) and [Figure 3](#)). In the post-test

of target academic words, the treatment group performed significantly better than the control group ( $F_{(1, 97)} = 4.40, p = .04, \text{partial } \eta^2 = .05$ ).

**Table 1.** Means, Standard Deviations, and ANOVAs for Target Academic Words

| Group     | Pre-Test |       |    | Post-Test |       |    | df  | F    | p   | d   | Partial $\eta^2$ |
|-----------|----------|-------|----|-----------|-------|----|-----|------|-----|-----|------------------|
|           | M        | SD    | N  | M         | SD    | N  |     |      |     |     |                  |
| Treatment | 36.76    | 14.64 | 42 | 44.41     | 11.77 | 37 | 78  | 6.47 | .01 | .58 | .08              |
| Control   | 37.14    | 14.39 | 50 | 39.85     | 12.39 | 58 | 107 | 1.10 | .30 | .20 | .01              |

*Note.* Though there were 48 students in the treatment group and 60 in the control group, not all of them took both tests. For instance, there were 50 participants in the pre-test and 58 in the post-test for the control group.



**Figure 3.** Student performance on target academic vocabulary in pre-tests and post-tests

To answer Research Question 2, the ANOVA for general academic vocabulary test scores suggested that there was not any significant difference for the treatment group between the pre-test ( $M = 25.16, SD = 4.27$ ) and post-test scores ( $M = 24.88, SD = 5.32, F_{(1, 67)} = 0.05, p = .82$ ). Similarly, there was no significant difference in the control group between the pre-test ( $M = 24.80, SD = 5.72$ ) and the post-test ( $M = 23.04, SD = 6.55, F_{(1, 95)} = 1.94, p = .17$ ; see [Table 2](#) and [Figure 4](#)). The findings suggested that no intervention effect was present for general academic vocabulary, which was not instructed in the intervention. In fact, the treatment group’s general academic vocabulary mean score in the post-test was slightly lower than in the pre-test, with a mean difference (*MD*) of  $-0.27$  (pre-test  $SD = 4.27$ , post-test  $SD = 5.31$ ). Similarly, the control group’s general academic vocabulary mean score in the post-test also was lower than in the pre-test, with a *MD* of  $-1.76$ , (pre-test  $SD = 5.73$ , post-test  $SD = 6.55$ ). However, the lower scores were likely caused by random errors within reasonable fluctuation, and these decreases were not statistically significant or meaningful.

**Table 2.** Means, Standard Deviations, and ANOVAs for General Academic Words

| Group     | Pre-Test |      |    | Post-Test |      |    | df | F    | p   | d    | Partial $\eta^2$ |
|-----------|----------|------|----|-----------|------|----|----|------|-----|------|------------------|
|           | M        | SD   | N  | M         | SD   | N  |    |      |     |      |                  |
| Treatment | 25.16    | 4.27 | 32 | 24.89     | 5.31 | 35 | 66 | .05  | .82 | -.06 | .00              |
| Control   | 24.80    | 5.73 | 46 | 23.04     | 6.55 | 49 | 94 | 1.94 | .17 | -.29 | .02              |

*Note.* Though there were 48 students in the treatment group and 60 in the control group, not all of them took both tests. For instance, there were 46 participants in the pre-test and 49 in the post-test for the control group.

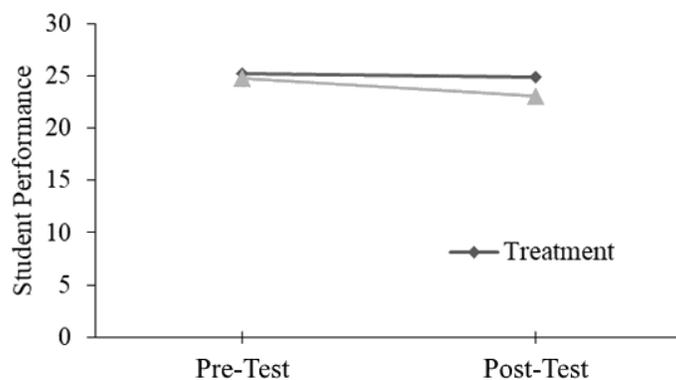


Figure 4. Student performance on general academic vocabulary in pre-tests and post-tests

## Discussion

### Effect of the Push Mode and Texting Content

The present results are consistent with previous studies which show significant learning gains of target words and thus support the use of texting over web-based vocabulary instruction, self-paced learning, and the independent use of dictionaries. In addition, the results reinforce active engagement and positive perceptions of the intervention (see [Appendix](#)). Except for the study by Chen et al. (2008), all the studies we reviewed were intervention studies focusing on measuring the learning gains of target words or idioms. Nine of these studies showed that with texting-instruction students learned significantly more words when compared to their performance before the intervention treatments and to their peers' performance in control groups (see [Appendix](#)). The web- and paper-based control conditions in many of the previous studies shared the characteristics of the self-paced and independent learning (e.g., Alemi et al., 2012; Hayati et al., 2013; Saran et al., 2012; Zhang et al., 2011).

