



Measuring Research Participants' EFL Proficiency: What Practical Options do Researchers in Thailand Have?

Sarut Supasiraprapa

National Institute of Development Administration

Researchers in Thailand often face the challenge of objectively measuring the EFL proficiency of their research participants as most Thais do not have scores from an internationally accepted standardized test, such as TOEFL or IELTS. This paper first presents a discussion of the necessity of assessing the English proficiency of non-native English speaking participants in ESL/EFL research. Following this, based on observations of published EFL studies conducted within the Thai context, the author discusses the lack of uniformity in how adult Thais' EFL proficiency has been assessed and reported. The author subsequently proposes two tests, Brown's (1980) cloze test and LexTALE (Lemhöfer & Broersma, 2012), as practical tests which quantitative researchers in Thailand can administer to obtain a reasonable proxy for Thais' EFL proficiency. For each test, a test description is given, including how it has been used in previous ESL/EFL research. The benefits and caveats associated with the tests are also identified. While recognizing that *language proficiency* is a complex construct, this paper aims to present researchers in Thailand with practical EFL proficiency assessment tests which they can adopt in further enhancing their research rigor.

Keywords: Thailand, EFL learners, ESL/EFL research, proficiency assessment, cloze tests, LexTALE

The Importance of Objective Measurement of ESL/EFL Proficiency in Research

There are several reasons why reporting the English proficiency of non-native English speakers (NNSs) is important in quantitative research on second language acquisition (SLA)¹, including research on ESL/EFL learning and teaching.² First, as Brown and Grüter (2020) observed, one major reason is the researcher's need to divide participants into proficiency groups (e.g., high and low proficiency groups) in cross-sectional studies investigating ESL/EFL development over time due to the impracticalities of a longitudinal study. Second, measuring participants' English proficiency is necessary even when proficiency is not a variable researchers focus on because ignoring its effect can pose a threat to internal research validity. For example, in studies that divide ESL/EFL learners into groups with each group receiving different treatment (e.g., teaching activities or task conditions), a lack of control of participants' initial English proficiency can cast doubt on whether the treatment is the sole determinant of the research

¹ As Loewen (2020) points out, SLA covers the acquisition of a second language both in a naturalistic context and in a classroom context.

² In this paper, an English-as-a-second language (ESL) context refers to a context in which NNSs learn English in an English-speaking environment (e.g., the US), while an English-as-a-foreign-language (EFL) context is a context in which NNSs learn English in a non-English speaking setting (e.g., Thailand).



outcome (e.g., English learning or task performance). Thus, in this type of studies, researchers typically make efforts to ensure that participant groups have comparable initial English proficiency (e.g., Ko, 2012; Nguyen & Boers, 2019). Third, one major concern in quantitative research is external validity, or whether the results are applicable to the wider population that the sample represents. Therefore, researchers should report participants' ESL/EFL proficiency to help readers determine the generalizability of the findings and to allow for the comparison of results across studies (Mackey & Gass, 2022; Norris & Ortega, 2000).

How Do Researchers Measure ESL/EFL Proficiency?

The construct of *language proficiency* is not easy to define (Hulstijn, 2011), and English proficiency has been operationalized differently in ESL/EFL research. To date, a great deal of such research has been conducted with adult NNSs who are university students in ESL contexts, particularly in the US and UK. To gauge participants' ESL proficiency, one common practice in such research is to use readily available standardized test scores which participants have previously obtained. In particular, scores from the Test of English as a Foreign Language (TOEFL) or the International English Language Testing System (IELTS) are often used (e.g., Jiang & Nekrasova, 2007; Siyanova-Chanturia et al., 2011; Supasiraprapa, 2019), although these tests have been developed to measure NNSs' language readiness to study at a university where English is the medium of instruction (Chapelle, 2011; Kerstjens & Nery, 2000). Alternatively, researchers assess ESL/EFL proficiency independently with a test, typically a commercial standardized proficiency placement test or a section of such a test—including the Oxford Placement Test (e.g., Tsimpli & Dimitrakopoulou, 2007), the Michigan Test of English Language Proficiency (e.g., Zhao & MacWhinney, 2018), and the reading comprehension section of the practice version of TOEFL (e.g., Akamatsu, 2003)—or oral interviews rated according to standard assessment guidelines, such as the Common European Framework of Reference for Languages, or CEFR (e.g., Foster et al., 2014). Finally, some studies do not rely on any proficiency test scores, but use other criteria, such as school placement (e.g., Laufer & Girsai, 2008) or educational level (e.g., Wolter & Gyllstad, 2013). While it might be impractical to expect all ESL/EFL researchers to adopt the same English proficiency test, if various proficiency criteria are used, a comparison of research participants across studies is certainly not straightforward.

