

i-lex v1 and v2: An Improved Method of Assessing L2 Learner Ability to See Connections between Words?

Ian Munby

Hokkai Gakuen University

doi: <http://dx.doi.org/10.7820/vli.v06.1.Munby>

Abstract

Knowing a word's associations is considered an aspect of word knowledge. It follows that L2 learner ability to see connections between words may improve with gains in vocabulary knowledge. Word association tests (WATs) may measure not only learner ability to see links between words, but they may also assess the degree of organization of L2 learner lexical knowledge which plays a role in the development of lexical competence. The aim of this study is to develop a new WAT wherein learners are presented with the three most common associates of a cue word. The task is to supply the missing cue word. Following this format, a test was developed using sets of three cue words chosen from the five most common associates to 50 target words (TWs) listed in the Edinburgh Associative Thesaurus, or EAT. Results of an initial study (i-lex v1) showed that, on average, a group of native speakers outperformed an experimental group of Japanese learners of English ranging in level from elementary to upper intermediate. Further, both in the initial study and a follow-up study (i-lex v2), significant and positive correlations were found among nonnative i-lex scores and a translation test. In i-lex v2, significant and positive correlations were also found among nonnative i-lex scores and the New Vocabulary Levels Test. These results indicate that the ability of these groups of participants to see links between highly frequent English words is related to their vocabulary knowledge.

1. Introduction

1.1 Knowledge of Word Associations

One complicating factor with research into L2 word associations is that some researchers have viewed word associations purely as an aspect of individual word knowledge. This line of thinking can be traced back to an influential paper by Richards (1976). He included knowing the “network of associations between that word and other words in language” (p. 81) as the sixth of eight word knowledge categories. It is interesting that his conclusion on this assumption focuses on the way words are stored in the mind “according to associative

bonds” (p. 87), or how words are learned and remembered, rather than whether or not knowing a word’s associations actually constitutes word knowledge. This is a vital consideration since associational knowledge differs from other forms of word knowledge in that it is not usually declarative knowledge, such as orthographic or morphological knowledge. It is for this reason that Meara (1996b) found this sixth assumption to be an exception to the set since it is not “driven exclusively by the concerns of descriptive linguistics, rather than by psycholinguistic concerns” (p. 3). Further, he believes associational knowledge has the potential to explain how word knowledge is acquired, or integrated into existing knowledge, while most descriptive knowledge does not.

While knowledge of word associations in L2 has since become established as an aspect of both productive and receptive word knowledge (Nation, 2001, p. 27), there appears to be a division of views among researchers regarding the degree of importance of associative links in the mind of the learner. On the one hand, associative knowledge is regarded as part of depth of individual word knowledge (e.g., Haastrup & Henriksen, 2000). On the other hand, while Meara and Wolter (2004) do view individual word knowledge as being important, they describe the breadth/depth model as “unfortunate,” preferring a “size/organization” model. The notion of organization is linked to a lexicon-focused or lexical network-oriented perspective on L2 lexical competence. This is based on the view that a learner’s L2 lexical competence is dependent not only on the number of words known but also on the state or quality of their interdependency in the mind of the learner. Underpinning this notion of lexical organization is the metaphor of the lexical store as a network, described by Aitchison (1994) as a “gigantic multi-dimensional cobweb” of words (p. 84). Wray (2002) also views lexical networks as central to the retention and production of language. These networks may play a role in determining lexical processing efficiency with both second language input (listening and reading) and output (speaking and writing). In other words, the dynamics of a learner’s lexical networks may tap into the core of an L2 learner’s overall proficiency and ability to make gains in the future.

This said, this view of L2 lexical competence has not gained much traction in the literature in the past two decades. For example, the section on word association in the famous 477-page, otherwise comprehensive book titled *Learning Vocabulary in another Language* (Nation, 2001) barely spans three pages, implying a limited role for word association studies. Similarly, at the end of a 7-page section on word associations in his book *Vocabulary and Language Teaching*, Schmitt (2000) concluded that research in this subfield has produced little that can inform the teaching of vocabulary.

1.2 Word Association Tests

Existing word association tests (WATs) in L2 vocabulary research seek to measure learner ability to make associations between words using a variety of formats. These can be broadly divided into productive WATs and receptive WATs. Given that associative knowledge cannot be classified as declarative knowledge, problems immediately surface in designing free, productive WATs.

A productive WAT typically involves inviting participants to supply a single response (in a discrete or continued WAT) or multiple responses (in a continuous WAT) to a set of stimuli with no restrictions placed on the type of response. In one strand of word association research, researchers such as Schmitt and Meara (1997), Schmitt (1998a), Schmitt (1999), Wolter (2002), Zareva (2005), and Higginbotham, Racine, and Munby (2015) have elicited responses from cue words and measured them against lists of associative norms generated from native speakers for scoring purposes.

While the results of these studies have generally indicated a link between performance on WATs and standard language tests or measures of overall language proficiency, there is some evidence to suggest that results depend on the cue words used to elicit responses and the norms list used to measure them for native-like stereotypy. For example, Kruse, Pankhurst, and Sharwood-Smith (1987) found no correlation between learner performance on a multiple response WAT and cloze test scores. They concluded there was no link between ability to produce word associations and proficiency. However, in a replication of this study using carefully selected cue words and new purpose-built norms lists in a WAT known as WAT20, Fitzpatrick and Munby (2014) found that learner WAT scores correlated significantly and positively with three different proficiency measures: a cloze test, a translation test based on Webb (2008), and the TOEIC (Test Of English for International Communication) test of listening and reading comprehension. These results indicate that with gains in proficiency, learner word associations become more native-like. Nevertheless, the issue raised by Schmitt (1998b) of what constitutes a native-like response is a concern that still challenges the validity of this approach to measuring associative competence. Indeed, Munby (2011) finds several instances of both native and nonnative participants providing the same nonnorms listed, and therefore nonscoring, responses to some cue words in WAT20.

With receptive WATs (e.g., The Word Associates Test, Read, 1993, 1998; V_Links, Meara & Wolter, 2004), different challenges to validity emerge. For example, with Read's "classic" 1998 test format, the task is to choose associated words to a TW from among two sets of four words each, one containing synonyms and the other containing collocates. As Read (2012) admitted, this format allows testees to guess word associations through elimination processes without knowing why they are associated. A validation study by Schmitt, Ng, and Garras (2011) provided evidence to support this weakness in the form of interviews with participants to examine test-taking strategies. These interviews sometimes revealed discrepancies between their successful answering of items and actual knowledge of the words being tested. A further issue highlighted by Read (2012) is that items are constructed through careful thought and dictionary study with the result that they bear no psycholinguistic reality, or are not based on data drawn from lists of associative norms. Finally, since many of the items in this test are low-frequency items, the test does not have the potential to test the ability to see connections between words with lower level learners. However, even if a test has weaknesses, it does not mean that it is not useful, and solutions can usually be found to most weaknesses in test design. Meanwhile, there is also a strong case for experimenting with new test formats. This certainly applies to WATs.