Additionally, two quasi-experimental studies (Cavus & Ibrahim, 2009; Song, 2008) without control groups showed significant vocabulary gains by students after experiencing interventions. As was the case in the present study, eight of the 11 studies mentioned above sent text messages containing vocabulary or idiom instruction to learners. Therefore, the push model is clearly effective in supporting students' target vocabulary or idiom learning (see [Appendix](#)). Our 2-month intervention study showed a significant difference in the learning gain of target words between the treatment and control groups ( $MD = 4.94$ ). This was comparable to some results of previous studies that had durations between two weeks and two months (e.g., Lu, 2008:  $MD = 2.8$ ; Suwantarathip & Orawiwatnakul, 2015:  $MD = 3.83$ ; Zhang et al., 2011:  $MD = 12.5$ ). It is worth mentioning that the present study taught more words than the three studies noted above. Some students in the present study, however, did express a preference for text messages in a two-way interaction. This echoes survey results by Cavus and Ibrahim (2009), who found that students believed that two-way texting instruction would have been more effective than the one-way approach. However, unless an adaptive, intelligent text messaging system can be developed, the push mode is more easily implemented due to the labor intensity required for interactive texting.

The present study, focusing on carefully crafted word definitions and sample sentences using simple language accessible to students, was in line with some previous studies (e.g., Hayati et al., 2013; Lu, 2008), supporting the principle of comprehensible input (Krashen, 1989, 2003). That is, vocabulary teaching content should be clear and comprehensible and avoid complicated explanations with unknown or less-known synonyms or definitions (Nation, 2001; Nation & Newton, 1996). During the interviews, most students in the present study expressed their appreciation for the simplicity and clarity of word definitions and sample sentences. When asked for suggestions for future interventions, one student said, "I'd suggest to do the similar thing ... actually, sending text messages with simple definitions." Hayati et al. (2013) reported on Iranian EFL students' appreciation of concise and helpful content of English idioms in text

messages. Similarly, Lu's (2008) high school students appreciated the clear content presentation and manageability of text messages and said, "The lessons are short and easily to read" (p. 521). On the contrary, Chinese university students in the study by Zhang et al. (2011) complained that long messages (five target words per message) were overwhelming, particularly when they were automatically segmented into mini-messages sometimes sent in the wrong sequence by the system. Overall, the present study is consistent with previous research results and further informs our understanding of the promise of push mode texting to facilitate ELLs' vocabulary learning. It also provides insight for the key principles in developing text message content conducive to maximizing student vocabulary learning.

### **Transfer Effect, Intervention Duration, and Reading Comprehension**

Our study did not find any significant transfer effect of the intervention. There was no significant difference between the control and intervention groups in the learning of academic vocabulary not taught in the intervention. The study by Motallebzadeh and Ganjali (2011) reported a significant gain on the total score of 10 target vocabulary and 20 reading comprehension questions embedded with target words that might have been perceived as a possible transfer effect. However, separate scores on vocabulary or reading questions were not reported, and no information was available on the reading measure. We were unable to locate any studies that reported on the intervention effect beyond the target words and idioms. Despite empirical evidence to the contrary, the qualitative data and observations of the present study, appeared to be supportive of the transfer effect. In the post-intervention survey, students reported that the intervention was helpful for them to learn not only target words, but vocabulary in general ( $M = 3.41$ ,  $SD = 0.85$ ). This disparity between empirical fact and student opinion may be due to the short term of the present intervention that fails to make any statistically significant transfer effect on acquisition of non-target words. Students further reported that the transfer effect of the intervention on their comprehension of assigned readings was somewhat helpful ( $M = 3.03$ ,  $SD = 0.82$ ). This was consistent with feedback collected from the students who participated in interviews. Most were satisfied with the content of the text messages, which they believed was highly relevant to and helpful for their reading comprehension, as one student described in detail:

When you have the page numbers (in the novel) for the words, it is very convenient for me to learn them, I mean, you know, the reading becomes more zhibai (直白, *straightforward*). Because the novel, *Crow Lake*, for some parts, is kind of yinhui (隐晦, *obscure, hard to understand*), when the words are sent to us in this format, I feel it makes it easier for me to understand the reading.

According to Nation (2001), when learning new words incidentally through reading, L2 learners need to know at least 95% of the words in a text, exclusive of proper nouns, to ensure comprehension and accurate guessing. It is plausible that the 2-month duration of the present intervention was still too short to provide students with sufficient repeated exposure to achieve long-term retention of target academic words. Also, the 189 target words taught were probably not adequate to reach the 95% word threshold in students' assigned readings that would enable them to learn other academic vocabulary autonomously through the readings. Text messaging's current technical capacity of 160 alphanumeric characters cannot provide more space for content that would enable students to learn words incidentally through reading paragraph-length messages. Therefore, longer durations with pre-determined optimal intervals and frequencies of texting that facilitate repeated, spaced word exposures and that are compatible with the needs and texting preferences of specific demographics of L2 learners are likely the key to enabling a transfer effect.