Current Practice in Reporting EFL Proficiency of Research Participants in Thailand

As in other EFL settings, recruiting research participants in Thailand who are adult NNSs with scores from international standardized tests such as TOEFL or IELTS often pose a practical challenge to researchers. Individuals taking these tests are typically those planning to pursue a degree in an English-speaking country; however, researchers may not be able to recruit these individuals as research participants. Moreover, at present, while many Thai universities offer international degree programs, in which English is the medium of instruction, and thus accept TOEFL or IELTS scores as part of admission requirements, these universities typically give applicants the option to take in-house English proficiency tests, which usually resemble the former paper-based version of TOEFL (e.g., Chulalongkorn Business School, Chulalongkorn University, n.d.). Because these tests cost much less than TOEFL or IELTS, the tests may be preferred by applicants, and thus the number of Thais taking TOEFL or IELTS remains limited.

Consequently, it may not be surprising that research studies conducted in Thailand have adopted various approaches to reporting adult Thais' EFL proficiency. The author's recent reviews of studies published in the journals *PASAA* and *LEARN Journal*, both focusing on English language teaching and acquisition and published by reputable Thai universities, serve to illustrate this point. While studies

conducted in the Thai context may be published in other journals (e.g., McDonough & Nekrasova-Becker, 2014), these two journals are selected as representative journals for two reasons. First, academic journals published in Thailand are typically managed by Thai universities, and these two journals are currently the highest ranked Thai journals in SCOPUS. Second, compared to other journals in SCOPUS, these two journals contain a high proportion of research in the Thai context. Between 2017 and 2021, the two journals published 41 studies conducted in Thailand using quantitative research methods or mixed methods which included inferential statistics.³ While these studies typically specified research participants' educational backgrounds, only 27% included an objective measure of participants' EFL proficiency. The measures also differed greatly, ranging from the scores from a university's in-house EFL proficiency test, the researcher's self-made test, to scores from an English exam which is part of the Ordinary National Educational Test (O-NET), Thailand's national set of exams for undergraduate college admissions. Only four out of the 41 studies reported scores from a test available in other countries as a proxy for participants' English proficiency, namely, TOEFL, the Oxford Placement Test, the Test of English for International Communication (TOEIC), and the Vocabulary Size Test (Nation & Beglar, 2007). Overall, therefore, comparing Thais across the 41 studies can be challenging, and comparing these participants against ESL/EFL learners in other countries is even more difficult.

There has been a recent attempt to benchmark scores from the Chulalongkorn University Test of English Proficiency (CU-TEP), a test administrated by a prestigious Thai university, against the CEFR (Wudthayagorn, 2018). However, not all Thai EFL learners are university students, and not all Thai university students take CU-TEP. Moreover, for researchers in Thailand, accessing commercial standardized tests used in previous studies may be impossible or costly, and asking Thais to perform an English task so that they can be assessed with scoring rubrics based on international frameworks such as the CEFR may be impractical due to the required rater training. Therefore, if researchers in Thailand would like to have an objective measure of participant's EFL proficiency, they still face a challenge as to what test is most apt for use. To help these researchers address this methodological challenge, in the following sections, Brown's (1980) cloze test and LexTALE (Lemhöfer & Broersma, 2012) are suggested as two possible options for these researchers.⁴