1.3 Aims and Research Questions

The aim of this study is therefore to pilot a new WAT inspired by Meara (1994), who mused upon possible uses of a Spanish word association norms list. He suggested presenting learners with the three most common associates of a cue word and asking them to supply the missing word. He added: “This is a task which native speakers find very easy, but one which turns out to be very difficult for non-native speakers” (p. 8). The author therefore decided to design a WAT with this format. The perceived strengths are firstly that the WAT is productive, with limited opportunity for guessing answers as with receptive formats, such as Read’s Word Associates Test. Second, there is no need to measure word association responses with norms lists for scoring purposes. Finally, it is appropriate for learners of all levels.

The instructions of this new WAT, i-lex, are as follows: What word is associated with the following sets of three words? Example: *drink, red, glass* > w _____ [4]. The first letter, “w,” is given for you and the word has four letters [4]. The answer is *wine*. In order to assess the validity of this WAT, the following four research questions were formulated to guide two versions of the test hereafter referred to as i-lex v1 and i-lex v2:

RQ1 Does i-lex v1 distinguish between native and nonnative speakers?

RQ2 Is there a significant, positive correlation between learner i-lex v1 and i-lex v2 scores and vocabulary test scores?

RQ3 Do i-lex v1 and i-lex v2 demonstrate internal reliability?

RQ4 Are nonnative speaker i-lex v2 results consistent between test and retest?

With reference to RQ4, note that a retest was not planned for i-lex v1 because it was not certain that the learner performance on the WAT would yield significant and positive correlations with scores on a standard vocabulary test.

2. Methodology

The following describes the methodology adopted in i-lex v1 and i-lex v2. The latter appears with the answer key in Appendix 1. Note that a brief report of the first version of i-lex (i-lex v1) appears in Munby (2013).

2.1 Participants

In i-lex v1, conducted in 2012–2013, the participants comprised a control group of 25 native speakers of English and 99 Japanese EFL (English as a Foreign Language) students who ranged in level from low elementary to upper intermediate. They were from four different universities, and were in their first, second, and third year. In the study on i-lex v2, conducted in 2016, there was no control group, but the experimental group of 164 Japanese participants was similar in range of level to the first study, and was drawn from first- and second-year EFL classes at three of the four universities where data from i-lex v1 were collected. These three

universities are classified nationally as high, middle, and low-ranking. All participants had studied English for at least 6 years.

2.2 Test Materials and Procedures

The three cue words (CWs) in each test item are chosen from the five most common associates on the Edinburgh Associative Thesaurus, or EAT (Kiss et al., 1973). The following criteria for item selection were established:

- (1) To minimize the possibility of words used in the test being unknown to lower level testees, all the TWs and CWs must be from the BNC (British National Corpus, 2007) 1K range, with words from this range not commonly known to low-level learners (e.g., *accept*, *account*, and *achieve*) also excluded according to the author's classroom experience.
- (2) To lend maximum transparency to the associations being tested, each set of CWs must include the most common associate on the EAT, listed first.
- (3) To ensure that dominant primary responses, including polar opposites such as *dark* > *light*, for example, do not excessively facilitate task success, the first CW must not account for more than 50% of the responses to the TW on the EAT.
- (4) To avoid priming, the TW must not also appear as a CW in another item.
- (5) To avoid conceptual repetition, the CWs in each set must not be part of the same word family, for example, *tell*, *teller* > *story*.
- (6) To avoid the need to accept alternative correct responses, all TWs which are verbs must not have a past tense form with the same number of letters, for example, *lose* > *lost*.
- (7) To ensure that the task is purely linguistic, TW–CW relationships must not require cultural knowledge such as *match*, *cricket* > *test*.
- (8) To reduce task difficulty, TWs must not be function words or noncontent words such as *through*.

For scoring purposes, one point was awarded for each correct TW supplied, whether it was spelled correctly or incorrectly, for example *club*, *up*, *together* > *joyn* for the TW *join*, but only if the specified number of letters was provided. A total of three trial versions of *i-lex v1* were conducted with two groups of learners ($n=22$ and $n=25$) in order to: (1) estimate appropriate time limits for a 50-item test, (2) identify and remove problematic items, such as items which no participants could answer, and (3) sort the items in order of difficulty from the easiest to the most difficult in order to limit incidences of lower level test-takers becoming stuck early on in the test and to facilitate consistency among items in a split-half reliability test.

Before beginning the test, participants were told that if they could not answer an item, they should leave it and move on to the next item. After completing *i-lex v1* (25 minutes), the nonnative group completed a translation test of controlled productivity adapted from Webb (2008) in 20 minutes (see Appendix 2). The answer key appears in Appendix 3. The task is to write English translations for a series of 160 single words of varying levels of word frequency written in L1(Japanese). In the original version of this translation test there were 180 items with a sample of 60 items from each of the following frequency bands in the BNC:

701st–1,900th, 1,901st–3,400th, and 3,401st–6,600th. In the original version of this translation test there were 180 items, but it was decided to shorten this to 120 items (40 in each band) because there are 36 loan words that are easily translated as transliterations of English loan words. Seventeen of these appear in the third band (3,401st–6,600th). Further, since a number of participants scored close to the maximum score in a previous study (Munby, 2011), it was felt that this test may not have the power to adequately differentiate the higher level students from their lower level peers. For this reason, it was decided to include an additional column of 40 Japanese words, with 10 each from the 6–7,000 K, 7,000–8,000 K, 8,000–9,000 K, and 9,000–10,000 K levels of the BNC. For scoring purposes, what Webb terms as “soft scoring” is applied and misspelled responses are accepted.

Following scoring of i-lex v1, a problem emerged, undetected during trialing, with one of the most challenging items: *boy, face, girl > baby* [4]. While 9 of the 99 nonnative participants supplied the target word successfully, 19 provided *body*. Among the 25 native participants, 8 responded successfully, but since 5 also provided *body*, it was decided to replace the item with a new set: *tree, fire, forest > wood* [4] in i-lex v2. Following i-lex v1, items were once again sorted for difficulty from easiest to most challenging according to test results and following a further trial with the new item with a group of nonnative participants ($n=25$).

The procedure for i-lex v2 was also modified slightly. The time allowed was reduced from 25 minutes to 20 minutes following an eleventh-hour conclusion that generally there was very limited pencil activity in the final 5 minutes of the test. Further, a second counter-proficiency measure, the New Vocabulary Levels Test (McLean & Kramer, 2015a), was administered following the translation test. This test was preferred to the Vocabulary Levels Test (Nation, 1990; Schmitt, Schmitt, & Clapham, 2001) for reasons described in McLean and Kramer (2015b), for example, there is no section which tests the first 1,000 word frequency band. Note that 30 minutes were allowed to complete Sections 1, 2, 3, 4, and 5 of this test. These sections test receptive knowledge of the first five frequency bands of the BNC/COCA corpus (Nation, 2012) with 20 multiple-choice items in each. Due to time constraints, it was decided to leave out Section 6 of the test which contains 30 items drawn from the Academic Word List (Coxhead, 2000). All three tests were completed in one 90-minute session during class time. A retest of i-lex v2 was conducted 2 weeks later to examine its reliability, and scores for participants who failed to attend both sessions were eliminated from the data set.

