COMPARATIVE CULTURE
The Journal of Miyazaki International College

Second-Language Vocabulary Growth

in the First Three Semesters of College

Phil Bennett & Tim Stoeckel

Abstract

This paper describes a project aimed at assessing English as a foreign language lexical knowledge of students in their first three semesters of study at college. Four equivalent forms of an instrument to test written receptive knowledge of the words in the General Service List (GSL) and the Academic List (AWL) were developed for this purpose. A validation study with 334 participants found that the instrument had satisfactory dimensionality, the vast majority of its items displayed good technical quality, and Rasch person reliability estimates ranged from .87 to .93 for the four forms. The main study used this instrument to track the vocabulary growth of 144 students from two cohorts over each of their first three semesters of college. On average, these students entered college with knowledge of approximately 1,440 (56.0%) of the words tested, a figure which increased to 1,790 (69.6%) after three semesters. With a minimum criterion set at 80% for demonstrating satisfactory knowledge of a word list, only a minority of students entered college with satisfactory knowledge of either the first or second half of the GSL or the AWL. After three semesters, the majority of students reached this threshold for the first half of the GSL but not for the other two word bands.

Introduction

Second language (L2) vocabulary was once viewed as both relatively easy to develop and less worthy of investigation than grammatical competence or the traditional four skills of reading, writing, listening, and speaking (Richards, 1976). However, when the focus of educational theorists began to fall on language learners as well as language itself, studies of metacognition found that successful learners prized lexical knowledge and recognized its multi-faceted nature (Wenden, 1986). Corpus linguistics has over the last twenty years taken lexical analysis out of the realms of researcher-intuition and provided it with an empirical footing. This work has revealed the complexity of lexical patterning and clarified the size of the challenge that L2 learners face (Hoey, 2005; Nation, 2006; Sinclair, 1991).

Lexical knowledge has been shown to be a strong predictor of performance of general language ability. Studies by Stæhr (2008) and Milton, Wade and Hopkins (2010) have found strong correlations between written and aural receptive vocabulary knowledge and all of the main language skills. The results of both studies, summarized in Table 1, indicate that written and aural vocabularies interact with other skills in different ways but also that there is a close overall relationship between vocabulary size and L2 language ability.

Table 1

Spearman Correlations for Vocabulary Size and the Four Language Skills

Type of Vocabulary Knowledge	Skill			
	reading	listening	writing	speaking
written[a]	.83	.69	.73	
written[b]	.70	.48	.76	.35
aural[b]	.22	.67	.44	.71

[a] From Stæhr, 2008. [b] From Milton et al., 2010.

Other studies have attempted to answer the question of how large a vocabulary is required to accomplish particular goals. A commonly-cited benchmark for reasonable, or gist comprehension of everyday discourse is 2,000-3,000 word families (Milton, 2009; Schmitt, 2010b), with a word family being defined as a headword plus its inflections and closely related derivations (e.g., *excite*, plus *excited*, *excites*, *exciting*, *excitement* and *excitedly*). For more demanding tasks, such as reading authentic text without difficulty, vocabularies as large as 8,000-9,000 word families may be required (Nation, 2006). These figures are supported by data from corpora, which show that the most frequently occurring 2,000 words of English cover almost 80% of the words used in typical academic text and around 90% of words in everyday conversation (Nation, 2001).

Another branch of inquiry has looked into the relationship between the percentage of words known in a text and overall comprehension (Hu & Nation, 2000; Laufer, 1992; Laufer & Ravenhorst-Kalovski, 2010; Schmitt, Jiang & Grabe, 2011). One finding common to all of these studies is that there appears to be a linear relationship between coverage, or the percentage of words known in a text, and reading comprehension; no evidence of a percentage at which comprehension markedly improves has been found, nor is there any indication of reading comprehension scores reaching asymptote, when increases in coverage no longer appear to affect comprehension. Unsurprisingly, the general conclusion is: the more vocabulary, the better the comprehension. This statement notwithstanding, two percentage figures have consistently been referred to as benchmarks for predicting learner comprehension of a given text. At present, 98% coverage is considered the point at which learners are likely to be able to read independently (Hu & Nation, 2000; Schmitt, Schmitt & Clapham, 2001). For assisted comprehension (i.e., with

teacher support and/or dictionary use), the picture is less clear, but a minimum of 95% coverage may be a reasonable estimate (Laufer & Ravenhort-Kalovski, 2010; Schmitt et al., 2011). Although the difference of only 3% between these figures may appear slight, it represents a change from one word in 50 being unknown to one word in 20. Over a text of 500 words, this would mean an increase from 10 unknown words to 25.