It should be pointed out that this paper does not aim to argue that scores from these two tests should replace scores from tests which have been demonstrated to be reliable and valid measures of English proficiency for a particular communicative purpose (e.g., TOEFL or IELTS scores as measures of academic English proficiency). Nor does this paper argue that these two tests should be used as high-stake tests in Thailand. Instead, while recognizing that *language proficiency* is a complex construct (Hulstijn, 2011), the paper proposes Brown's (1980) cloze test and LexTALE as practical research instruments which quantitative researchers in Thailand can adopt specifically to provide a reasonable proxy for their research participants' EFL proficiency.

The following sections describe the characteristics of the two tests, the benefits these tests offer to researchers, including the available validity evidence for the tests, how the tests have been used in previous research, and some caveats that researchers who plan to use these tests should be aware of.

³ Given that one major goal of quantitative research is generalizability of findings (Mackey & Gass, 2022), only empirical studies using inferential statistics are reported here.

⁴ Researchers conducting EFL vocabulary studies in Thailand may also consider the Vocabulary Levels Test (Schmitt et al., 2001), which has been used in many previous vocabulary studies to assess the number of English words ESL/EFL learners know receptively (e.g., Shiotsu & Weir, 2007; Webb et al., 2013).

Brown's (1980) Cloze Test

Test description

First used and validated in Brown's (1980) widely cited article, this written cloze test is probably the most well-known cloze test in the SLA literature. In this test, ESL/EFL learners are presented with an English passage entitled *Man and His Progress*, which contains three paragraphs. Every 7th word in the passage is deleted, and there are 50 embedded blanks in total. Test takers are instructed to first read the passage quickly to get the general meaning and then fill in each blank with exactly one word. The test is not provided in the article, but researchers can write an email to request it directly from the test developer. Based on recent experience of the author, two test versions are available—one in a Microsoft Word format, which can be printed out as a paper-based test, and the other in a fillable and savable pdf format. The test developer also provides an answer key based on two types of scoring: exact answer and acceptable answer. The former is based on whether test takers supply the deleted word replaced by each blank, while the latter includes several possible answers. The test does not come with a separate manual for researchers, but according to Brown (1980), the passage was created from a reading text for intermediate ESL learners, and each possible answer in the key for acceptable answer scoring was supplied by at least two out of the 77 native English speakers participating in the test development.

Key benefits to researchers

Using Brown's (1980) cloze test scores as a proxy for Thais' EFL proficiency offers several benefits to researchers in Thailand. First, while it is debatable whether cloze test scores truly reflect general ESL/EFL proficiency (e.g., Abraham & Chapelle, 1992), this cloze test has been shown to be concurrently valid with the English as a Second Language Placement Examination (ESLPE), a large-scale well-developed placement test administered by the University of California at Los Angeles to assess ESL reading, writing, listening, vocabulary, and grammar (see Davidson & Lynch, 2002 for details about the history and development of ESLPE). That is, Brown (1980) demonstrated a strong correlation between ESLPE scores and scores from the cloze test based on exact and acceptable answer scoring, as indicated by Pearson's correlation coefficients of .88 and .90, respectively. Second, because a new cloze test development requires knowledge, time, effort, and financial resources for pre-testing and test revisions (Hulstijn, 2010), by using this cloze test, researchers in Thailand can save the time and money required for new test development. Moreover, besides administering the test in a paper-based or a fillable pdf format, researchers can create a web-based test version for online data collection. For example, at present, some websites (e.g., www.jotform.com) allow researchers to construct a cloze test consisting of multiple blanks in a paragraph. In addition, comparison of scores from Brown's (1980) cloze test across studies is now more convenient due to Brown and Grüter's (2020) recent study, which provides standardized scores of the test derived from more than 1,000 ESL/EFL participants in previous studies adopting this test. Researchers using the test can refer to these standardized scores and describe the raw scores their participants obtain (e.g., 28-39 points based on exact answer scoring) in terms of percentile based on Brown and Grüter's (2020) large sample data (e.g., 95.7th - 99.9th percentile), derived from English learners in various contexts. Furthermore, Brown's (1980) cloze test offers scoring flexibility and can be used to measure a wide range of English proficiency levels. Based on a test item analysis, Brown and Grüter (2020) reported that, when research participants are a heterogenous group of ESL/EFL learners who are generally proficient in English, researchers could use exact answer scoring as it is better at distinguishing between more and less proficient participants and thus yields more reliable scores. In contrast, when participants are a heterogenous group of ESL/EFL learners with generally low English proficiency, researchers can use acceptable answer scoring because it is better at discriminating between more and less proficient students.