3. Results

In this section, with a view to answering the first two research questions concerning the validity of i-lex v1 and i-lex v2, the descriptive statistics for these two studies are presented in Table 1. In addition, native and nonnative performances are represented in a scatterplot for comparison in Figure 1. Correlational analysis is presented in Table 2, together with a scatterplot representation of i-lex v2 and translation test scores in Figure 2. To address the final two research questions concerning the reliability of i-lex v1 and i-lex v2, the results of the split-half reliability estimates for the learner scores are reported. Finally, test–retest reliability measures in learner i-lex v2 performance are presented.

Table 1. A Comparison of the Means and Standard Deviations of All Test Scores for All Participants

	Mean	SD	High	Low	MPS	R
i-lex 1 (ns, $n=25$)	41.76	4.28	49	33	50	0.96
i-lex 1 (nns, $n=98^*$)	25.00	7.23	42	6	50	0.99
Translation test	96.26	20.2	145	44	160	0.99
i-lex 2 (nns, $n=164$) T1	21.54	7.98	40	4	50	0.99
i-lex 2 T2	24.86	8.52	42	5	50	0.99
Translation test	89.83	22.76	152	31	160	0.99
New VLT	68.45	20.25	116	20	120	0.99

*Note that scores for one nonnative subject were removed because her i-lex score of 4 was an outlier.

High = Highest score achieved; Low = Lowest score achieved; MPS = Maximum Possible Score; R = Reliability coefficient; T1 = The first test; T2 = Retest after 2 weeks.

Table 2. Pearson Correlations among Scores for i-lex v1 and i-lex v2 and Proficiency Countermeasures

	Translation	New VLT
i-lex v1	0.729**	–
i-lex v2	0.804**	0.749**
New VLT	0.827**	–

Pearson 1-sided p -value: ** $p < 0.01$.

3.1 RQ1 Does i-lex v1 Distinguish Between Native and Nonnative speakers?

With reference to RQ1, the results in Table 1 indicate that, on average, native speakers outperform nonnatives on i-lex v1. A one-tailed unpaired t -test confirms that the difference between i-lex v1 scores for the two groups is significant at $t=4.199$ ($p < 0.0001$).

Figure 1 features a comparative representation of the distribution of i-lex v1 scores for the two subject groups: natives and nonnatives. One nonnative speaker scored above the native mean, but none of the native speakers scored below the nonnative mean. Note that, as Bachman (1990) points out, native speakers neither perform uniformly well, nor uniformly better than nonnatives in tests designed for L2 language learners.

3.2 RQ2 Is There a Significant, Positive Correlation between Learner i-lex v1 and i-lex v2 Scores and Vocabulary Test Scores?

Pearson correlations among all sets of scores are reported in Table 2. In both i-lex v1 and i-lex v2, these results indicate that the ability of these groups of participants to see links between highly frequent English words is related to their breadth of vocabulary knowledge. See Figure 2 for a scatter plot representation comparing learner performance on i-lex v2 and the translation test.

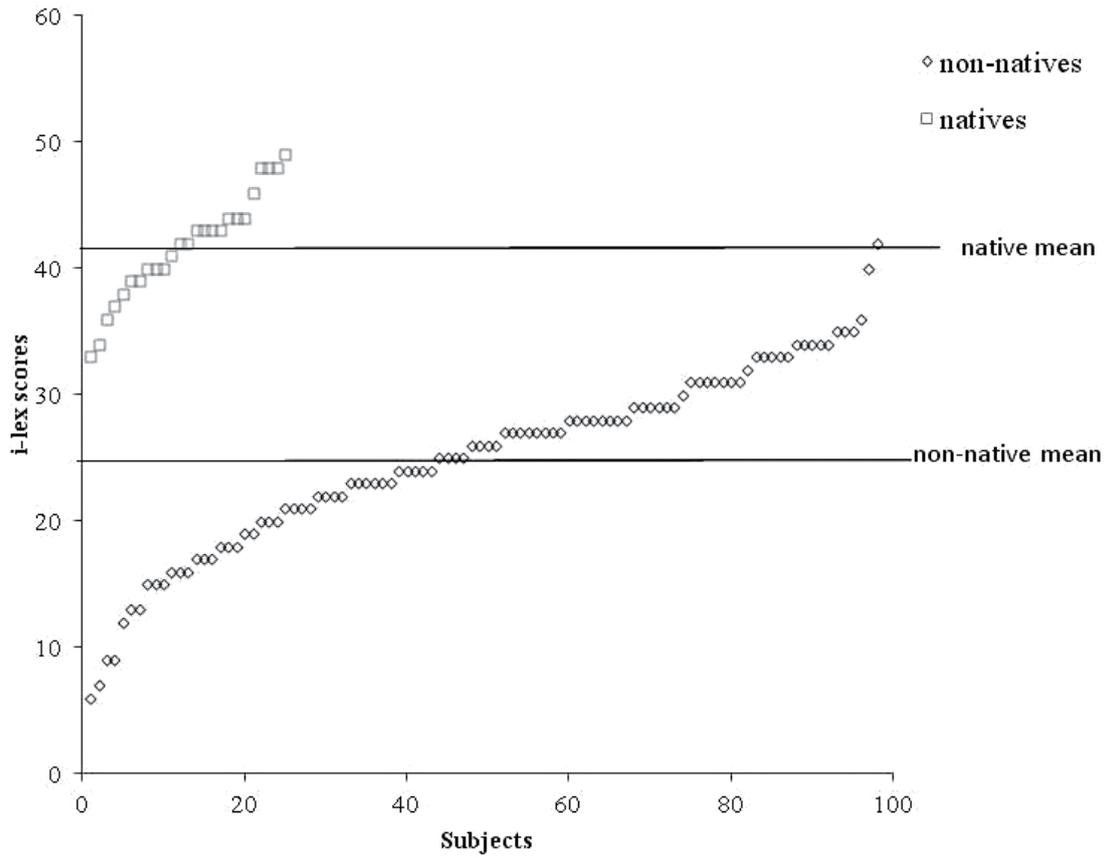


Figure 1. Distribution of nonnative and native speaker scores for i-lex v1.