### **Conclusions**

The unique contribution of the present study lies in its rigorous experimental design that substantiated the positive results reported in similar previous studies. This design includes a longer duration of the intervention treatment, a larger number of target words, and a randomly assigned, adequate sample size in both treatment and control groups with comparable language proficiency levels. In addition, it measured both the direct and transfer effect of the intervention and students' learning gains through a self-

development target academic vocabulary test and a general academic vocabulary levels test (Schmitt et al., 2001), the results of which were validated by internal consistent reliability tests. Many of these aspects were lacking in the previous studies.

There are some limitations in the present study. First, no measurement of the long-term word retention using delayed post-tests was made due to the time conflict with students' final exams. Future research using rigorous measures is needed to investigate the long-term effects of texting instruction, not only on students' learning of target academic words, but also on their learning of academic and low-frequency words in general. Though the present intervention, which lasted for two months and taught 189 target words, had a longer duration with more target word items than most studies in the area, we believe that the duration was not long enough to cover an adequate number of unknown words that would have enabled students to learn new academic words independently—the transfer effect we expected to occur. In addition, and to our knowledge, no other studies have reported on the transfer effect of texting-based intervention. Longitudinal intervention studies that systematically teach academic words would provide more insight into the pedagogical capacity of texting-based instruction to have a transfer effect on students' overall vocabulary acquisition and to develop reading comprehension skills. Furthermore, future studies on enriched MMS messages are needed, as Chen et al. (2008) and Lin and Yu (2016) found that students who received messages presenting word information in enriched multimedia outperformed those who received messages presenting word information in text-only, sound-only, or picture-only modes.

To sum up, the distinctive features of text messages, such as mobility, easy access, instant communication, and bit-sized modules, have great potential to help students learn vocabulary. The features in themselves, however, do not automatically ensure optimal or motivated academic vocabulary learning among students. In order to effectively enhance students' learning, the design—particularly the content—of text messages must be carefully developed (Levy & Kennedy, 2005) in line with the needs of specific demographic groups of language learners to encourage their frequent reading of text messages that warrant their repeated, spaced exposures of target words.

## Acknowledgements

This study was funded by a Standard Research Grant of the Social Sciences and Humanities Research Council of Canada. We would like to thank all student participants and the research assistants Linying Ji, Yan Wang, Qizhen Deng, and Amale Aldarwish. Special thanks also go to the anonymous reviewers for providing us with detailed suggestions and comments during the revisions of the manuscript.

## Notes

1. At the time of this study, the curricula for the EAP courses was unchanged from when the lead researcher had begun teaching them.

## References

- Alemi, M., Sarab, M. R. A., & Lari, Z. (2012). Successful learning of academic word list via MALL: Mobile assisted language learning. *International Education Studies*, 5, 99–109.
- American Council on the Teaching of Foreign Languages. (2012). *The ACTFL proficiency Guidelines*. Retrieved from [https://www.actfl.org/sites/default/files/pdfs/public/ACTFLProficiencyGuidelines2012\\_FINAL.pdf](https://www.actfl.org/sites/default/files/pdfs/public/ACTFLProficiencyGuidelines2012_FINAL.pdf)
- Anderson, J., & Rainie, L. (2012). *Main findings: Teens, technology, and human potential in 2020*. Washington, DC: Pew Research Center. Retrieved from <http://www.pewinternet.org/2012/02/29/main-findings-teens-technology-and-human-potential-in-2020/>