Use in previous research

Scores from Brown's (1980) cloze test have been reported in previous studies to indicate ESL/EFL proficiency or to ensure that participants are homogenous in terms of their proficiency (e.g., Chrabaszcz & Jiang, 2014; Connell et al., 2018; Dekydtspotter & Miller, 2013; Kim & Rah, 2019). Scores from this test have also been used to classify ESL/EFL learners in a single study into proficiency groups or are used in statistical analyses as an English proficiency measure (e.g., Kim et al., 2019; Malone, 2018; Tremblay, 2008). These studies have been published in highly ranked journals indexed in SCOPUS, including *Studies in Second Language Acquisition*, *Language Learning*, *Second Language Research*, and *Applied Psycholinguistics*.

Caveats

ESL/EFL learners may spend a great deal of time on Brown's (1980) cloze test if they are not advanced learners because the test contains as many as 50 blanks and because the deletion of 50 words may reduce their passage comprehension. Thus, in a research project, if ESL/EFL learners complete other tests or tasks prior to this cloze test, they may possibly become fatigued or inattentive by the time they do the cloze test. Because fatigue or inattention can affect learners' test performance, and hence the internal validity of a test (Mackey & Gass, 2022), in such a case, the researchers should consider giving participants a break before the cloze test. Moreover, as Brown and Grüter (2020) observed, although one reason for the popularity of Brown's (1980) cloze test is the high scoring reliability as reported in Brown's (1980) paper, such reliability is not inherent to the test, but is specific to the test scores in each study. Therefore, researchers who use the test should calculate and report the reliability of the test. Brown and Grüter (2020) further argued that if the test is used to classify learners into proficiency groups and the test reliability is high, researchers can be reasonably confident in the classification. On the other hand, if the test is used only to document ESL/EFL proficiency, even when test reliability is low due to small sample sizes and/or a narrow range of test scores, researchers can still compare participants' scores to the previously discussed standardized scores the two researchers published.

LexTALE

Test description

Developed by Lemhöfer and Broersma (2012), LexTALE, or Lexical Test for Advanced Learners of English, is a lexical decision test which can be completed online at <http://www.lextale.com>. Based on this website, the test is intended for cognitive researchers studying ESL/EFL participants who are relatively proficient in English (i.e., medium to advanced English learners) in an experimental setting. On average, the test can be completed in approximately 3.5 minutes and is particularly suitable for SLA psycholinguistics experiments. In the test, each test taker sees 60 letter strings in English, one at a time, on a computer screen, and uses a mouse to press a YES or NO button on the screen to judge whether each string is an English word. The 60 items consist of 40 words and 20 non-words. The test is untimed; each test taker completes the test at their own pace but is instructed not to use a dictionary during testing. The test development stemmed from Lemhöfer and Broersma's (2012) observation that (1) previously there was no consensus as to how ESL proficiency is assessed in psycholinguistics research, (2) many ESL researchers relied on proficiency self-ratings or second language experience questionnaires, which may not be valid or reliable proficiency measurement tools, and (3) proficiency tests used in previous psycholinguistics ESL studies varied and may be impractical, expensive, or inaccessible to researchers.