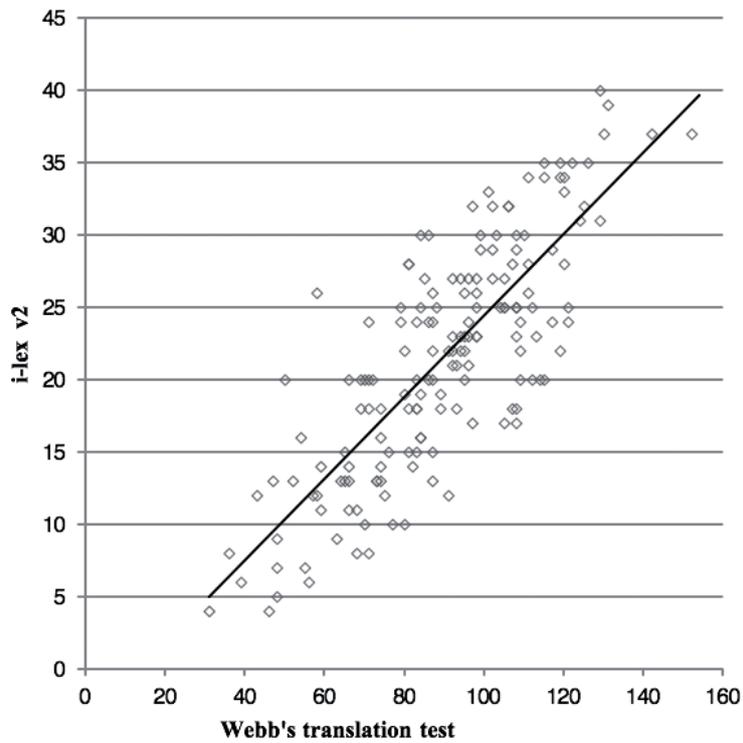


Figure 2. Scatter plot representation comparing learner performance on i-lex v2 and the translation test.

3.3 RQ3 Do *i-lex v1* and *i-lex v2* Demonstrate Internal Reliability?

In order to rule out the possibility of the participants producing correct responses in much greater quantity for some items than for others, affecting the balance of the set, a split-half reliability test was performed to check results for internal consistency. In this analysis, following Bachman (1990, p. 173), two parallel sets were constructed for both *i-lex v1* and *i-lex v2*. The “odd” set consisted of an analysis of the number of correctly answered items for the whole group for the odd-numbered sets (first, third, fifth, etc.). The even set comprised an analysis of scores for the even-numbered items (second, fourth, sixth, etc.). With reference to Table 3, the correlations for both *i-lex v1* and *i-lex v2* indicate that the items in each sub-test set of 25 items were assessing ability to see connections in a similar way. Further, results of a *t*-test do not indicate a significant difference between the means of the two sets in both versions of *i-lex*.

3.4 RQ4 Are Nonnative Speaker *i-lex v2* Results Consistent between Test and Retest?

Correlations between *i-lex v2* scores at T1 and the T2 retest after 2 weeks are $r = 0.871$ ($p < 0.01$) which suggest a satisfactory level of test–retest reliability of the WAT. As a further reliability check, a paired *t*-test between the pairs of means of the test at time 1 and time 2 was conducted, and a significant difference in *i-lex* scores was found ($t = 10.0967$, $p < 0.0001$). This indicates that these gains are consistent due to a practice effect that benefitted the majority of nonnative participants. See Figure 3 for a scatter plot representation of test and retest performance.

4. Discussion

This section presents an evaluation of the results in the light of the research questions and correlational analysis. It continues with discussion of the potential limitations of *i-lex* and suggests avenues for future research

Regarding RQ1, in *i-lex v1*, the native control group outperforms the non-native control group since, as Meara (1994) predicted, learners often find this kind of task more difficult than native speakers. With regard to RQ3, both *i-lex v1* and *i-lex v2* demonstrate internal consistency through a split-half reliability check. With regard to RQ4, with *i-lex v2*, a test–retest confirms that the WAT has yielded reliable results. Finally, concerning the key research question (RQ2), the results of these two studies indicate a relationship between learner ability to

Table 3. Means and Standard Deviations of Nonnative Participants Scores for Odd-Numbered and Even-Numbered Items in *i-lex v1* and *i-lex v2*, Pearson Correlations between the Two Sets, and One-tailed Paired *t*-test

	ODD Mean (SD)	EVEN Mean (SD)	Correlations	<i>t</i>
<i>i-lex v1</i> ($n=98$)	51.08 (28.3)	47.00 (25.0)	0.879**	1.5118 ns
<i>i-lex v2</i> ($i=164$)	70.76 (39.7)	70.84 (46.0)	0.733**	0.9949 ns

Pearson 1-sided *p*-value: Significant at ** $p < 0.01$; ns= not significant.

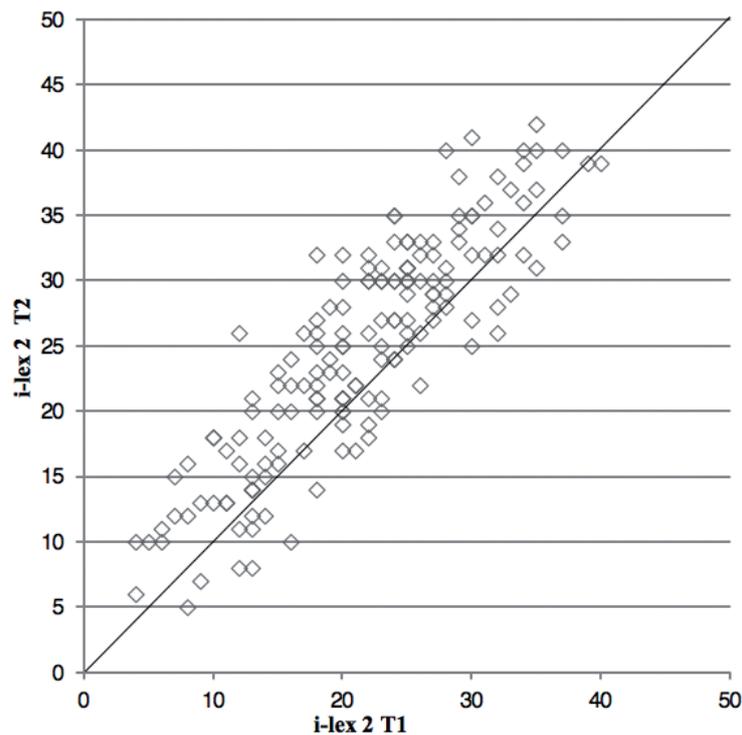


Figure 3. Comparison of i-lex v2 test–retest scores ($r = 0.871$ [$p < 0.01$]).

see connections between words and breadth of vocabulary knowledge. However, caution is required in claiming the relationship is strong on account of the high correlations between both versions of i-lex and the countermeasures. As Brown (2005) explains, there is a potential problem with correlational analysis since: “If a tester chooses to base a correlational analysis on a sample that is made up of fairly homogenous language proficiency levels ... the range of talent may have been restricted, and such a restriction will tend to make any resulting correlation coefficients much lower” (p. 161). Indeed, previous researchers (e.g., Wolter, 2002) have found low, but positive correlations when comparing results of their WATs with proficiency countermeasures in groups of nonnative participants of similar level. Unfortunately, their interpretation of their correlational analyses led to premature conclusions that the weak links were not promising for WA research. Conversely, if a broad range of proficiency levels is used, as in these two studies, the resulting correlation coefficients will tend to be much higher with the suggestion that the tests are essentially measuring the same kind of lexical knowledge.