Dictionary use might be considered capable of compensating for poor lexical knowledge, but the research has been inconclusive. Whereas some studies have found no improvement in understanding of L2 texts with dictionary use (Bensoussan, Sim, & Weiss, 1984; Hulstijn, 1993; Nesi & Meara, 1991), others have found significant gains in comprehension (Knight, 1994; Shieh & Freiermuth, 2010). In these latter studies, however, learners with poor lexical knowledge understood less than learners with high vocabulary knowledge (Knight, 1994), even when given ample time to complete the reading task (Shieh & Freiermuth, 2010). This suggests that while dictionary use can be helpful, it is no panacea for limitations in vocabulary knowledge. A likely reason for this is that humans have a finite amount of processing ability which cannot be dedicated to comprehension of a text until automaticity of lexical processing has been achieved (Browne, 2008).

Frequency lists

As it is now widely recognized that vocabulary knowledge is closely tied to overall language competence, researchers have tried to identify the words most beneficial for learners to acquire. Word frequency is a useful guide in this regard as there is tremendous variation in the rate at which words occur in language. In a typical text, a relatively small proportion of different word families comprises a large percentage of all of the words in the text. Figure 1 (adapted from Nation, 2001) shows

how the text coverage provided by successively less frequent groups of words

becomes progressively smaller.

Figure 1. Percentage of text coverage provided by successive 1,000-word frequency
bands in the Brown Corpus (adapted from Nation, 2001).
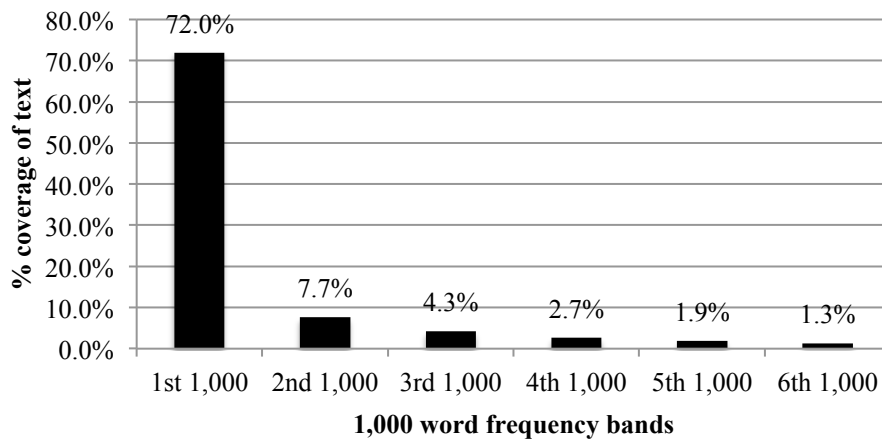
This implies that while frequency offers a useful initial guide for determining

which vocabulary to study, at some point learners will benefit from switching to a

more focused approach to lexical development that takes into account their individual

study purposes. Such an approach could be catered to by the General Service List

(GSL) (Bauman & Culligan, 1995; West, 1953) and the Academic Word List (AWL)

(Coxhead, 2000). The GSL was originally developed for use in writing simplified

reading materials, but since its composition took into account frequency, range of use,

and lack of specialized terms, it has become widely used as a notional core

vocabulary list. The GSL consists of 2,284 word families divided into two sublists

which approximate the first and second thousand words of English. The AWL was

developed to meet the needs of learners studying in an academic environment who

already demonstrate adequate knowledge of the GSL. As with the GSL, the AWL

used frequency and range of use as criteria for inclusion. The AWL contains 570

word families that are not in the GSL but that appear frequently in texts drawn from

four broad academic fields (arts, commerce, law, and science). Within these fields,

texts were sampled equally from 28 subject areas (Coxhead, 2000). Table 2 shows

coverage across a range of genres provided by the GSL and AWL.

Table 2

Text Type and Coverage Provided by the GSL and AWL

Levels	Conversation	Fiction	Newspapers	Academic
1st half of GSL	84.3%	82.3%	75.6%	73.5%
2nd half of GSL	6.0%	5.1%	4.7%	4.6%
AWL	1.9%	1.7%	3.9%	8.5%
Total	92.2%	89.1%	84.2%	86.6%

Note. Adapted from Nation, 2001.

