

# Vocabulary Knowledge and Vocabulary Use in Second Language Writing

MARK D. JOHNSON  
East Carolina University

ANTHONY ACEVEDO  
LEONARDO MERCADO  
Instituto Cultural Peruano Norteamericano

Research has consistently shown diversity of vocabulary to be an important indicator of second language (L2) writing development as well as L2 writing performance. These studies underscore the importance of vocabulary to L2 writing. However, they provide little to indicate what kind of vocabulary learners of English may need to know in order to develop writing proficiency. This small-scale pilot study examined the relationships among vocabulary knowledge, vocabulary use, and L2 writing performance. The results suggest that accurate productive knowledge of high-frequency word families was associated with L2 writing performance. However, actual use of high-frequency word families was negatively associated with L2 writing performance. Based on the results, the authors present potential uses of lexical frequency information to help students develop (a) accurate productive knowledge of high-frequency word families and (b) a repertoire of low-frequency word families based on their communicative needs.

*doi: 10.1002/tesj.238*

**T**o say that vocabulary is important to writing seems obvious. However, to the practicing teacher of second language (L2) writing, vocabulary takes on added importance because L2 writers often struggle with limited vocabulary or with vocabulary that may have been only partially learned (Nation, 2001). Despite the importance of vocabulary to L2 writing performance, L2 writing

researchers have offered little to inform L2 writing instructors' practice aside from the consistent finding that a greater diversity of vocabulary is associated with stronger L2 writing performance (Engber, 1995; Grant & Ginther, 2000). However, research is slowly beginning to inform the selection of target vocabulary for instruction and the types of knowledge that need to be developed to improve L2 writers' performance.

Originally designed to examine the size and development of L2 writers' productive vocabularies, early research used lexical frequency profiles to compare the vocabulary in L2 writers' texts to word frequency lists, typically the first (1K) and second (2K) 1,000 most frequent word families according to the General Services List (GSL; West, 1953) in combination with either (a) the University Word List (UWL; Xue & Nation, 1984) or (b) the Academic Word List (AWL; Coxhead, 2000). Vocabulary that did not appear in any of these lists was simply described as "beyond" the 2K frequency band.

The handful of studies using lexical frequency profiles to analyze L2 writers' texts have found that the use of vocabulary that appears less frequently in the English language—usually measured as decreased use of words from the 1K and 2K GSL list and increased use of words from the AWL or UWL and from beyond the 2K GSL list—is associated with L2 writing development over time as well as L2 writing performance (Laufer, 1994; Laufer & Nation, 1995). More recent studies using frequency lists based on the Bank of English corpus (Coniam, 1999), British National Corpus (Johnson, Acevedo, & Mercado, 2013), or specialty frequency lists generated from the participants' own writing (Verspoor, Schmid, & Xu, 2012) appear to support the findings of earlier studies using the GSL and UWL or AWL. To summarize the findings of these studies concisely, the use of words that appear less frequently in the English language is associated with stronger L2 writing performance as well as L2 writing development over time. Missing from research on vocabulary and its role in L2 writing performance is an exploration of vocabulary knowledge and its relationship to vocabulary use. In order to better understand this relationship, the following small-scale pilot study was driven by the following research questions:

1. What are the relationships among receptive, productive, and aural vocabulary knowledge and L2 writing performance?
2. What are the relationships among vocabulary use and L2 writing performance?
3. How do stronger and weaker L2 writers differ in receptive/productive/aural vocabulary knowledge?
4. How do stronger and weaker L2 writers differ in their use of vocabulary?

## METHOD

### Setting and Participants

The study was conducted in five advanced-level classes at a large binational center located in Lima, Peru. The participants were 100 learners of English as a foreign language, all of whom were just under 20 years of age ( $M = 19.58$ ,  $SD = 3.53$ ). Ninety-one participants indicated that the language they used in their homes was Spanish. One participant indicated that he or she spoke a language other than Spanish in the home, and eight participants did not identify the language they spoke in the home.

### Materials

The following materials were used to examine the relationship(s) among vocabulary knowledge, vocabulary use, and L2 writing performance. With the exception of the aural vocabulary test, all are freely available at [www.lex tutor.ca](http://www.lex tutor.ca). However, because of limited availability of computers in the classrooms at the research site, the materials were adapted as print materials.