Nevertheless, the following four potential limitations in i-lex are identifiable. These limitations concern the construct validity of the test defined by Daller, Milton, and Treffers-Daller (2007) as a question of “whether the test measures the skill or construct it is meant to” (p. 16). First, it is by no means certain that all nonnative participants had receptive knowledge of the meanings of all the CWs. It is also possible that they knew the L1 equivalent of the TW, but were unable to supply it due to lack of productive knowledge of the TW in L2. In other words, with some items, with lower level participants, i-lex may also be testing receptive and productive L2 vocabulary knowledge rather than pure ability to see associative links between words. A second potential problem is that it may be possible for

participants to answer items successfully without seeing links between all three CWs as intended. Taking the example of item 28 (*money, account, book > bank*), it is not inconceivable that testees could guess the answer from the first CW *money*. The third issue concerns test-taking strategies. Although test-takers are advised to move on if they are unable to answer an item in order to avoid getting stuck and losing precious time required to answer other items, there is no way of knowing if they took this advice. Alternatively put, low scores may indicate an unwillingness to abandon time-consuming struggles to solve a limited number of problematic items rather than inability to see associative links between words in general. The fourth issue concerns test format familiarity, a threat to validity mentioned by Daller et al. (2007, p. 17). Since the test format of *i-lex* is nonstandard, and the participants are unlikely to have completed any similar test, performances may be affected by unfamiliarity with the task of supplying a target word from its associates. Admittedly, the high test–retest correlations appear to rule this out to some extent. However, varied performance in *i-lex* may indicate varied need for practice with task type rather than varied ability in seeing connection between words. Regarding the effects of task familiarization, Akamatsu (2008) finds that a 7-week training program of once a week tests in L2 word recognition with a group of 49 first-year Japanese university students resulted in significant gains in both reaction times and response accuracy. It is therefore possible that participants could make similar gains with training in the *i-lex* format.

Turning to avenues of future research, there is no question that the issues described above concerning the validity of *i-lex* require addressing. One way to investigate the first issue of whether the CWs and TWs are known to the nonnative participants would be to give a translation test to a sample of lower level participants. In order to investigate the second issue of guessing TWs, it would be necessary to interview participants post-test to determine how they arrived at their successful answers following an approach adopted by Schmitt et al. (2011) in their validation of the Word Associates Test. Regarding the third potential limitation of test–taking strategies, it may be useful to give a parallel version of *i-lex* to learners who perform relatively poorly and adopt think aloud procedures to investigate whether or not poor test–taking strategies influence performance. Finally, replicating the approach taken by Akamatsu with *i-lex* could shed light on the extent to which format familiarity influences results.

It should not be forgotten that because the correlations are positive and significant, this apparent link between ability to see connections between high-frequency words and breadth of vocabulary knowledge requires explanation and further investigation. Of particular interest is whether or not participants who perform relatively well on *i-lex* are displaying above average lexical processing skills. Longitudinal studies to track learner development over time would be necessary to examine this possibility. In addition, studies involving a battery of tests combining *i-lex* with other WATs which investigate similar constructs such as WAT20 (Fitzpatrick & Munby, 2014) may shed light on whether similar patterns in learner performance emerge across tests.

In order to support the view that the efficiency of L2 learner lexical processing may be determined by L2 lexical networks, it is necessary to consider language learning from theories and findings in two other fields: foreign language

(FL) aptitude theory and neuroscience. To begin with FL aptitude theory, according to Wen, Biedroń, and Skehan (2017), the work of John Carroll in the 1950s and 1960s has proved enduring to this day. Carroll (1962) viewed specific talent for learning foreign or second languages as dependent on phonetic coding ability, grammatical sensitivity, inductive language learning ability, and associative memory. Clearly, it is the latter, or learner capacity to form associative links in memory, that should interest researchers in the field of L2 vocabulary learning. It may also be useful to draw parallels between associative networks as a metaphor for vocabulary learning in an L2 and evidence from studies in neurological science within the field of FL aptitude. For example, Wen et al. cite empirical studies by Li and Grant (2015) which found that “brain connectivity networks can indeed serve as a reliable predictor for learning both L2 novel words and (artificial) grammar/sequence rules in different learning contexts (natural, virtual etc.)” and that there was clear evidence of “more efficient and more flexible brain connectivity detected among more successful learners as opposed to their less successful counterparts” (p. 14).

Further, recent research conducted by neuroscientists Huth, Heer, Griffiths, Theunissen, and Gallant (2016) found that an “atlas” of the brain can be created by examining activity in the cerebral cortex. The technology employed (voxel-wise modelling of functional MRI) allows for visual imaging of how single words in stories light up spots in different areas of the brain of the listener. These combine to illuminate networks of spots representing how meanings of individual words may be constructed. The shape and form of these networks were found to differ from subject to subject. Judging from what we know about L2 word associations, these networks are likely to be less stable or more tenuous in learners than in native speakers of the language (Meara, 1983). In view of this recent research, lexical networks may not simply be confined to metaphorical representations of L2 language learning. In a similar vein, during the course of a presentation by Anthony, Schmitt, and Nation (2016), Anthony claimed that words are known by the company they keep. Although this comment was related to how studies in corpora can reveal relationships between words, the same claim can be applied to words in associative networks in the minds of language learners.

During the same presentation, in reference to word association research, Schmitt commented that most association responses were idiosyncratic and so we could not learn much about the learners. However, the proportion of idiosyncratic to nonidiosyncratic responses on norms lists of word associations depends on a number of factors such as the CWs selected, whether single or multiple responses are being elicited, and the number of participants. With many CWs, most responses are nonidiosyncratic. Further, the following two points need to be borne in mind. First, considerable gains in our understanding of WA have been made in the last three decades not only in this particular strand of WA research, as evidenced in the work cited in this paper, but also in a wealth of research left uncited here. See Meara (2009) for examples. Second, supported by newly available corpora, such as the BNC, and recent frequency profiling tools, L2 vocabulary research in the last three decades has been primarily motivated by a clearly identifiable need to investigate, for example, the readability of texts for L2 learners and what words need to be learned rather than by a desire to discover how they

are learned. If minor skirmishes regarding the influence on results of guessing in multiple-choice vocabulary size tests do not continue for too much longer, might it not be time to begin a new “post-size” era and to focus more on questions like “why do some learners learn more words than others?”