Knowledge of the GSL and AWL could be expected to provide coverage of 84

to 92% of the words in a typical text, depending on genre. Although these figures are

still short of the 95% estimate for assisted comprehension described earlier, the

addition of proper nouns and the probability that students will also know some words

beyond these word levels will bring the targets closer.

Word frequency has also been identified as a predictor of whether a lexical

item is likely to be known by L2 learners (Milton, 2009). Several studies have shown

that if learners are tested on their knowledge of a range of words at differing

frequency bands, the most well-known words will be those of highest frequency, and

the least recognized will be those that appear least often in the language (Beglar,

2009; Richards & Malvern, 2007). Plotting the results of such tests produces charts

similar to that shown in Figure 2.

It should be noted that this is a general trend across groups. At an individual

level, it is not uncommon for learners to show a deficit at a particular frequency band.

Indeed, some studies (e.g., Milton, 2009) have found up to 40% of learners with

vocabularies that deviate to some degree from this trend. Likewise, all of the words

from a given frequency band should not be assumed to be of equal difficulty; factors

such as morphology, cognate status, part of speech and concreteness have also been shown to affect word recognition (Daulton, 2008; de Groot, 2006; Hayashi & Murphy, 2011, Stoeckel & Bennett, in press). Moreover, Milton (2009) has shown how Zipf's Law implies that the effects of word frequency diminish in less frequent word bands. Vocabulary tests that cover a broad range of frequency bands have shown evidence of this, with mean scores in mid- to low-frequency bands occasionally deviating from the general trend, as shown in Figure 3.

Figure 2. Frequency profile of word recognition by British learners of French. Adapted from Richards & Malvern, 2007.
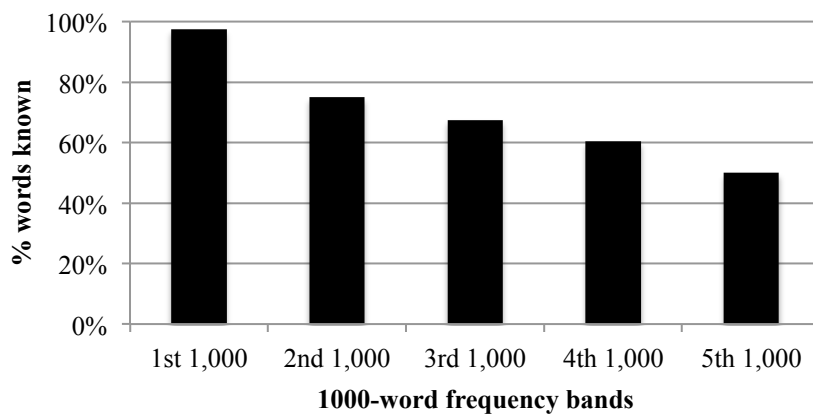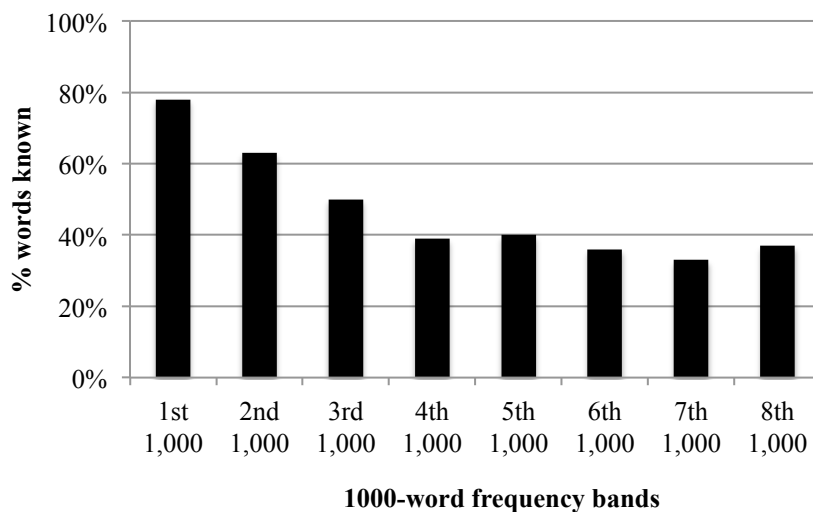

Figure 3. Frequency profile of word recognition by Japanese learners of English. Adapted from Aizawa, 2006.