- Barcroft, J. (2012). *Input-based incremental vocabulary instruction*. Alexandria, VA: TESOL International.
- Basal, A., Yilmaz, S., Tanriverdi, A., & Sari, L. (2016). Effectiveness of mobile applications in vocabulary teaching. *Contemporary Educational Technology, 7*, 47–59.
- Bauman, J. (1995). *About the general service list*. Retrieved from <http://jbauman.com/aboutgsl.html>
- Baumann, J. F., & Graves, M. F. (2010). What is academic vocabulary? *Journal of Adolescent & Adult Literacy, 54*, 4–12.
- Biemiller, A. (2001). Teaching vocabulary: Early, direct, and sequential. *American Educator, 25*, 24–28.
- Biemiller, A. (2009). *Words worth teaching*. Columbus, OH: McGraw-Hill.
- Birjandi, P., Anabisarab, M., & Samimi, D. (2007). *English for pre-university students: Learning to read*. Tehran, Iran: Ministry of Education.
- Braun, K., & Rubin, D. (1998). The spacing effect depends on an encoding deficit, retrieval, and time in working memory: Evidence from once-presented words. *Memory, 6*(1), 37–65.
- Cavus, N., & Ibrahim, D. (2009). m-Learning: An experiment in using SMS to support learning new English language words. *British Journal of Educational Technology, 40*(1), 78–91.
- Chen, N. S., Hsieh, S. W., & Kinshuk (2008). Effects of short-term memory and content representation type on mobile language learning. *Language Learning & Technology, 12*(3), 93–113.
- Cobb, T. (2016). VocabProfilers. Retrieved from <http://www.lex tutor.ca/vp/comp/>
- Coxhead, A. (2000). A new academic word list. *TESOL Quarterly, 34*(2), 213–238.
- Derakhshan, A., & Kaivanpanah, S. (2011). The impact of text-messaging on EFL freshmen's vocabulary learning. *The EUROCALL Review, 19*, 47–56.
- Dixson, R. J. (2004). *Essential idioms in English: Phrasal verbs and collocations*. London, UK: Longman.
- Douglas, S. R. (2010). *Non-native English speaking students at university: Lexical richness and academic success* (Unpublished dissertation). University of Calgary, Calgary, AB.
- Erasmus, G. (2002). Why can't we talk? *The Globe and Mail*, F6–7.
- Fowler, W. S., & Coe, N. (1976). *Nelson English language test (NELT)*. London, UK: Nelson.
- Gaetz, S., Tarasuk, V., Dachner, N., & Kirkpatrick, S. (2006). "Managing" homeless youth in Toronto: Managing food access and nutritional well-being. *Canadian Review of Social Policy, 58*, 43–61.
- Hayati, A., Jalilifar, A., & Mashhadi, A. (2013). Using short message service (SMS) to teach English idioms to EFL students. *British Journal of Educational Technology, 44*(1), 66–81.
- Hulstijn, J. H. (2001). Intentional and incidental second-language vocabulary learning: A reappraisal of elaboration, rehearsal, and automaticity. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 258–286). Cambridge, UK: Cambridge University Press.
- Kasesniemi, E. (2003). *Mobile messages: Young people and a new communication culture*. Tampere, Finland: Tampere University Press.
- Kennedy, C., & Levy, M. (2008). L'italiano al telefonino: Using SMS to support beginners' language learning. *ReCALL, 20*(3), 315–330.
- Kline, R. B. (2004). *Beyond significance testing: Reforming data analysis methods in behavioral research*. Washington, DC: American Psychological Association.

- Krashen, S. (1989). We acquire vocabulary and spelling by reading: Additional evidence for the input hypothesis. *The Modern Language Journal*, 73, 440–464.
- Krashen, S. (2003). *Explorations in language acquisition and use: The Taipei lectures*. Portsmouth, NH: Heinemann.
- Laufer, B., & Nation, I. S. P. (1999). A vocabulary-size test of controlled productive ability. *Language Testing*, 16(1), 33–51.
- Laufer, B., & Ravenhorst-Kalovski, G. C. (2010). Lexical threshold revisited: Lexical text coverage, learners' vocabulary size, and reading comprehension. *Reading in a Foreign Language*, 22(1), 15–30.
- Lawrence, J. F., Crosson, A. C., Paré-Blagoev, E. J., & Snow, C. E. (2015). Word generation randomized trial discussion mediates the impact of program treatment on academic word learning. *American Educational Research Journal*, 52, 750–786.
- Lawson, M. (2003). *Crow lake*. Toronto, ON: Random House of Canada.
- Lenhart, A. (2010). Cell phones and American adults. Washington, DC: Pew Research Center. Retrieved from <http://www.pewinternet.org/2010/09/02/cell-phones-and-american-adults/>
- Levy, M., & Kennedy, C. (2005). Learning Italian via mobile SMS. In A. Kukulska-Hulme & J. Traxler (Eds.), *Mobile learning: A handbook for educators and trainers* (pp. 76–83). New York, NY: Routledge.
- Li, J., Cummins, J., & Deng, Q. Z. (2018). The effectiveness of texting to enhance academic vocabulary learning: English language learners' perspective. *Computer Assisted Language Learning*, 30(8), 816–843. <https://doi.org/10.1080/09588221.2017.1366923>
- Lin, C. C., & Yu, Y. C. (2016). Effects of presentation modes on mobile-assisted vocabulary learning and cognitive load. *Interactive Learning Environments*, 25(4), 528–542.
- Ling, R. (2010). Texting as a life phrase medium. *Journal of Computer-Mediated Communication*, 15, 277–292. <https://doi.org/10.1111/j.1083-6101.2010.01520.x>
- Lu, M. (2008). Effectiveness of vocabulary learning via mobile phone. *Journal of Computer Assisted Learning*, 24, 515–525. <https://doi.org/10.1111/j.1365-2729.2008.00289.x>
- Motallebzadeh, K., & Ganjali, R. (2011). SMS: Tool for L2 vocabulary retention and reading comprehension ability. *Journal of Language Teaching and Research*, 2, 1111–1115.
- Motallebzadeh, K., Beh-Afarin, R., & Rad, S. D. (2011). The effect of short message service on the retention of collocations among Iranian lower intermediate EFL learners. *Theory and Practice in Language Studies*, 1, 1514–1520.
- Motiwalla, L. F. (2007). Mobile learning: A framework and evaluation. *Computers & Education*, 49, 581–596.
- Nagy, W., Herman, P., & Anderson, R. (1985). Learning words from context. *Reading Research Quarterly*, 20(2), 233–253.
- Nation, I. S. P. (2001). *Learning vocabulary in another language*. Cambridge, UK: Cambridge University Press.
- Nation, I. S. P., & Newton, J. (1996). Teaching vocabulary. In J. Coady & T. Huckin (Eds.), *Second language vocabulary acquisition* (pp. 238–254). Cambridge, UK: Cambridge University Press.
- Porter, D. (2008). *Check your vocabulary for academic English: All you need to pass your exams*. Oxford, UK: Macmillan Education.