5. Conclusions

The present study investigated two versions of a new WAT called *i-lex*. Although conclusions remain tentative pending further validation, it was revealed that the ability of learners to make connections between highly common English words appears to be dependent on the number of words they know. The more words they know, the more connections they are able to identify. At present, it is not known whether this ability to make connections is a cause or a result of knowing the meanings of more words, or if it is a combination of both. Hopefully the next three decades will draw WA research out of the shadows, and it will receive more attention in the field of L2 vocabulary acquisition. Indeed, it is also hoped that new avenues shall be explored that focus more deeply on what it means to know a word and the role of lexical retrieval and memory in L2 lexical processing. At present, to its detriment, the field of L2 vocabulary studies remains remarkably insular. With new research techniques becoming available, this will surely change.

References

- Akamatsu, N. (2008). The effects of training on automatization of word recognition in English as a foreign language. *Applied Psycholinguistics*, *29*, 175–193. doi:10.1017/S0142716408080089
- Aitchison, J. (1994). *Words in the mind*. Oxford: Blackwell.
- Anthony, L., Schmitt, N., & Nation, I. (2016). Unanswered Questions in L2 Vocabulary Acquisition and Research Agendas for the Next 10 Years. Speeches presented at Vocab@tokyo, Tokyo, September 14th.
- Bachman, L.F. (1990). *Fundamental considerations in language testing*. Oxford: Oxford University Press.
- Brown, J.D. (2005). *Testing in language programs*. New York: McGraw-Hill.
- Carroll, J.B. (1962). The prediction of success in intensive foreign language training. In R. Glaser (Ed.), *Training research and education* (pp. 87–136). Pittsburgh, PA: University of Pittsburgh Press.
- Coxhead, A. (2000). A new academic word list. *TESOL Quarterly*, *34*, 213–238. doi: 10.2307/3587951
- Daller, H., Milton, J., & Treffers-Daller, J. (2007). *Modelling and assessing vocabulary knowledge*. Cambridge: Cambridge University Press.
- Fitzpatrick, T., & Munby, I. (2014) Word associations and the L2 lexicon. In J. Milton & T. Fitzpatrick. (Eds.), *Dimensions of vocabulary knowledge*. Basingstoke: Palgrave Macmillan, 92–105.

- Haastrup, K., & Henriksen, B. (2000). Vocabulary acquisition: Acquiring depth of knowledge through network building. *International Journal of Applied Linguistics*, 10(2), 221–239. doi:10.1111/j.1473-4192.2000.tb00149.x
- Higginbotham, G., Munby, I., & Racine, J.P. (2015). A Japanese word association database of English. *Vocabulary Learning and Instruction*, 4(2), 1–20.
- Huth, A.G., Heer, W.A., Griffiths, T.L., Theunissen, F.E., & Gallant, J.L. (2016). Natural speech reveals the semantic maps that tile human cerebral cortex. *Nature*, 532(7600), 453–458. doi:10.1038/nature17637
- Kiss, G.R., Armstrong, C., & Milroy, R. (1973). *An associative thesaurus of English*. EP Microfilms, Wakefield.
- Kruse, H., Pankhurst, J., & Sharwood-Smith, M. (1987). A multiple word association probe. *Studies in Second Language Acquisition*, 9(2), 141–154. doi:10.1017/S0272263100000449
- Li, P., & Grant, A. (2015). Second language learning success revealed by brain networks. *Bilingualism: Language and Cognition*, 19(4), doi:10.1017/s1366728915000280
- Meara, P.M. (1983). Word associations in a second language. *Nottingham Linguistics Circular*, 11, 28–38. doi:10.1177/026553229301000308
- Meara, P.M. (1994). Word associations in Spanish. *Vida Hispánica* 10, 12–22. Retrieved from www.lognostics.co.uk/vlibrary/meara1994.pdf
- Meara, P.M. (1996a). The dimensions of lexical competence. In G. Brown, K. Malmkjaer, & J. Williams (Eds.), *Performance and competence in second language acquisition* (pp. 35–53). Cambridge: Cambridge University Press.
- Meara, P.M. (1996b). *The vocabulary knowledge framework*. Retrieved from <http://www.swan.ac.uk/cals/calsres.vlibrary/pm96d.htm> revised 1999
- Meara, P.M. (2009). *Connected words: Word associations and second language vocabulary acquisition*. Amsterdam: John Benjamins.
- Meara, P.M., & Wolter, B. (2004). V_Links: Beyond vocabulary depth. *Angles on the English Speaking World*, 4, 85–97.
- McLean, S., & Kramer, B. (2015a). *The new vocabulary levels test*. Retrieved from <http://www.lex tutor.ca/>
- McLean, S., & Kramer, B. (2015b). The creation of a new vocabulary levels test. *Shiken*, 19(2), 1–11. Retrieved from <http://www.lex tutor.ca/tests/levels/recognition/nvlt/paper.pdf>
- Munby, I. (2011). Development of a multiple response word association test for learners of English as an L2. Unpublished PhD thesis, University of Wales, Swansea.
- Munby, I. (2013). I-lex: An improved method of assessing L2 learner ability to see connections between words. *Vocabulary Education & Research Bulletin*, 2(2), 11–13.
- Nation, I.S.P. (1990). *Teaching and learning vocabulary*. Rowley, MA: Newbury House.

- Nation, I.S.P. (2001). *Learning vocabulary in another language*. Cambridge: Cambridge University Press.
- Nation, I.S.P. (2012). *The BNC/COCA word family lists*. Retrieved from http://www.victoria.ac.nz/lals/about/staff/publications/paul-nation/Information-on-the-BNC_COCA-word-family-lists.pdf
- Read, J. (1993). The development of a new measure of L2 vocabulary knowledge. *Language Testing*, 10(3), 355–371.
- 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: Erlbaum.
- Read, J. (2012). Piloting vocabulary tests. In G. Fulcher & F. Davidson (Eds.), *The Routledge handbook of language testing* (pp. 287–299). London: Routledge.
- Richards, J. (1976). The role of vocabulary teaching. *TESOL Quarterly*, 10, 77–89. doi:10.2307/3585941
- Schmitt, N. (1998a). Tracking the incremental acquisition of second language vocabulary: A longitudinal study. *Language Learning*, 48, 281–317. doi:10.1111/1467-9922.00042
- Schmitt, N. (1998b). Quantifying word association responses: What is native-like? *System*, 26, 389–401. doi:10.1016/S0346-251X(98)00019-0
- Schmitt, N. (1999). The relationship between TOEFL vocabulary items and meaning, association, collocation and word-class knowledge. *Language Testing*, 16(2), 189–216. doi:10.1177/026553229901600204
- Schmitt, N. (2000). *Vocabulary in language teaching*. Cambridge: Cambridge University Press.
- Schmitt, N., & Meara, P.M. (1997). Researching vocabulary through a word knowledge framework: Word associations and verbal suffixes. *Studies in Second Language Acquisition*, 19, 17–36. doi:10.1017/S0272263197001022
- Schmitt, N., Ng, J.W., & Garras, J. (2011). The word associates format: Validation evidence. *Language Testing*, 28(1), 105–126. doi:10.1177/0265532210373605
- Schmitt, N., Schmitt, D., & Clapham, C. (2001). Developing and exploring the behaviour of two new versions of the Vocabulary Levels Test. *Language Testing*, 18(1), 55–88. doi:10.1177/026553220101800103
- The British National Corpus, version 3.2 (BNC XML Edition). (2007). *Distributed by Oxford University Computing Services on behalf of the BNC Consortium*. Retrieved from <http://www.natcorp.ox.ac.uk/>
- Webb, S. (2008). Receptive and productive vocabulary sizes Of L2 learners. *Studies in Second Language Acquisition*, 30(1), 79–95. doi:10.1017/s0272263108080042
- Wen, Z., Biedroń, A., & Skehan, P. (2017). Foreign language aptitude theory: Yesterday, today and tomorrow. *Language Teaching*, 50(1), 1–31. doi:10.1017/s0261444816000276.
- Wolter, B. (2002). Assessing proficiency through word associations: Is there still hope? *System*, 30, 315–329. doi:10.1016/S0346-251X(02)00017-9