Key benefits to researchers

First, although vocabulary knowledge constitutes only a part of English language proficiency, Lemhöfer and Broersma (2012) offered evidence based on their large-scale study to argue for LexTALE's criterion-related validity.⁵ That is, in the study, LexTALE scores obtained by Dutch and Korean learners of English correlated significantly with their general English proficiency as measured by the Quick Placement Test (Quick Placement Test, 2001), a commercial English proficiency test which has been validated on a large number of ESL/EFL learners and has been used to classify students into different proficiency groups based on the CEFR. Researchers in Thailand who use LexTALE can therefore consult the guidelines in Table 1, created by the test developers, to compare LexTALE scores against the CEFR levels.

TABLE 1
Comparison between CEFR Levels and LexTALE Scores

CEFR Level	CEFR Description	LexTALE score
C1 and C2	Upper and lower advanced/ proficient user	80%-100%
B2	Upper intermediate	60%-80%
B1 and lower	Lower intermediate or lower	Below 59%

Note. This table is adapted from a table in Lemhöfer and Broersma's (2012) study.

The second major benefit of LexTALE is that it is freely available and quick and easy to administer. Moreover, each test taker automatically sees their score, calculated from the number of correct responses, instantly following test completion. Details about the test scoring is also available for researchers at <http://www.lextale.com/scoring.html>. As the test website indicates, this test can also be implemented in experimental software, such as Praat and Matlab, in a paper-pencil format, or as part of an experiment created on an online platform. Finally, given the increasing number of experimental studies adopting LexTALE in other settings, researchers in Thailand who use this test can compare their participants against ESL/EFL learners in other contexts in terms of English proficiency.

Use in previous research

LexTALE scores have been used as a proxy for research participants' general ESL/EFL proficiency as relates to the CEFR (e.g., Peltonen, 2018). The scores can also be used as a basis for participant recruitment; in some experimental ESL/EFL studies researchers recruited only individuals who obtained a specified minimum LexTALE score to ensure that participants were similarly proficient in English and/or they were able to complete a given English task (e.g., Elgort et al., 2018; Hwang et al., 2018; Öksüz et al., 2021). In addition, LexTALE scores have been demonstrated to predict ESL/EFL learners' behavior in experimental research, such as their English morphological awareness (Wu & Juffs, 2021) and their reaction time and error rates in an English word recognition task (e.g., Lemhöfer & Broersma, 2012). As in the case of Brown's (1980) cloze test, studies using LexTALE scores as a proxy for ESL/EFL proficiency have been published in various highly ranked journals in SCOPUS, such as *Studies in Second Language Acquisition*, *Language Learning*, *Second Language Research*, and *The Modern Language Journal*, particularly in recent years.

Caveats

In the previously discussed study by Lemhöfer and Broersma (2012), the test developers offered validity evidence for the test based only on test takers who were relatively proficient ESL/EFL learners,

⁵ The full report based on this study can be viewed online at https://www.lextale.com/pdf/Lemhofer_Broersma_2012.pdf

that is, Dutch university students with frequent exposure to English and Korean university students with a mean TOEIC score of 887 out of 990 ($SD = 40$). Lemhöfer and Broersma (2012) therefore concluded that LexTALE is a useful and valid test for “medium to highly proficient speakers of [English as a second language]” (p.326). Thus, compared to Brown’s (1980) cloze test, LexTALE is arguably appropriate for measuring a more restricted range of ESL/EFL proficiency.