- Read, J. (1998). Validating a test to measure depth of vocabulary knowledge. In A. Kunnan (Ed.), *Validation in language assessment* (pp. 41–60). Mahwah, NJ: Lawrence Erlbaum.
- Saran, M., Seferoğlu, G., & Çağiltay, K. (2012). Mobile language learning: Contribution of multimedia messages via mobile phones in consolidating vocabulary. *The Asia-Pacific Educational Researcher*, 21(1), 181–190.
- Schmitt, N. (2008). Instructed second language vocabulary learning. *Language Teaching Research*, 12, 329–363.
- Schmitt, N., Schmitt, D., & Clapham, C. (2001). Developing and exploring the behavior of two new versions of the vocabulary levels test. *Language Testing*, 18, 55–88.
- Seabrook, R., Brown, G. D. A., & Solity, J. E. (2005). Distributed and massed practice: From laboratory to classroom. *Applied Cognitive Psychology*, 19, 107–122.
- Smith, A., & Page, D. (2015). *U.S. smartphone use in 2015*. Washington, DC: Pew Research Center. Retrieved from <http://www.pewinternet.org/2015/04/01/us-smartphone-use-in-2015/>
- Snow, C. E., & Uccelli, P. (2009). The challenge of academic language. In D. R. Olson & N. Torrance (Eds.), *The Cambridge handbook of literacy* (pp. 112–133). Cambridge, UK: Cambridge University Press.
- Song, Y. (2008). SMS enhanced vocabulary learning for mobile audiences. *International Journal of Mobile Learning and Organisation*, 2(1), 81–98.
- Spagnolli, A., & Gamberini, L. (2007). Interacting via SMS: Practices of social closeness and reciprocation. *British Journal of Social Psychology*, 46(2), 343–364.
- Stahl, S. A., & Nagy, W. E. (2006). *Teaching word meanings*. Mahwah, NJ: Lawrence Erlbaum.
- Steckley, J. (1997). Aboriginal people. In P. U. Angelini (Ed.), *Our society: Human diversity in Canada* (pp. 131–158). Toronto, ON: Nelson.
- Steeves, V. (2014). Young Canadian in a wired world, Phase III: Life online. Ottawa, ON: MediaSmarts. Retrieved from [http://mediasmarts.ca/sites/mediasmarts/files/pdfs/publication-report/full/YCWWIII\\_Life\\_Online\\_FullReport.pdf](http://mediasmarts.ca/sites/mediasmarts/files/pdfs/publication-report/full/YCWWIII_Life_Online_FullReport.pdf)
- Stockwell, G. (2010). Using mobile phones for vocabulary activities: Examining the effect of the platform. *Language Learning & Technology*, 14(2), 95–110.
- Suwantarathip, O., & Orawiatnakul, W. (2015). Using mobile-assisted exercises to support students' vocabulary skill development. *TOJET*, 14(1), 163–171.
- Swanborn, M., & De Glopper, K. (1999). Incidental word learning while reading: A meta-analysis. *Review of Educational Research*, 69(3), 261–285.
- Tabatabaei, O., & Goojani, A. H. (2012). The impact of text-messaging on vocabulary learning of Iranian EFL learners. *Cross-Cultural Communication*, 8, 47–55.
- Thornton, P., & Houser, C. (2005). Using mobile phones in English education in Japan. *Journal of Computer Assisted Learning*, 21, 217–228.
- Watson, D. (1991). *Practicing idioms*. Hong Kong: Thomas Nelson and Sons.
- Wilkinson, L. (1999). Statistical methods in psychology journals: Guidelines and explanations. *American Psychologist*, 54(8), 594–604.
- Zhang, H., Song, W., & Burstson, J. (2011). Reexamining the effectiveness of vocabulary learning via mobile phones. *TOJET*, 10(3), 203–214.