- Wray, A. (2002). *Formulaic language and the Lexicon*. Cambridge: Cambridge University Press.
- Zareva, A. (2005). Models of lexical knowledge assessment of second language learners of English at higher levels of language proficiency. *System*, 33, 547–562. doi:10.1016/j.system.2005.03.005

Appendix 1. i-lex v2

Word associations puzzle

Name _____ Student number _____

What word is associated with the following sets of 3 words?

Example: drink, red, glass >>>> w i n e [4]

The first letter, "w", is given for you and the word has 4 letters [4]. The answer is "wine"

Write the answers on the spaces below. If you don't know the answer, leave it and go on to the next one.

- | | | | |
|-----------------------------|-------------|-----------------------------|-------------|
| 1. coffee, cup, time | t _____ [3] | 26. study, trees, mother | n _____ [6] |
| 2. club, up, together | j _____ [4] | 27. round, left, over | t _____ [4] |
| 3. eat, drink, good | f _____ [4] | 28. money, account, book | b _____ [4] |
| 4. post, write, box | l _____ [6] | 29. hand, shopping, paper | b _____ [3] |
| 5. work, money, employment | j _____ [3] | 30. teach, read, book | l _____ [5] |
| 6. science, painting, music | a _____ [3] | 31. see, hospital, come | v _____ [5] |
| 7. tree, fire, forest | w _____ [4] | 32. leader, me, behind | f _____ [6] |
| 8. country, city, village | t _____ [4] | 33. run, talk, over | w _____ [4] |
| 9. truth, down, tell | l _____ [3] | 34. game, birthday, playing | c _____ [4] |
| 10. bed, door, space | r _____ [4] | 35. god, think, trust | b _____ [7] |
| 11. long, head, cut | h _____ [4] | 36. now, later, near | s _____ [4] |
| 12. minute, day, time | h _____ [4] | 37. car, fast, light | s _____ [5] |
| 13. man, people, individual | p _____ [6] | 38. play, stage, now | a _____ [3] |
| 14. street, car, way | r _____ [4] | 39. fun, game, drink | p _____ [5] |
| 15. thought, good, mind | i _____ [4] | 40. see, people, girl | m _____ [4] |
| 16. time, clock, out | w _____ [5] | 41. past, future, here | p _____ [7] |
| 17. big, large, small | s _____ [4] | 42. push, out, hard | p _____ [4] |
| 18. wrong, left, away | r _____ [5] | 43. of, piece, time | p _____ [4] |
| 19. go, home, out | l _____ [5] | 44. red, test, cut | b _____ [5] |
| 20. hear, to, radio | l _____ [6] | 45. paper, reader, good | n _____ [4] |
| 21. for, stop, bus | w _____ [4] | 46. center, end, east | m _____ [6] |
| 22. front, back, top | s _____ [4] | 47. boy, teacher, college | s _____ [6] |
| 23. good, fortune, bad | l _____ [4] | 48. shopping, words, names | l _____ [4] |
| 24. stick, football, game | m _____ [5] | 49. over, out, fast | m _____ [4] |
| 25. fish, ball, drop | c _____ [5] | 50. out, kind, of | s _____ [4] |

Answer key.

1. Tea 2. Join 3. Food 4. Letter 5. Job 6. Art 7. Wood 8. Town 9. Lie 10. Room 11. Hair 12. Hour 13. Person 14. Road 15. Idea 16. Watch 17. Size 18. Right 19. Leave 20. Listen 21. Wait 22. Side 23. Luck 24. Match 25. Nature 26. Turn 27. Bank 28. Bag 29. Learn 30. Visit 31. Follow 32. Walk 33. Card 34. Middle 35. Believe 36. Soon 37. Speed 38. Act 39. Party 40. Meet 41. Present 42. Pull 43. Part 44. Blood 45. News 46. Middle 47. School 48. List 49. Move 50. Sort

Appendix 2. Translation Test

Write (single word) English translations for the following Japanese words.

Band 1	Band 2	Band 3	Band 4
1. 島 i_____	1. 薬 m_____	1. 霧 f_____	1. 2カ国語 b_____
2. 動物 a_____	2. 宮殿 p_____	2. 糊(のり) g_____	2. 赤道 e_____
3. 住所 a_____	3. 英雄 h_____	3. 靴下 s_____	3. 化石 f_____
4. 地球 e_____	4. 冬 w_____	4. 青あざ b_____	4. エビス s_____
5. 医者 d_____	5. 耳 e_____	5. 誓う s_____	5. いやみ s_____
6. 安い c_____	6. 嫌う h_____	6. 頬 c_____	6. 迷信 s_____
7. 弱い w_____	7. 空 s_____	7. ふくろう o_____	7. 歯磨き粉 t_____
8. 安全 s_____	8. 隠す h_____	8. 潜る d_____	8. カメ t_____
9. 農場 f_____	9. 歌手 s_____	9. 葬式 f_____	9. 滝 w_____
10. 深い d_____	10. 屋根 r_____	10. 泡 b_____	10. やなぎ w_____
11. 鉄砲 g_____	11. 牛 c_____	11. 雷 t_____	11. 弁護士 l_____
12. 科学者 s_____	12. 砂 s_____	12. 抱き合う h_____	12. ウナギ e_____
13. 雨 r_____	13. 野菜 v_____	13. はしご l_____	13. 天皇 e_____
14. 柔らかい s_____	14. 甘い s_____	14. 数学 m_____	14. 絶滅 e_____
15. 椅子 c_____	15. 草 g_____	15. 天井 c_____	15. 湿気 h_____
16. 女王 q_____	16. 骨 b_____	16. 熟した r_____	16. 熟語 i_____
17. 秘密 s_____	17. 親戚 r_____	17. まくら p_____	17. いかだ r_____
18. 機械 m_____	18. 葉 l_____	18. 豚肉 p_____	18. おみやげ s_____
19. 切符 t_____	19. たんぱく質 p_____	19. 訳す t_____	19. 日光浴 s_____
20. 雑誌 m_____	20. 妊娠 p_____	20. 蝶 b_____	20. マグロ t_____
21. 鳥 b_____	21. 大使 a_____	21. 知らない人 s_____	21. 人類学者 a_____
22. 寝る s_____	22. 公害 p_____	22. 階段 s_____	22. 天文学 a_____
23. 穴 h_____	23. 冗談 j_____	23. 絨毯 c_____	23. 簿記 a_____

Appendix Table continued.