Frequency can help to predict both the likelihood of a word being encountered and of it being recognized by learners. As such, it is a highly useful criterion to consider in text analysis and in establishing achievement goals for language learners. If learners were provided with specific frequency-based targets and their standing in relation to them, they would have salient goals by which they could evaluate their own progress in developing lexical knowledge. Goals that are specific and challenging yet attainable have been described by Dörnyei (2001) as important components of goal-setting theory. Goal-setting – along with planning, self-monitoring, metacognitive awareness, and use of learning strategies – is one of the components of Tseng, Dörnyei and Schmitt's (2006) construct of self-regulation in vocabulary acquisition.

Similarly, if teachers and administrators were aware of students' vocabulary profiles and growth over time, they would be better able to judge the developmental appropriateness of program materials and to assess program efficacy in relation to vocabulary goals.

Vocabulary Testing Instruments

Among the instruments widely used to provide estimates of receptive vocabulary knowledge are yes/no tests, the Vocabulary Levels Test (VLT), and the Vocabulary Size Test (VST). Yes/no tests have been used in education since at least the 1940s (Bear & Odbert, 1941). In this test design, learners are simply presented with a list of words and asked to indicate which words they know. Anderson and Freebody (1982) enhanced this design by adding a number of pseudowords to check and correct for overestimation of word knowledge. In L2 studies, Meara (1992) has made extensive use of yes/no tests, arguing that their simplicity allows for a far greater number of words to be tested than multiple-choice formats. Criticisms of

yes/no tests have centered on the use of formulas to correct for the selection of pseudowords, and the tendency for test takers from different language backgrounds to respond to the pseudowords in varying ways (Schmitt, 2010b). It should also be noted that yes/no tests require respondents only to indicate word recognition rather than to demonstrate receptive knowledge.

In the VLT, learners must match three out of six words to given definitions (see Figure 4). This test assesses vocabulary knowledge at the 2,000, 3,000, 5,000 and 10,000-word frequency bands, as well as having questions on words drawn from the AWL (Nation, 1983; Schmitt et al., 2001). Validation studies have been carried out on this instrument (Beglar & Hunt, 1999; Schmitt et al., 2001), and it has been widely used as both a diagnostic tool and in research to estimate vocabulary size (Laufer, 1998; Qian, 2002; Stæhr, 2008).

Figure 4. Sample item from the Vocabulary Levels Test. From Schmitt et al, 2001.

```
1  debate

2  exposure
                    _____        plan
3  integration
                    _____        choice
4  option
                    _____        joining something into a whole
5  scheme

6  stability
```

The third widely-used vocabulary measurement instrument is the VST (Nation & Beglar, 2007; Beglar, 2009). This is a multiple choice test in which learners read a sentence containing the target word in a natural yet non-defining context and select the correct definition from four choices (see Figure 5). The VST assesses knowledge of 10 words from each of the first to the fourteenth 1,000-word frequency bands of the British National Corpus. As a measure of overall vocabulary size, the VST is

designed to provide an indication of how successfully students will perform with certain materials and tasks and to assess how vocabulary grows over time.

Figure 5. Sample item from the Vocabulary Size Test. From Beglar, 2009.

1. **miniature**: It is a **miniature**.

a   a very small thing of its kind

b   an instrument for looking at very small objects

c   a very small living creature

d   a small line to join letters in handwriting

An important aspect of test design is that consideration should be given to the needs of test takers and the context in which they are learning. In academic settings, it is imperative that efforts are made to maximize recognition of the most frequent 2,000 words of English and to improve recognition of words on the AWL. This might suggest that the VLT would be a suitable instrument for such contexts, but there are reasons why it is less than ideal. First, the format of the VLT described above does not assess knowledge of the first 1,000 words of English. There is a separate VLT form containing picture items to assess these words, but to our knowledge, it has never been validated. Second, there are only two forms available of the most recent version of the VLT, meaning that any attempt to assess vocabulary growth longitudinally would risk a testing effect influencing the results as learners became familiar with the tested words.

For the VST, the same problem exists. Only two forms have been published, and a validation study has been conducted for only the first of these (Beglar, 2009). In addition, having only ten items to estimate knowledge of each 1,000-word frequency band may raise questions over the test's reliability if not over the entire 140-item form then certainly at each frequency band. In academic contexts with learners of low to

intermediate L2 proficiency, knowledge of total vocabulary size is of lesser importance than judging students' understanding of particular word bands. Finally, the lack of questions specifically targeting academic vocabulary means that students would not necessarily be tested on words they are likely to encounter in such contexts.

The following sections describe the development and initial validation evidence of a new test of vocabulary knowledge and then report on a study of English lexical development with students in the first three semesters at Miyazaki International College (MIC), a small English-medium liberal arts college in Kyushu, Japan.