**Receptive vocabulary size test (Nation & Beglar, 2007).** Originally designed to estimate students' receptive vocabulary size, the multiple-choice vocabulary size test presents 10 items in each of 14 different frequency bands, the first 1,000 (1K) through the fourteenth 1,000 (14K) most frequent word families according to the British National Corpus. In each of the frequency bands, the student is presented with 10 multiple-choice items. Each item presents a key word in all capital letters, followed by a sample sentence with the key word in bold (see Figure 1). Each sample sentence is constructed such that the meaning of the key word cannot be inferred from context. The sample sentence is

- |  |   |
|--|---|
| <p>1. SEE<br/>They <b>saw</b> it.</p> <ul style="list-style-type: none"> <li>a. cut</li> <li>b. waited for</li> <li>c. looked at</li> <li>d. started</li> </ul>                | <p>2. TIME<br/>They have a lot of <b>time</b>.</p> <ul style="list-style-type: none"> <li>a. money</li> <li>b. food</li> <li>c. hours</li> <li>d. friends</li> </ul>    |
| <p>3. PERIOD<br/>It was a difficult <b>period</b>.</p> <ul style="list-style-type: none"> <li>a. question</li> <li>b. time</li> <li>c. thing to do</li> <li>d. book</li> </ul> | <p>4. FIGURE<br/>Is this the right <b>figure</b>?</p> <ul style="list-style-type: none"> <li>a. answer</li> <li>b. place</li> <li>c. time</li> <li>d. number</li> </ul> |

Figure 1. Four sample items from the first section (1K) of the vocabulary size test (Nation & Beglar, 2007)

followed by four alternatives, from which the student must choose the alternative that most closely matches the meaning of the key word. Based on the results of previous research among learners in the same instructional context (Johnson et al., 2013), the participants were tested only on 1K through 5K word families.

**Aural vocabulary test (Fountain & Nation, 2000).** Developed as a vocabulary-based placement test, the five-paragraph dictation test was designed to test participants' aural knowledge of progressively less frequent vocabulary. For example, the first paragraph tests participants' knowledge of 10 key words from the first 500 most frequent word families. The second paragraph tests participants' knowledge of 20 key words from the second 500 most frequent word families. The third and fourth paragraphs test participants' knowledge of 10 key words each from the second and third 1,000 most frequent word families, respectively. Finally, the last paragraph tests participants' knowledge of 20 key words from the fourth through the sixth 1,000 most frequent word families, respectively. The key words are presented in context as part of a text on food production and distribution. Students are to transcribe the dictation test as it is read to them.

**Productive levels test (Laufer & Nation, 1999).** The productive levels test was used to measure the participants' productive vocabulary knowledge. The test was designed to estimate

students' productive knowledge of vocabulary in four frequency bands: words from the 2K frequency band, words from the 2K–3K frequency bands, words from the 3K–5K frequency bands, and words from the 5K–10K frequency bands.

In the test, students are presented with 18 sentence completion items in each frequency band. As can be seen in Figure 2, each completion item presents a sentence with the target word missing. The first few letters of the missing word are provided to prompt the appropriate response. Students complete the missing word by writing in the missing letters.

**L2 writing performance.** In order to determine the participants' L2 writing performance, the participants wrote essays in response to one of two counterbalanced writing prompts from past TOEFL independent writing tasks (see Appendix A). The prompts were chosen to address similar topics in different rhetorical modes: argumentation (Prompt A) and comparison (Prompt B).

## Procedures

The three vocabulary tests and the writing task were administered at 1-week intervals by the participants' regular classroom teachers, following explicit written instructions from the researchers. The productive vocabulary test, the receptive vocabulary test, and the composition task were all administered under timed conditions; participants had 30 minutes to complete each. The aural vocabulary test was dictated to the participants by the teachers. Per the instructions of the test's designers (Fountain & Nation,

### Section A

1. La\_\_\_\_\_ of rain led to a shortage of water in the city.
2. The rich man died and left all of this we\_\_\_\_\_ to his son.
3. Pup\_\_\_\_\_ must hand in their papers by the end of the week.
4. This sweater is too tight. It needs to be stret\_\_\_\_\_.
5. If you blow up that balloon any more, it will bur\_\_\_\_\_.