Write (single word) English translations for the following Japanese words.

24. 橋 b_____	24. 競馬の騎手 j_____	24. 襟 c_____	24. まつげ e_____
25. 夢 d_____	25. お面 m_____	25. 誘拐 k_____	25. 遠足 e_____
26. 鍵 k_____	26. 式 c_____	26. 香料 s_____	26. タコ o_____
27. 温かい w_____	27. 城 c_____	27. 苳 s_____	27. 薬剤師 p_____
28. 燃える b_____	28. 茹でる b_____	28. 酸っぱい s_____	28. 学期 s_____
29. 人口 p_____	29. 鏡 m_____	29. 虹 r_____	29. 株式仲買人 s_____
30. 科学 s_____	30. 胃 s_____	30. かつら w_____	30. 竜巻 t_____
31. 約束 p_____	31. 黄色 y_____	31. 浅い s_____	31. 水槽 a_____
32. 気象 w_____	32. 認可 l_____	32. 垂直の v_____	32. 雪崩 a_____
33. 箱 b_____	33. 宝石 j_____	33. 菓 n_____	33. 学生寮 d_____
34. 二月 F_____	34. 俳優 a_____	34. 百合(ゆり) l_____	34. キリン g_____
35. 風 w_____	35. おもちや t_____	35. 怠け者 l_____	35. 旅程 i_____
36. 角 c_____	36. 塔 t_____	36. 動物園 z_____	36. ヒョウ l_____
37. 台所 k_____	37. 双子 t_____	37. 奴隸 s_____	37. サソリ s_____
38. 王 k_____	38. 盗む s_____	38. 奇跡 m_____	38. 継母 s_____
39. 明日 t_____	39. ささやく w_____	39. ろうそく c_____	39. ひまわり s_____
40. 電車 t_____	40. 隣人 n_____	40. 液体 l_____	40. 台風 t_____

Appendix 3. Translation test answer key

1. 島 island	1. 薬 medicine	1. 霧 fog	1. 2カ国語 bilingual
2. 動物 animal	2. 宮殿 palace	2. 糊(のり) glue	2. 赤道 equator
3. 住所 address	3. 英雄 hero	3. 靴下 socks	3. 化石 fossil
4. 地球 earth	4. 冬 winter	4. 青あざ bruise	4. エビ shrimp
5. 医者 doctor	5. 耳 ear	5. 誓う swear	5. いやみ sarcasm
6. 安い cheap	6. 嫌う hate	6. 頬 cheek	6. 迷信 superstition
7. 弱い weak	7. 空 sky	7. ふくろう owl	7. 歯磨き粉 toothpaste
8. 安全 safe	8. 隠す hide	8. 潜る dive	8. カメ turtle
9 農場 farm	9. 歌手 singer	9. 葬式 funeral	9 滝 waterfall
10. 深い deep	10. 屋根 roof	10. 泡 bubble	10. やなぎ willow
11. 鉄砲 gun	11. 牛 cow	11. 雷 thunder	11. 弁護士 lawyer
12. 科学者 scientist	12. 砂 sand	12. 抱き合う hug	12. ウナギ eel
13. 雨 rain	13. 野菜 vegetable	13. はしご ladder	13. 天皇 emperor
14. 柔らかい soft	14. 甘い sweet	14. 数学 mathematics	14. 絶滅 extinct
15. 椅子 chair	15. 草 grass	15. 天井 ceiling	15. 湿気 humidity
16. 女王 queen	16. 骨 bone	16. 熟した ripe	16. 熟語 idiom
17. 秘密 secret	17. 親戚 relative	17. まくら pillow	17. いかだ raft
18. 機械 machine	18. 葉 leaf	18. 豚肉 pork	18. おみやげ souvenir
19. 切符 ticket	19. たんぱく質 protein	19. 訳す translate	19. 日光浴 sunbathe
20. 雑誌 magazine	20. 妊娠 pregnant	20. 蝶 butterfly	20. マグロ tuna
21. 鳥 bird	21. 大使 ambassador	21. 知らない人 stranger	21. 人類学者 anthropologist
22. 寝る sleep	22. 公害 pollution	22. 階段 stairs	22. 天文学 astronomy
23. 穴 hole	23. 冗談 joke	23. 絨毯 carpet	23. 簿記 accounting
24. 橋 bridge	24. 競馬の騎手 jockey	24. 襟 collar	24. まつげ eyelash
25. 夢 dream	25. お面 mask	25. 誘拐 kidnap	25. 遠足 excursion
26. 鍵 key	26. 式 ceremony	26. 香料 spice	26. タコ octopus
27. 温かい warm	27. 城 castle	27. 苺 strawberry	27. 薬剤師 pharmacist
28. 燃える burn	28. 茹でる boil	28. 酸っぱい sour	28. 学期 semester
29. 人口 population	29. 鏡 mirror	29. 虹 rainbow	29. 株式仲買人 stockbroker
30. 科学 science	30. 胃 stomach	30. かつら wig	30. 竜巻 tornado
31. 約束 promise	31. 黄色 yellow	31. 浅い shallow	31. 水槽 aquarium
32. 気象 weather	32. 認可 license	32. 垂直の vertical	32. 雪崩 avalanche
33. 箱 box	33. 宝石 jewel	33. 巣 nest	33. 学生寮 dormitory
34. 二月 February	34. 俳優 actor	34. 百合(ゆり) lily	34. キリン giraffe
35. 風 wind	35. おもちゃ toy	35. 怠け者 lazy	35. 旅程 itinerary
36. 角 corner	36. 塔 tower	36. 動物園 zoo	36. ヒョウ leopard
37. 台所 kitchen	37. 双子 twins	37. 奴隷 slave	37. サソリ scorpion
38. 王 king	38. 盗む steal	38. 奇跡 miracle	38. 継母 stepmother
39. 明日 tomorrow	39. ささやく whisper	39. ろうそく candle	39. ひまわり sunflower
40. 電車 train	40. 隣人 neighbor	40. 液体 liquid	40. 台風 typhoon