Test Development

The primary purpose of test development was to address the shortcomings of existing instruments for repeated diagnostic assessment of L2 vocabulary knowledge in academic settings. Specifically, we aimed to develop four equivalent forms of a test of written receptive knowledge of the words on the GSL and AWL. This section briefly describes item and test form development and initial validation work. See Bennett and Stoeckel (in press) for a more complete description.

Item Development

Given their applicability to academic contexts, the GSL and AWL were selected as the word lists that would be used as the basis for the test. Eighty headwords were randomly sampled from each of the first 1,000 words of the GSL (hereafter GSL1), the second 1,000 words of the GSL (GSL2) and the AWL. Most test items were written following a blueprint similar to that of the VST. Each sampled headword was presented in bold-face type, followed by a short sentence that used the word in a natural, non-defining context. The Corpus of Contemporary American English (http://corpus.byu.edu/coca/) was consulted to confirm that one of the most

commonly occurring members of the target word family was used in these example sentences. This was followed by a definition of the target word and three distractors. Care was taken to use simplified language in test items so as not to introduce construct-irrelevant difficulty (Messick, 1995). That is, for test items targeting knowledge of words in the GSL, only words from the GSL1 were used, and for items targeting knowledge of words in the AWL, only words in the GSL were used. To avoid construct-irrelevant easiness (Messick, 1995), the distractors were written so as to be plausible substitutes for the target word.

A small number of test items differed in format from that of the VST in that pictures rather than text were used for the four answer choices. This approach was adopted in order to avoid having to use words in the answers which were of lower frequency than the target word itself. Two other deviations from the VST format were made after early piloting of the instrument revealed that, despite written directions to skip unknown words, many students had a high ratio of wrongly-answered to skipped items, suggesting that they were guessing and thereby inflating scores and reducing reliability (Zimmerman & Williams, 1965). To reduce guessing, a fifth choice (hereafter Choice E) was added which reads, "I don't know this word," together with the threat of a penalty in the test instructions. These two changes were retained when an examination of the data revealed both a reduced ratio of wrongly answered to skipped items (Bennett & Stoeckel, 2012) and an improvement in reliability estimates from .86 to .92. Sample text and picture items are provided in Figures 6 and 7.
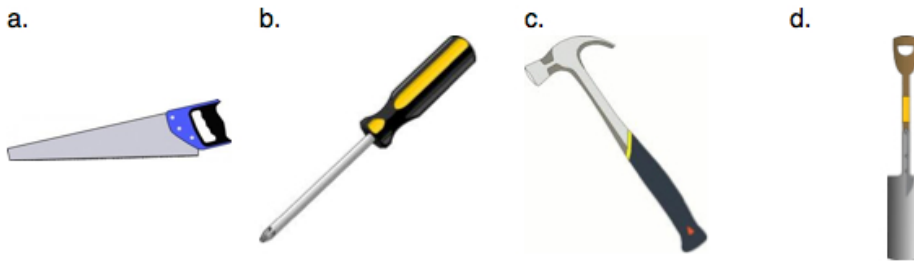
Figure 6. Example vocabulary test item with text-based answer choices.

bias: Be careful of **bias** in your writing.

a. grammar mistakes
b. language that is not exact
c. unfair opinions
d. informal language

e. I DON'T KNOW THIS WORD.

Figure 7. Example vocabulary test item with pictures as answer choices.

hammer: I need a **hammer**.

e. I DON'T KNOW THIS WORD.

Test Form Development

The data from piloting were also used to initially estimate item difficulties. Items were divided into four forms of 60 items each, with 20 items at each of the GSL1, GSL2 and AWL levels. The forms were balanced according to item difficulty, picture items, the parts of speech of the target words, and whether target words existed as loanwords in Japanese.

In order to assess the initial item difficulty calibrations and to judge the equivalence of test forms, some items were taken from their original forms and shared across the other forms to act as anchors (Wolfe, 2000). The end result was four 90-item test forms with 30 items at each level.

Summary of Validation study

An initial validation study was then conducted with students enrolled in two colleges in Japan ($n = 334$). The first was MIC, and the second was a medium-sized university of foreign language study in Osaka. The four 90-item test forms were spiraled in each of the participating 21 class sections. The data were analyzed using the Rasch dichotomous model with the Winsteps software package to determine construct dimensionality, the technical quality of items, the reliability of the four test forms, and the relative difficulty of the forms.