Figure 2. A sample item from the first section (1K) of the receptive levels test (Laufer & Nation, 1999)

2000), the test passage was dictated one time—and one time only—to the students. Students were instructed to transcribe as much of the passage as they could.

### **Data Coding and Analysis**

We collected complete sets of vocabulary tests and essays from 62 of the 100 participants. The tests and essays were analyzed in the following manner.

**Receptive vocabulary knowledge.** In each of the five frequency bands, we calculated the total number of correct items from 10 possible. Missing answers were given a score of zero. Once a score was calculated in each of the frequency bands, a total receptive vocabulary score was calculated by adding the total number correct in all of the frequency bands. All scores were then rescaled as decimal scores by dividing the total number correct by the total number possible.

**Aural vocabulary knowledge.** We calculated scores following the procedure described by Fountain and Nation (2000). Per this procedure, we calculated the number of key words correctly transcribed in each paragraph. No scores were assigned for key words in the first paragraph, because Fountain and Nation designed the first paragraph as an introductory warm-up paragraph. We totaled the number of correctly transcribed key words in each of the subsequent paragraphs and assigned a score for each paragraph. Further, we calculated a total score by adding participants' scores in each paragraph. All scores were rescaled as decimal scores by dividing the total correct by the total possible.

**Productive vocabulary knowledge.** We calculated four component scores in each of the frequency bands for each of the participants: (1) total number of correct words, spelled correctly, in the correct form (see Figure 3, item 1); (2) total number of correct words, spelled correctly, in the incorrect form (see Figure 3, items 2 and 4); (3) total number of correct words, spelled incorrectly, in the correct form (see Figure 3, item 3); and (4) total number of correct words, spelled incorrectly, in the incorrect form. Missing and incorrect words (see Figure 3, item 5) were given a score of zero. A total score was then calculated in each of the frequency bands by adding the four component scores. We also calculated a

## Section A

1. Lack of rain led to a shortage of water in the city.
2. The rich man died and left all of this welthy to his son.
3. Pupe/s must hand in their papers by the end of the week.
4. This sweater is too tight. It needs to be stretch.
5. If you blow up that balloon any more, it will burn.

Figure 3. An example illustrating how productive test answers were coded

grand total by adding the total scores for each of the frequency bands. All scores were rescaled as decimal scores by dividing the total correct by the total possible.

**L2 writing performance.** Each essay was assessed for holistic quality by two experienced teachers of first language writing using the six-point (0–5) holistic rubric used by ETS to assess performance on the independent writing task of the TOEFL (see Appendix B). After a 90-minute norming session, each teacher worked independently to rate each essay, and a high rate of inter-rater reliability was achieved ( $r = .98$ ). Ratings were averaged to produce a single score for each essay.

**Vocabulary use.** Each essay was analyzed using the Range program (Heatley, Nation, & Coxhead, 2002) to determine the number of word types—as opposed to word tokens—from the 1K through the 5K frequency bands according to the British National Corpus. The number of word types was normed to 100 words so that essays of differing lengths could be more accurately compared to one another.

## Statistical Analysis

To determine the relationship between various types of vocabulary knowledge, vocabulary use, and L2 writing performance, correlations were calculated between holistic writing score and (a) each of the vocabulary test scores and (b) the normed frequency of 1K–5K word types in the participants' essays. Significant correlations ( $\alpha = 0.01$ ) were then chosen for a comparison of high-performing L2 writers and low-performing L2 writers.

As can be seen in Table 1, the participants scored somewhat low in the raters' assessment of their writing performance. Further, the participants appear to have been quite closely matched in their writing performance, with a majority of participants receiving holistic scores of 2 or 3. However, a multifacets Rasch analysis indicated two significantly different levels of writing performance. Based on this information, the decision was made to assign participants with holistic scores of 3 and higher to the "high" writing performance group ( $n = 30$ ) and to assign participants with holistic scores of 2.5 and lower to the "low" writing performance group ( $n = 32$ ). Such a division is supported by an analysis of descriptors used in the holistic rubric (see Appendix B).