Conclusion

As Tremblay (2011) aptly argued, in research which aims to explain second language knowledge or behavior, reporting or controlling for second language learners’ proficiency in the target language is a necessity, and not an option. Given the challenge in recruiting adult Thais with scores from internationally accepted standardized tests and the cost of commercial EFL proficiency tests, Brown’s (1980) cloze test and LexTALE are two practical and freely available tests which researchers in Thailand can administer to obtain a reasonable proxy for research participants’ EFL proficiency. As discussed, Brown’s (1980) cloze test can be used to assess the different levels of EFL proficiency. However, test takers, particularly those with low EFL proficiency, may require substantial time for test completion, and the test needs to be manually scored. In the case of LexTALE, the test is quick and easy to administer and scores from the test have been benchmarked against the CEFR. Each test taker also obtains a score automatically at the end of the test, but the test may be more appropriate for assessing relatively proficient Thai EFL learners than low proficiency learners. Researchers in Thailand should therefore consider these pros and the cons when deciding whether either of these tests is appropriate for their research purposes. Regardless of choice, with these tests, researchers in Thailand can operationalize Thais’ EFL proficiency in a more objective manner and thus further improve the methodological rigor of their studies. Moreover, the tests will not only accommodate comparisons between different studies conducted in Thailand but also allow for comparisons between research participants in Thailand and elsewhere, thereby helping readers determine the generalizability of research findings.

Acknowledgement

The author would like to thank Dr. James Dean Brown for sharing the cloze test and for providing the author with some useful information about the test. Moreover, the manuscript benefited from the constructive feedback from Dr. Cecilia Zhao Guanfang, the section editor, and two anonymous reviewers.

The Author

Sarut Supasiraprapa is an assistant professor at the Graduate School of Language and Communication at the National Institute of Development Administration in Bangkok, Thailand. His research interests include second language acquisition, second language psycholinguistics, vocabulary learning, and research methods. His previous work has appeared in *TESOL Quarterly*, *Bilingualism: Language and Cognition*, and *Applied Psycholinguistics*.

Graduate School of Language and Communication
National Institute of Development Administration
13F Sayam Boromrajkumari BLDG
118 Seri-thai Road, Klongchan, Bangkok
Bangkok 10240 Thailand
Email: sarut.sup@nida.ac.th