The instrument as a whole was found to be unidimensional, and all but four of the 240 test items displayed good technical quality. That is, more difficult items were likely to be answered correctly only by persons of higher ability, and easier items were typically answered incorrectly only by persons of low ability. The four items with poor technical quality were found to contain ambiguity or overly complex grammatical constructions in the wording. These four items were revised and will be monitored in future test administrations.

Rasch person reliabilities for the 90-item forms ranged from .92 to .95, and with the anchor items removed, the reliabilities for the 60-item forms ranged from .87 to .93.

Rasch analysis transforms raw scores to a logit scale, meaning that for every possible raw score, the logit values from each test form can be compared to assess equivalency. We found that for any given raw score across the four test forms, the logit values were within one standard error of each other. More specifically, Forms A and C were nearly identical in terms of difficulty, Form B was the most difficult, and Form D was the easiest. The difference in difficulty between Forms B and D was equivalent to about three points over the 60-item form.

Overall, the test forms appear to be good measures of the construct of written receptive vocabulary knowledge. The research method we have chosen allows for underperforming items to be identified, revised, and then placed back into the item pool to be re-assessed. In this way, the test forms undergo a process of continual refinement (Wise & Kingsbury, 2000). Similarly, because Rasch analysis produces quantitative estimates of difficulty for each test item, it will be possible to redistribute items among the four forms to more closely approximate the goal of equivalency. This is another step that we plan to take once we have obtained sufficient data on our revised items.

Study of Vocabulary Development at MIC

Purposes

The primary purpose of this aspect of the project was to estimate MIC students' knowledge of the words in the GSL and AWL both upon entry into MIC and as they progress through the first three semesters.

Method

Participants. Members of the MIC first-year cohort of 2011 (n = 74) participated in the study for two years. This group included 71 Japanese and three Korean students. Members of the 2012 cohort (n = 70), consisting of 67 Japanese, two Chinese, and one Korean, participated for the second year of the study.

Instrument. The instrumentation included Forms A, B, C, and D of the vocabulary test. During the study, the length of these test forms varied from 60 to 90 items as some items were shared across other test forms to (a) act as anchors for the validation study described above and (b) increase test reliability. A paper and pencil format was used except in the cases of 47 students with whom a computer-based version was piloted in December 2012.

Protocol. The test was administered six times during the 2011 and 2012 academic years, at the beginning and end of the spring semesters and at the end of the fall semesters. The 2011 cohort participated for the first five of these, and the 2012 cohort the final three. The basic protocol was for examinees to encounter the test forms sequentially such that those who used Form A for one administration would use Form B for the next, and so on. This protocol was followed except for in April 2012 when we changed from using the same test for each given administration (i.e., Form A in April, Form B in July, etc.) to spiraling all four test forms in each class section for each administration. Care was taken to record which test forms students received in April 2012, and the basic protocol has been followed since.

Measures. A formula which penalizes guessing was used to score the tests. Correct answers earned one point, skipped items or those answered with Choice E earned zero points, and wrongly answered items earned minus one-third. The result of this formula was used to calculate the percentage of items correct for each of the three sections of the test, which in turn were used to estimate the approximate number of words each student knew at the GSL1, GSL2, and AWL levels. For group means, these calculations were made twice, once with the complete data set and once removing students who missed one or more test administrations. The results were nearly identical, and those utilizing the full data set are reported below.

Results

Table 3 presents the means, standard deviations, and Rasch person-reliability estimates for each vocabulary test form in each administration. The instrument displayed satisfactory reliability, with coefficients ranging from .82 to .96 throughout the study and no lower than .89 after the addition of Choice E in December 2011.

Table 3

Measurement Properties of the Vocabulary Tests

test date	Test form (no. items)	n^a	M	SD	Rasch Person reliability
2011 April	A (60)	73	37.0	8.2	0.82
2011 July	B (90)	74	63.6	11.7	0.86
2011 December	C (90)	69	60.8	12.3	0.92
2012 April	A (90)	33	52.9	14.2	0.92
	B (90)	32	56.2	14.2	0.93
	C (90)	31	58.6	15.5	0.94
	D (90)	33	52.8	20.4	0.96
2012 July	A (72)	33	49.9	13.8	0.94
	B (72)	33	50.5	9.7	0.89
	C (72)	33	49.0	11.2	0.93
	D (72)	33	49.2	11.2	0.91
2012 December	A (72)	19	41.9	14.9	0.93
	B (72)	19	44.6	14.6	0.93
	C (72)	17	39.0	12.3	0.90
	D (72)	11	42.7	13.1	0.93

[a]For each of the three 2011 administrations, one version of the test was given to the available members of the 2011 cohort. In April and July 2012, the four versions were given at each administration to available members of both cohorts. In December 2012, the four versions were given to members of the 2012 cohort; members of the 2011 cohort were on study abroad.