The distinction between a holistic score of 2 and a holistic score of 3 lies largely in the rubric's description of weakness at a holistic score of 2. At a holistic score of 2, L2 writers' performance indicates overall weaknesses in areas of L2 writing such as vocabulary, sentence structure, organization, and cohesion. In contrast, at a holistic score of 3, L2 writers' performance indicates greater facility in these areas, though some weakness may be evident in the writers' texts.

Each of the vocabulary measures that was identified in the correlational analysis was then compared between the two L2 writing performance groups using an independent-samples  $t$ -test ( $\alpha = 0.05$ ). For each significant comparison, effect sizes were calculated as Cohen's  $d$  using means and standard deviations.

## RESULTS

Significant correlations ( $\alpha = 0.01$ ) are presented in Table 2. Not surprisingly, the grand total of productive vocabulary test scores correlated moderately with holistic writing quality scores ( $r = .38$ ,  $p = .003$ ). Surprisingly, perhaps, productive knowledge of 2K word families also correlated strongly with holistic writing scores ( $r = .31$ ,  $p = .01$ ). However, *accurate* productive knowledge (i.e., the

TABLE 1. Participants' Holistic Writing Scores

Avg. score	1.0	1.5	2.0	2.5	3.0	3.5	4.0	4.5	5.0
Total participants	5	0	26	1	29	0	1	0	0



TABLE 2. Significant Correlations ( $\alpha = 0.01$ ) With Average Holistic Writing Scores

	Total productive vocabulary test score	Total 1K productive test score	Accurate 1K productive test score	Total 2K productive test score	Normed use of 1K word families	Normed use of 2K word families
Average holistic writing score	.38 ( $p = .003$ )	.31 ( $p = .01$ )	.42 ( $p = .001$ )	.41 ( $p = .001$ )	-.48 ( $p < .001$ )	-.35 ( $p = .005$ )

correct word, in its correct form, spelled correctly) of 2K word families correlated even more strongly with L2 writing performance ( $r = .42, p = .001$ ). Further, total productive knowledge of 2K–3K word families correlated moderately with holistic writing quality scores ( $r = .41, p = .001$ ).

In contrast, the strongest correlations were negative correlations between holistic writing scores and the use of 1K word types ( $r = -.48, p < .001$ ) and 2K word types ( $r = -.35, p = .005$ ), suggesting that the use of words from among the most frequent words in the English language was associated with poorer L2 writing performance. Taken together, the findings of the correlational analysis suggest that stronger L2 writers had accurate productive knowledge of word families from the 2K and 2K–3K lists, though they apparently relied on words from the 1K and 2K lists less than participants who received lower holistic writing scores.

As can be seen in Table 3, the results of the independent-samples *t*-test indicated that the higher performing L2 writers examined in this study had significantly higher total productive vocabulary test scores. Furthermore, the higher performing L2 writers in this study had significantly higher scores in the accurate 2K and total 2K–3K subsections of the productive vocabulary test but used significantly fewer 1K and 2K word types in their writing. Effect sizes ranged from moderate ( $d = .56$ ) to strong ( $d = -.88$ ), suggesting not only a significant difference but also a meaningful difference in productive vocabulary knowledge and vocabulary use between the two groups.

TABLE 3. Independent-Samples *t*-Tests Result ( $\alpha = 0.05$ )

	Low performing <i>M (SD)</i>	High performing <i>M (SD)</i>	<i>t</i>	<i>df</i>	Cohen's <i>d</i>
Productive vocabulary test, total score	0.21 (0.06)	0.25 (0.08)	-2.15*	60	0.56
Productive vocabulary test, accurate 1K word families	0.34 (0.13)	0.43 (0.12)	-2.80**	60	0.69
Productive vocabulary test, total 2K word families	0.21 (0.08)	0.27 (0.11)	-2.39*	60	0.63
1K word types/100 words	42.54 (5.49)	38.26 (4.18)	3.44**	60	-0.88
2K word types/100 words	3.60 (1.13)	2.75 (0.87)	3.33**	60	-0.84

\* $p < .05$ ; \*\* $p < .01$ .