## Appendix. Experimental and Quasi-Experimental Studies on English Vocabulary Interventions Using Texting

| Study                                      | Participants <sup>1</sup>                                             | Proficiency        | Target Words                                                             | Treatment Duration | Intervention Model and Treatment <sup>2</sup>                                                                                                                                                                                                                                                                                                                               | Control                                                                        | Baseline Comparability Test                                               | Measures and Item Reliability Index <sup>3</sup>                                                                      | Results                                                                                                                                                |
|--------------------------------------------|-----------------------------------------------------------------------|--------------------|--------------------------------------------------------------------------|--------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------|---------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------|
| Alemi et al. (2012)                        | 45 freshmen from Iran<br>(treatment $n = 28$ )                        | Upper-intermediate | 320 AWL words                                                            | 16 weeks           | <ul style="list-style-type: none"> <li>• Push model, receiving two SMS messages per week</li> <li>• Each message included 10 target words, definitions in Persian and English, and example sentences.</li> </ul>                                                                                                                                                            | Independently looking up 10 words using a dictionary twice per week            | N/A                                                                       | Self-developed post-tests and delayed post-tests (40 AWL target words) with ICRT, Cronbach's alpha = .89 (pilot test) | Significantly greater vocabulary gains only on delayed post-test, supporting SMS intervention                                                          |
| Basal, Yilmaz, Tanriverdi, and Sari (2016) | 50 freshmen from Turkey<br>(treatment $n = 25$ )<br>Convenient sample | Upper-intermediate | 40 idioms from MCASE <sup>4</sup>                                        | 4 weeks            | <ul style="list-style-type: none"> <li>• Push model, receiving 10 MMS messages per week</li> <li>• Each included an idiom, definition, visual support, and at least three sample sentences.</li> </ul>                                                                                                                                                                      | Handouts including idiom meanings, usage examples, fill-in-the-blank exercises | Self-developed pre-test on 40 target idioms                               | Self-developed pre-test and post-test (40 target idioms)                                                              | Significantly greater vocabulary gains, supporting MMS intervention                                                                                    |
| Cavus and Ibrahim (2009)                   | 45 freshmen from Cyprus                                               | N/A                | N/A                                                                      | 9 days             | <ul style="list-style-type: none"> <li>• Push model, receiving 16 SMS messages per day, one every half an hour, with three clusters of 16 messages repeated three days per week</li> <li>• Each included words and definitions in Turkish.</li> </ul>                                                                                                                       | N/A                                                                            | N/A                                                                       | Self-developed pre-tests and post-tests (number of target words unknown)                                              | Significantly greater vocabulary gains after SMS intervention                                                                                          |
| Chen et al. (2008)                         | 156 university students from Taiwan                                   | N/A                | 24 words from the most common 2284-word list (Bauman, 1995) <sup>5</sup> | 50 minutes         | <ul style="list-style-type: none"> <li>• Push model with four treatments</li> <li>• Student received messages including (a) words with their spelling, phonetics, and Chinese translation via SMS; (b) Treatment A plus a sample sentence via SMS; (c) Treatment A plus a picture denoting word meaning via MMS; or (d) Treatments A, B, and C combined via MMS.</li> </ul> | See treatments                                                                 | Self-developed screen test on 50 target words (Bauman, 1995) <sup>5</sup> | Self-developed post-tests (24 target words in two formats: recognition and recall)                                    | Significantly greater vocabulary gains for students with lower verbal and higher visual ability, supporting MMS intervention with pictorial annotation |