Table 4 shows the means and standard deviations of the scores in each word level for both cohorts combined; the first row of data displays vocabulary knowledge upon entry into MIC, and the subsequent rows exhibit change over time. Keeping in mind that the GSL is roughly 2,000 word families and the AWL 570, the average student demonstrated knowledge of roughly 1,290 GSL and 150 AWL word families upon entry to MIC and experienced an average gain of about 280 words from these lists during the first semester, with a much smaller gain of about 70 words for the second and third semesters combined.

During the second semester, there was a loss in all three word bands, for which there is no clear explanation. Choice E was introduced at the end of the second

semester in 2011 (Year 1 December); this was a significant change in test format for the cohort of that year which could be a plausible cause of lower test scores. However, this explanation is unsatisfactory because the 2012 cohort, which did not experience a change in test format, also experienced a decrease during the second semester.

Table 4

Mean Percentage of GSL and AWL Words Known Through Three Semesters

date	n^a	% of Words Known (*SD*)		
		GSL1	GSL2	AWL
Year 1 April	139	70.7 (21.4)	58.3 (23.1)	27.0 (21.3)
Year 1 July	142	79.6 (17.2)	67.2 (19.3)	44.2 (22.4)
Year 1 December	129	77.5 (17.4)	63.4 (19.5)	39.2 (21.6)
Year 2 April	63	81.5 (14.0)	67.3 (15.6)	41.1 (18.9)
Year 2 July	64	82.9 (15.5)	69.6 (19.7)	46.7 (20.0)

[a]The discrepancy in the size of *n* between years 1 and 2 is due to the fact that the 2011 cohort has participated in the study for two years, and the 2012 cohort for just one.

There was considerable individual variation in vocabulary scores, which can be seen in Figures 8-10. These figures show the percentage of students who achieved satisfactory knowledge as opposed to moderate or large gaps in knowledge of each word band over time. The criterion for satisfactory knowledge was set at 80% based on Milton (2009), who demonstrated that groups of high ability learners achieve an average score of 85 to 90%, and not 100%, even for high frequency word bands. Given that this is an average with some learners scoring lower, we feel that 80% is an appropriate standard. The criterion for having large gaps in knowledge was arbitrarily defined as knowledge of less than 60% of a word band.

Figure 8 shows that less than half of the students demonstrated satisfactory comprehension of the GSL1 upon entry into college, and roughly three quarters did after three semesters. Figure 9 reveals that only a minority of students achieved satisfactory knowledge of the GSL2 after three semesters and Figure 10 that just 5% of students reached this level for the AWL in the same time frame.

Figure 8. Percentage of students at three levels of GSL1 knowledge during the first three semesters of college.
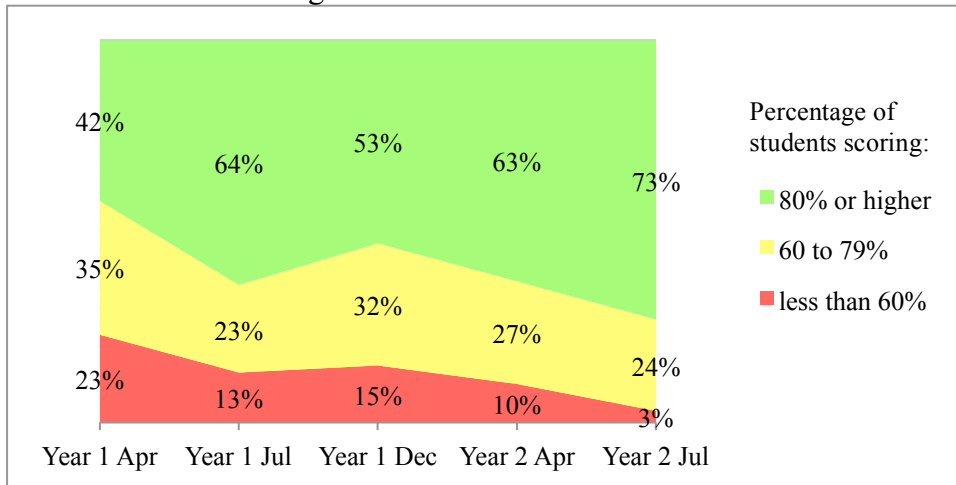