## DISCUSSION

The results of the current study point to a conclusion that many L2 writing researchers and teachers have intuitively known for some time: Development of productive vocabulary knowledge is important to the development of L2 writing performance. However, the results of the current study suggest that an emphasis on *accurate* productive knowledge of high-frequency vocabulary may be more strongly associated with L2 writing performance. The matter is complicated, however, because actual use of high-frequency vocabulary items was even more strongly associated with weaker L2 writing performance in this study. This finding appears to support earlier research that indicated the use of low-frequency word forms was associated with stronger L2 writing performance (Coniam, 1999; Johnson et al., 2013; Laufer, 1994; Laufer & Nation, 1995).

It would appear, then, that teachers of L2 writing must balance two opposing forces: (1) the need to develop L2 writers' accurate productive knowledge of high frequency word families and (2) the need to develop L2 writers' productive knowledge of low-frequency word families. Complicating matters further for L2 writing instructors are the recommendations of many in the area of L2 vocabulary development (Nation, 2001), which suggest a focus on receptive word knowledge rather than productive word knowledge. How are teachers of L2 writing to achieve these seemingly contrary goals?

The following uses of lexical frequency profiles are proposed for the L2 writing classroom in order to achieve two broad ends: to develop L2 writers' accurate productive knowledge of high-frequency lexical items and to extend L2 writers' lexical repertoires beyond the most frequent lexical items. Teachers should, of course, adapt these ideas to the specific needs of their students.

### **Development of Accurate Productive Knowledge of High-Frequency Items**

First, per the recommendation of Nation (2001) and others, class time should be devoted to explicit vocabulary instruction in the most frequent word families, selection of which may be informed through teachers' use of frequency information. Lexical frequency profiles can be conducted on course materials—whether published or authentic—using the Range (Heatley, Nation, & Coxhead, 2002) profiler, which is freely available via Paul Nation's website, or the more user-friendly vocabprofiler available via [www.lextutor.ca](http://www.lextutor.ca). Teachers can then target vocabulary instruction on the first 2,000 to 3,000 most frequent word families.

Commonly used learning techniques such as word cards can be easily manipulated in such a way that productive, rather than receptive, word knowledge is practiced. For example, students can be paired to play a guessing game in which one student must explain the target word to his or her partner, who must then guess the word. Such a game provides students with receptive and productive word knowledge of the target vocabulary and can be further adapted so that students must write their answers. Peer correction may then promote accuracy and attention to spelling.

Teachers may consult word frequency lists in order to determine the affixes needed to teach their students the accurate use of word families rather than individual words. For example, the word *answer* appears in the first 1,000 word families together with *answered*, *unanswered*, *answering*, *answers*, *answerable*, and *unanswerable*. The members of this word family may be presented in a rich, natural context to model appropriate use of each of the word forms. Teachers may then direct students' attention to

inflectional and derivational morphology to further broaden students' productive vocabulary of word families.

As part of the revision process, teachers and students of L2 writing can also use lexical frequency profiles together with dictionaries, thesauruses, and other reference materials to highlight high-frequency vocabulary for replacement with low-frequency lexical items.

### **Expanding Vocabulary Beyond the Most Frequent Lexical Items**

As many L2 writing teachers are well aware, most writing tasks are not performed under timed conditions, and multiple revisions are the norm in most contemporary writing classrooms. For L2 writing classrooms with ready computer and Internet access, students can create lexical frequency profiles of their own texts.

The profiler available at [www.lex tutor.ca](http://www.lex tutor.ca) provides a number of output options that students and teachers can use to better understand students' reliance on vocabulary from each of the frequency bands. For example, students receive color-coded versions of their essays to indicate the use of lexis from each of the frequency bands. Further, students receive lists of words from their essays organized by frequency.

Students and teachers may work together to consider carefully the intended audience of each student's text, then use this information together with frequency information and reference materials to determine which words are most suitable for replacement with a more advanced, lower frequency word. For example, a search for synonyms of the word *animal* uncovered the following four synonyms: *critter*, *beast*, *creature*, and *being*. Using the vocabprofile tool at [www.lex tutor.ca](http://www.lex tutor.ca), it was determined that *animal* and *being* are listed among the 1K word families. *Creature* is among the 2K word families. *Beast* is among the 4K word families. *Critter* is among the 14K word families. Though *critter* and *beast* are low-frequency lexical items, they may not be appropriate to some academic audiences, whereas the word *creature* may be considered appropriate, depending on the writer's intended impact on the audience. Teachers can work together with students to refine their notions of intended audience and select vocabulary appropriate to that audience.