References

- Abraham, R. G., & Chapelle, C. A. (1992). The meaning of cloze test scores: An item difficulty perspective. *Modern Language Journal*, 76, 468-479. <https://doi.org/10.2307/330047>
- Akamatsu, N. (2003). The effects of first language orthographic features on second language reading in text. *Language Learning*, 53, 207-231. <https://doi.org/10.1111/1467-9922.00216>
- Brown, J. D. (1980). Relative merits of four methods for scoring cloze tests. *Modern Language Journal*, 64, 311-317. <https://doi.org/10.1111/j.1540-4781.1980.tb05198.x>
- Brown, J. D., & Grüter, T. (2020). The same cloze for all occasions? Using the Brown (1980) Cloze Test for measuring proficiency in SLA research. *International Review of Applied Linguistics in Language Teaching (IRAL)*. Advance online publication. <https://doi.org/10.1515/iral-2019-0026>
- Chapelle, C. A. (2011). The TOEFL validity argument. In C. A. Chapelle, M. K. Enright, & J. M. Jamieson (Eds.), *Building a validity argument for the Test of English as a Foreign Language* (pp. 333-366). Routledge.
- Chrabaszcz, A., & Jiang, N. (2014). The role of the native language in the use of the English nongeneric definite article by L2 learners: A cross-linguistic comparison. *Second Language Research*, 30, 351-379. <https://doi.org/10.1177/0267658313493432>
- Chulalongkorn Business School, Chulalongkorn University (n.d.). Admission requirements 2022. <https://bba.acc.chula.ac.th/index.php/2014-06-27-06-06-17/admission-announcement/admission-requirements>
- Connell, K., Hüls, S., Martínez-García, M. T., Qin, Z., Shin, S., Yan, H., & Tremblay, A. (2018). English learners' use of segmental and suprasegmental cues to stress in lexical access: An eye-tracking study. *Language Learning*, 68, 635-668. <https://doi.org/10.1111/lang.12288>
- Davidson, F., & Lynch, B. K. (2002). *Testcraft: A teachers guide to writing and using language test specifications*. Yale University Press.
- Dekydspotter, L., & Miller, A. K. (2013). Inhibitive and facilitative priming induced by traces in the processing of *wh*-dependencies in a second language. *Second Language Research*, 29, 345-372. <https://doi.org/10.1177/0267658312467030>
- Elgort, I., Brysbaert, M., Stevens, M., & Van Assche, E. (2018). Contextual word learning during reading in a second language: An eye-movement study. *Studies in Second Language Acquisition*, 40, 341-366. <https://doi.org/10.1017/S0272263117000109>
- Foster, P., Bolibaugh, C., & Kotula, A. (2014). Knowledge of nativelike selections in a L2: The influence of exposure, memory, age of onset, and motivation in foreign language and immersion settings. *Studies in Second Language Acquisition*, 36, 101-132. <https://doi.org/10.1017/S0272263113000624>
- Hulstijn, J. H. (2010). Measuring second language proficiency. In E. Blom & S. Unsworth (Eds.), *Experimental methods in language acquisition* (pp.185-199). John Benjamins.
- Hwang, H., Shin, J. A., & Hartsuiker, R. J. (2018). Late bilinguals share syntax unsparingly between L1 and L2: Evidence from crosslinguistically similar and different constructions. *Language Learning*, 68, 177-205. <https://doi.org/10.1111/lang.12272>
- Jiang, N. A., & Nekrasova, T. M. (2007). The processing of formulaic sequences by second language speakers. *Modern Language Journal*, 91, 433-445. <https://doi.org/10.1111/j.1540-4781.2007.00589.x>
- Kerstjens, M., & Nery, C. (2000). Predictive validity in the IELTS test: A study of the relationship between IELTS scores and students' subsequent academic performance. *IELTS Research Reports*, 3, 86-108. <https://doi.org/10.1111/j.1540-4781.2007.00589.x>
- Kim, Y., Jung, Y., & Skalicky, S. (2019). Linguistic alignment, learner characteristics, and the production of stranded prepositions in relative clauses: Comparing FTF and SCMC contexts. *Studies in Second Language Acquisition*, 41, 937-969. <https://doi.org/10.1017/S0272263119000093>
- Kim, H., & Rah, Y. (2019). Constructional processing in a second language: The role of constructional knowledge in verb-construction integration. *Language Learning*, 69, 1022-1056. <https://doi.org/10.1111/lang.12366>