|                                   |                                                                                   |                                            |                                         |         |                                                                                                                                                                                                                                                                                                                                                |                                                              |                                                                                          |                                                                                                                                                   |                                                                                                                                 |
|-----------------------------------|-----------------------------------------------------------------------------------|--------------------------------------------|-----------------------------------------|---------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------|------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------|
| Derakhshan and Kaivanpanah (2011) | 43 freshmen from Iran (treatment $n = 21$ )<br>Random assignment                  | N/A                                        | About 200 words from Porter (2000)      | 6 weeks | <ul style="list-style-type: none"> <li>• Sending instructor SMS messages including a sentence using the target word and then receiving feedback</li> <li>• Texting 3 partners to exchange one sentence</li> <li>• 15–20 words per class meeting, two meetings per week</li> </ul>                                                              | Equivalent learning activities done in person and on paper   | Three vocabulary screening tests <sup>5</sup> with ICRT reported on three original tests | • Self-developed post-test and delayed post-test (40 target words)                                                                                | No significant differences in vocabulary gains on both post-test and delayed post-test                                          |
| Hayati et al. (2013)              | 45 EFL learners from Iran<br>Ages 19–24<br>Convenient sample                      | Intermediate (based on MTELP) <sup>6</sup> | 80 idioms from Watson (2009)            | 20 days | <ul style="list-style-type: none"> <li>• Push model, receiving four SMS messages per day</li> <li>• Each included an idiom, definitions, and sample sentences in English.</li> </ul>                                                                                                                                                           | Idioms and short passages with exercises in printed pamphlet | One <i>SD</i> below and above mean on standardized MTELP                                 | Self-developed multiple-choice pre-test and post-test (50 target idioms)                                                                          | Significantly greater vocabulary gains, supporting SMS intervention                                                             |
| Lin and Yu (2016)                 | 32 8th graders from Taiwan                                                        | Elementary                                 | 36 words (9 for each treatment mode)    | 4 weeks | <ul style="list-style-type: none"> <li>• Push model, receiving two MMS messages per week with four treatments</li> <li>• Student received messages including (a) text only (i.e., word, part of speech, Chinese translation, and example sentence), (b) text plus picture, (c) text plus sound, or (d) text plus picture and sound.</li> </ul> | See treatments                                               | N/A                                                                                      | Self-developed post-test and delayed post-test (18 target words)                                                                                  | Significantly greater vocabulary gains only on delayed post-tests, supporting combined MMS mode                                 |
| Lu (2008)                         | 30 Grade 10 vocational students from Taiwan                                       | Intermediate                               | 28 target words (14 for each treatment) | 2 weeks | <ul style="list-style-type: none"> <li>• Push model, receiving two SMS messages per day</li> <li>• Each included two target words, Chinese translations, and part of speech.</li> </ul>                                                                                                                                                        | Paper materials on same content of 14 other words            | N/A                                                                                      | Self-developed pre-test, post-test, and delayed post-tests (28 target words)                                                                      | Significantly greater vocabulary gains only on post-test, supporting SMS intervention                                           |
| Motallebzadeh, and Ganjali (2011) | 40 adult female EFL learners from Iran (treatment $n = 21$ )<br>Random assignment | Lower-intermediate                         | 50 words                                | 5 weeks | <ul style="list-style-type: none"> <li>• Push model, receiving three SMS messages per week</li> <li>• Each included three to four words, definitions, and sample sentences.</li> </ul>                                                                                                                                                         | Board and paper                                              | Nelson English Language Test (NELT; Fowler & Coe, 1976); ICRT, Cronbach's alpha = .82    | Self-developed pre-test and post-test (10 target words, 20 reading) with ICRT, $KR-21 = .82$ (pilot vocabulary test) and .71 (pilot reading test) | Significantly greater gains on the total score, supporting SMS intervention; no report on separate vocabulary or reading scores |

|                                        |                                                                                 |                                 |                                                             |          |                                                                                                                                                                                                                                                                                                       |                                                                    |                                                                                                  |                                                                                                                                                                     |                                                                                                                                                             |
|----------------------------------------|---------------------------------------------------------------------------------|---------------------------------|-------------------------------------------------------------|----------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------|--------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Motallebzadeh et al. (2011)            | 40 EFL learners from Iran (treatment $n = 21$ )<br>Random assignment            | Lower-intermediate              | 70 collocations from Dixson (2004)                          | 5 weeks  | <ul style="list-style-type: none"> <li>• Push model, receiving two SMS messages per week</li> <li>• Each included seven collocations, descriptions, and sample sentences.</li> </ul>                                                                                                                  | Receiving paper materials with the same content two times per week | Self-developed pre-tests on 40 target words and the NELT with ICRT, Cronbach's alpha = .82       | Self-developed multiple-choice pre-test and post-test (40 target words) with ICRT, Cronbach's alpha = .54 (post-test)                                               | Significantly greater vocabulary gains, supporting SMS intervention                                                                                         |
| Saran et al. (2012)                    | 103 university students from Turkey                                             | Elementary and pre-intermediate | 80 words supporting regular classroom activities            | 4 weeks  | <ul style="list-style-type: none"> <li>• Push model, receiving MMS messages</li> <li>• Each included words, pronunciations, definitions, sample sentences, related visuals, and information on word formation.</li> </ul>                                                                             | Paper- and web-based instruction                                   | Self-developed pre-test on 80 target words with ICRT, Cronbach's alpha = .84-.85                 | Self-developed pre-test, post-test, and delayed post-tests (80 target words for each English proficiency level, respectively) with ICRT, Cronbach's alpha = .80-.85 | Significantly greater vocabulary gains on both post-test and delayed post-tests, supporting MMS intervention (however, only for the pre-intermediate group) |
| Song (2008)                            | 10 adult EFL learners from Hong Kong                                            | N/A                             | Words about nightlife, food, weather, and travel            | 16 days  | <ul style="list-style-type: none"> <li>• Push model, receiving four to six SMS messages per day</li> <li>• Each included words and short explanations in English and Chinese.</li> <li>• A website included longer explanations and other resources.</li> </ul>                                       | N/A                                                                | N/A                                                                                              | Self-developed multiple choice pre-test and post-tests                                                                                                              | Significant vocabulary gains after SMS intervention                                                                                                         |
| Suwantharithip and Orawiatnakul (2015) | 80 university students from Thailand (treatment $n = 40$ )<br>Random assignment | N/A                             | 100 target words in a textbook                              | 6 weeks  | <ul style="list-style-type: none"> <li>• Receiving SMS messages including fill-in-the-blank exercises and paragraph writing exercises after class</li> <li>• Students sent the teacher complete assignments via SMS during the week. They received the teachers' instant feedback via SMS.</li> </ul> | Paper-based equivalent exercises in class and teacher feedback     | Self-developed pre-tests on 50 target words with ICRT, $KR-20 = .79$ (pilot test)                | Self-developed multiple choice pre-test and post-test (50 target words) with ICRT, $KR-20 = .79$ (pilot test)                                                       | Significantly greater vocabulary gains, supporting SMS intervention                                                                                         |
| Tabatabaei and Goojani (2012)          | 60 male high school students from Iran                                          | N/A                             | About 65 words from Birjandi, Anabisarab, and Samimi (2007) | 2 months | <ul style="list-style-type: none"> <li>• Sending five to six SMS messages per 1.5-hour session</li> <li>• Each included a sentence using a target word.</li> <li>• The instructor replied with feedback.</li> </ul>                                                                                   | Writing sentences using target words on paper                      | Self-developed multiple-choice pre-test on 40 target words with ICRT, $KR-21 = .79$ (pilot test) | Self-developed multiple choice pre-test and post-tests (40 word items) with ICRT, $KR-21 = .79$ (pilot test)                                                        | Significantly greater vocabulary gains, supporting SMS intervention                                                                                         |