## CONCLUSION

Despite a number of limitations (i.e., the nature of correlation, the homogeneity and size of the sample), the results of the current study suggest that *accurate* productive knowledge of the most frequent word families—rather than total productive vocabulary—is associated with stronger L2 writing performance among L2 writers who are closely matched in general L2 proficiency. In contrast, the use of high-frequency words may be negatively associated with L2 writing performance. Based on these results, it would appear that teachers of L2 writing must strike a balance between equal but opposite forces among their students: developing an accurate base of high-frequency word families *and* developing a repertoire of mid- to low-frequency word families.

To address these seemingly opposing goals, we presented a number of ideas for incorporating lexical frequency profiles in L2 writing instruction. Doing so may help L2 writing students develop accurate productive knowledge of high-frequency word families and help students attend to word choice and strategic use of vocabulary reference materials to develop an individualized lexicon of low-frequency items.

## ACKNOWLEDGMENTS

We wish to express our deepest gratitude to Drs. Sidney A. McPhee and Brad Bartell of Middle Tennessee State University for their generous support of this project as well as the teachers and academic staff of Instituto Cultural Peruano Norteamericano for their hard work in implementing this study. We wish to also express our gratitude to Dr. Margo DelliCarpini and the anonymous reviewers of *TESOL Journal* for their feedback and input.

## THE AUTHORS

[Mark D. Johnson](#) is an assistant professor of TESOL and applied linguistics at East Carolina University. His research interests are in cognition and vocabulary in second language writing.

[Anthony Acevedo](#) works at Instituto Cultural Peruano Norteamericano, in Lima, Peru, where he is in charge of the academic program. He has worked in the English language

teaching field as both a teacher and an administrator with a continued interest in reading and curriculum design.

**Leonardo Mercado**, originally from Queens, New York, is the former academic manager at the Instituto Cultural Peruano Norteamericano and has worked in the field of English as a second or foreign language as a teacher, teacher trainer, program administrator, and certified language proficiency tester for almost two decades.

## REFERENCES

- Coniam, D. (1999). An investigation into the use of word frequency lists in computing vocabulary profiles. *Hong Kong Journal of Applied Linguistics*, 4(1), 103–123.
- Coxhead, A. (2000). A new academic word list. *TESOL Quarterly*, 34, 213–238. doi:10.2307/3587951
- Engber, C. (1995). The relationship of lexical proficiency to the quality of ESL composition. *Journal of Second Language Writing*, 4(2), 139–155. doi:10.1016/1060-3743(95)90004-7
- Fountain, R. L., & Nation, I. S. P. (2000). A vocabulary-based graded dictation test. *RELC Journal*, 3(2), 29–44. doi:10.1177/003368820003100202
- Grant, L., & Ginther, A. (2000). Using computer-tagged linguistic features to describe L2 writing differences. *Journal of Second Language Writing*, 9, 123–145. doi:10.1016/S1060-3743(00)00019-9
- Heatley, A., Nation, I. S. P., & Coxhead, A. (2002). *Range [Computer software]*. Wellington, New Zealand: Victoria University of Wellington.
- Johnson, M. D., Acevedo, A., & Mercado, L. (2013). What vocabulary should we teach? Lexical frequency profiles and lexical diversity in second language writing. *Writing and Pedagogy*, 5(1), 83–103. doi:10.1558/wap.v4i5.1
- Laufer, B. (1994). The lexical profile of second language writing: Does it change over time? *RELC Journal*, 25(2), 21–33. doi:10.1177/003368829402500202
- Laufer, B., & Nation, I. S. P. (1995). Vocabulary size and use: Lexical richness in L2 written production. *Applied Linguistics*, 16, 307–322. doi:10.1093/applin/16.3.307

- Laufer, B., & Nation, I. S. P. (1999). A vocabulary size test of controlled productive ability. *Language Testing*, 16(1), 33–51. doi:10.1093/applin/16.3.307
- Nation, I. S. P. (2001). *Learning vocabulary in another language*. Cambridge, England: Cambridge University Press.
- Nation, I. S. P., & Beglar, D. (2007). A vocabulary size test. *The Language Teacher*, 31(7), 9–13.
- Verspoor, M., Schmid, M. S., & Xu, X. (2012). A dynamic usage based perspective on L2 writing. *Journal of Second Language Writing*, 21, 239–263. doi:10.1016/j.jslw.2012.03.007
- West, M. (1953). *A general service list of English words with semantic frequencies and a supplementary word-list for the writing of popular science and technology*. London, England: Longman.
- Xue, G., & Nation, I. S. P. (1984). A university word list. *Language Learning and Communication*, 3, 215–229.