- Ko, M. H. (2012). Glossing and second language vocabulary learning. *TESOL Quarterly*, 46, 56-79. <https://doi.org/10.1002/tesq.3>
- Laufer, B., & Girsai, N. (2008). Form-focused instruction in second language vocabulary learning: A case for contrastive analysis and translation. *Applied linguistics*, 29, 694-716. <https://doi.org/10.1093/applin/amn018>
- Lemhöfer, K., & Broersma, M. (2012). Introducing LexTALE: A quick and valid lexical test for advanced learners of English. *Behavior Research Methods*, 44, 325-343. <https://doi.org/10.3758/s13428-011-0146-0>
- Loewen, S. (2020). *Introduction to instructed second language acquisition* (2nd ed.). Routledge.
- Mackey, A., & Gass, S. M. (2022). *Second language research: Methodology and design*. Routledge.
- Malone, J. (2018). Incidental vocabulary learning in SLA: Effects of frequency, aural enhancement, and working memory. *Studies in Second Language Acquisition*, 40, 651-675. <https://doi.org/10.1017/S0272263117000341>
- McDonough, K., & Nekrasova-Becker, T. (2014). Comparing the effect of skewed and balanced input on English as a foreign language learners' comprehension of the double-object dative construction. *Applied Psycholinguistics*, 35, 419-442. <https://doi.org/10.1017/S0142716412000446>
- Nation, I. S. P., & Beglar, D. (2007). A vocabulary size test. *The Language Teacher*, 31, 9-13.
- Nguyen, C. D., & Boers, F. (2019). The effect of content retelling on vocabulary uptake from a TED talk. *TESOL Quarterly*, 53, 5-29. <https://doi.org/10.1002/tesq.441>
- Norris, J. M., & Ortega, L. (2000). Does type of instruction make a difference? Substantive findings from a meta-analytic review. *Language Learning*, 51, 157-213. <https://doi.org/10.1111/j.1467-1770.2001.tb00017.x>
- Öksüz, D., Brezina, V., & Rebuschat, P. (2021). Collocational processing in L1 and L2: The effects of word frequency, collocational frequency, and association. *Language Learning*, 71, 55-98. <https://doi.org/10.1111/lang.12427>
- Peltonen, P. (2018). Exploring connections between first and second language fluency: A mixed methods approach. *Modern Language Journal*, 102, 676-692. <https://doi.org/10.1111/modl.12516>
- Quick Placement Test. (2001). Oxford University Press.
- Schmitt, N., Schmitt, D., & Clapham, C. (2001). Developing and exploring the behaviour of two new versions of the Vocabulary Levels Test. *Language Testing*, 18, 55-88. <https://doi.org/10.1177/026553220101800103>
- Shiotsu, T., & Weir, C. J. (2007). The relative significance of syntactic knowledge and vocabulary breadth in the prediction of reading comprehension test performance. *Language Testing*, 24, 99-128. <https://doi.org/10.1177/0265532207071513>
- Siyanova-Chanturia, A., Conklin, K., & Schmitt, N. (2011). Adding more fuel to the fire: An eye-tracking study of idiom processing by native and non-native speakers. *Second Language Research*, 27, 251-272. <https://doi.org/10.1177/0267658310382068>
- Supasiraprapa, S. (2019). Frequency effects on first and second language compositional phrase comprehension and production. *Applied Psycholinguistics*, 40, 987-1017. <https://doi.org/10.1017/S0142716419000109>
- Tremblay, A. (2008). Is second language lexical access prosodically constrained? Processing of word stress by French Canadian second language learners of English. *Applied Psycholinguistics*, 29, 553-584. <https://doi.org/10.1017/S0142716408080247>
- Tremblay, A. (2011). Proficiency assessment standards in second language acquisition research: "Clozing" the gap. *Studies in Second Language Acquisition*, 33, 339-372. <https://doi.org/10.1017/S0272263111000015>
- Tsimpli, I. M., & Dimitrakopoulou, M. (2007). The interpretability hypothesis: Evidence from *wh*-interrogatives in second language acquisition. *Second Language Research*, 23, 215-242. <https://doi.org/10.1177/0267658307076546>

- Webb, S., Newton, J., & Chang, A. (2013). Incidental learning of collocation. *Language Learning*, 63, 91-120. <https://doi.org/10.1111/j.1467-9922.2012.00729.x>
- Wolter, B., & Gyllstad, H. (2013). Frequency of input and L2 collocational processing: A comparison of congruent and incongruent collocations. *Studies in Second Language Acquisition*, 35, 451-482. <https://doi.org/10.1017/S0272263113000107>
- Wu, Z., & Juffs, A. (2021). Effects of L1 morphological type on L2 morphological awareness. *Second Language Research*. <https://doi.org/10.1177/0267658321996417>
- Wudthayagorn, J. (2018). Mapping the CU-TEP to the Common European Framework of Reference (CEFR). *LEARN Journal: Language Education and Acquisition Research Network Journal*, 11, 163-180.
- Zhao, H., & MacWhinney, B. (2018). The instructed learning of form-function mappings in the English article system. *Modern Language Journal*, 102, 99-119. <https://doi.org/10.1111/modl.12449>

(Received October 29, 2021; Revised February 24, 2022; Accepted March 18, 2022)