|                            |                                                         |     |                                          |         |                                                                                                                                                                                                                                                                   |                                          |                                             |                                                                              |                                                                                        |
|----------------------------|---------------------------------------------------------|-----|------------------------------------------|---------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------|---------------------------------------------|------------------------------------------------------------------------------|----------------------------------------------------------------------------------------|
| Thornton and Houser (2005) | 44 female 3rd- and 4th-year college students from Japan | N/A | 20 words                                 | 2 weeks | <ul style="list-style-type: none"> <li>• Push model, receiving three mini-lesson emails via mobile phone per day, five target words per week</li> <li>• Each introduced and reviewed words in multiple contexts and incorporated the words in stories.</li> </ul> | Paper- or web-based activities           | N/A                                         | N/A                                                                          | Significantly greater vocabulary gains, supporting treatment                           |
| Zhang et al. (2011)        | 78 sophomores from China<br>Convenient sample           | N/A | 130 words from the TOEFL vocabulary test | 3 weeks | <ul style="list-style-type: none"> <li>• Push model, receiving two SMS message per day</li> <li>• Each included five words, part of speech, phonetics, Chinese translation, and sample sentences.</li> </ul>                                                      | Self-paced learning with paper materials | Self-developed pre-test on 130 target words | Self-developed pre-test, post-test, and delayed post-test (130 target words) | Significantly greater vocabulary gains only on post-tests, supporting SMS intervention |

<sup>1</sup>If reported, either convenient sample or randomly assignment is indicated in the table; otherwise the study didn't report on participant selection and assignment procedure.

<sup>2</sup>Texting frequency is indicated in the table unless it was not reported in the study.

<sup>3</sup>If reported, internal consistency reliability test (ICRT) either from self-developed or standardized tests is indicated in the table; otherwise ICRT was not reported in the study.

<sup>4</sup>MCASE = Michigan Corpus of Academic Spoken English

<sup>5</sup>Associates Test (Read, 1998), Levels Test (Laufer & Nation, 1999), and Academic Vocabulary Levels Test (Schmitt et al., 2001).

<sup>6</sup>MTELP = The Michigan Test of English Language Proficiency.

## About the Authors

Jia Li is an Associate Professor at the Faculty of Education at the University of Ontario Institute of Technology. She currently holds a John A. Sproul Research Fellowship at the University of California, Berkeley. She was a Canada–U.S. Fulbright Scholar at the Harvard Graduate School of Education from 2011 to 2012. Her research focuses on data-driven innovative language and literacy instruction using emerging technologies for linguistically diverse students.

**E-mail:** [jia.li@uoit.ca](mailto:jia.li@uoit.ca)

Jim Cummins is a Professor Emeritus at the Ontario Institute for Studies in Education of the University of Toronto. He held a Canada Research Chair (Tier 1) and was a recipient of the International Reading Association's Albert J. Harris award. He also received an honorary doctorate in Humane Letters from the Bank Street College of Education in New York City. His research focuses on literacy and multiliteracies, technology and academic language learning, and English as a second language.

**E-mail:** [james.cummins@utoronto.ca](mailto:james.cummins@utoronto.ca)