## APPENDIX A

### PAST TOEFL INDEPENDENT WRITING TASK PROMPTS USED IN THE PILOT STUDY<sup>1</sup>

#### PROMPT A

It has been said, “Not everything that is learned is contained in books.” Compare and contrast knowledge gained from experience with knowledge gained from books. In your opinion, which source is more important? Why?

#### PROMPT B

Some people believe that the best way of learning about life is by listening to the advice of family and friends. Other people believe that the best way of learning about life is through personal experience. Compare the advantages of these two different ways

---

<sup>1</sup>Reprinted by permission of Educational Testing Service (ETS), the copyright owner. No endorsement of any kind by ETS should be inferred.

of learning about life. Which do you think is preferable? Use specific examples to support your preference.

## APPENDIX B

### HOLISTIC ASSESSMENT RUBRIC USED TO EVALUATE PARTICIPANTS' WRITING PERFORMANCE<sup>2</sup>

Score	Task Description
5	<p><b>An essay at this level largely accomplishes all of the following:</b></p> <ul style="list-style-type: none"> <li>effectively addresses the topic and task</li> <li>is well organized and well developed, using clearly appropriate explanations, exemplifications, and/or details</li> <li>displays unity, progression, and coherence</li> <li>displays consistent facility in the use of language, demonstrating syntactic variety, appropriate word choice, and idiomaticity, though it may have minor lexical or grammatical errors</li> </ul>
4	<p><b>An essay at this level largely accomplishes all of the following:</b></p> <ul style="list-style-type: none"> <li>addresses the topic and task well, though some points may not be fully elaborated</li> <li>is generally well organized and well developed, using appropriate and sufficient explanations, exemplifications, and/or details</li> <li>displays unity, progression, and coherence, though it may contain occasional redundancy, digression, or unclear connections</li> <li>displays facility in the use of language, demonstrating syntactic variety and range of vocabulary, though it will probably have occasional noticeable minor errors in structure, word form, or use of idiomatic language that do not interfere with meaning</li> </ul>
3	<p><b>An essay at this level is marked by one or more of the following:</b></p> <ul style="list-style-type: none"> <li>addresses the topic and task using somewhat developed explanations, exemplifications, and/or details</li> <li>displays unity, progression, and coherence, though connection of ideas may be occasionally obscured</li> <li>may demonstrate inconsistent facility in sentence formation and word choice that may result in lack of clarity and occasionally obscure meaning</li> <li>may display accurate but limited range of syntactic structures and vocabulary</li> </ul>
2	<p><b>An essay at this level may reveal one or more of the following weaknesses:</b></p> <ul style="list-style-type: none"> <li>limited development in response to the topic and task</li> <li>inadequate organization or connection of ideas</li> <li>inappropriate or insufficient exemplifications, explanations, or details to support or illustrate generalizations in response to the task</li> <li>an accumulation of errors in sentence structure and/or usage</li> </ul>
1	<p><b>An essay at this level is seriously flawed by one or more of the following weaknesses:</b></p> <ul style="list-style-type: none"> <li>serious disorganization or underdevelopment</li> <li>little or no detail, or irrelevant specifics, or questionable responsiveness to the task</li> <li>serious and frequent errors in sentence structure or usage</li> </ul>
0	<p><b>An essay at this level</b> merely copies words from the topic, rejects the topic, or is otherwise not connected to the topic, is written in a foreign language, consists of keystroke characters, or is blank.</p>

<sup>2</sup> Reprinted with the permission of Educational Testing Service (ETS; 1996) the copyright owner. No endorsement of this study by ETS should be inferred.