Figure 9. Percentage of students at three levels of GSL2 knowledge during the first three semesters of college.
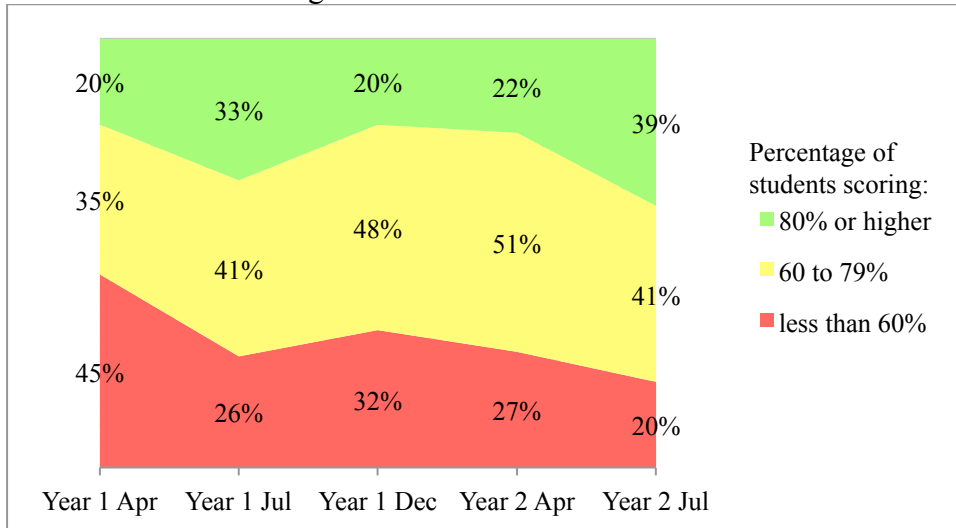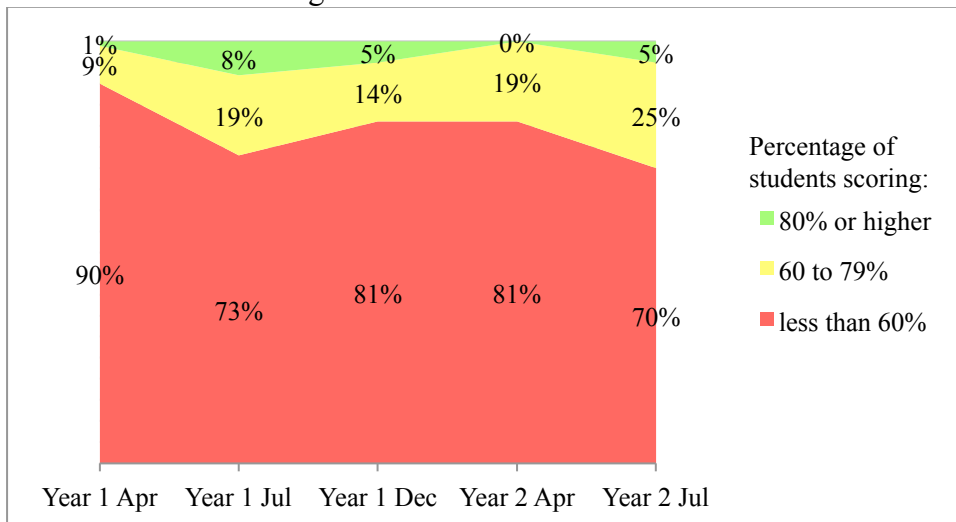

Figure 10. Percentage of students at three levels of AWL knowledge during the first three semesters of college.

Discussion

We have described a project aimed at assessing students' lexical knowledge in order to provide them with feedback for goal setting, to assess the efficacy of the curriculum in terms of promoting vocabulary development, and to help instructors make informed decisions regarding the lexical demands of materials and classroom activities. The instrument designed to accomplish these tasks has displayed satisfactory reliability and dimensionality, and the vast majority of its items have demonstrated good technical quality. Regarding vocabulary knowledge, our results indicate that few students enter MIC with mastery of the GSL, the core vocabulary necessary for accomplishing everyday tasks in English, and virtually none with mastery of the AWL. There is clear evidence of vocabulary growth, but after three semesters most students do not appear to achieve satisfactory understanding of the GSL2 or the AWL.

Implications

Keeping in mind the figures presented in Table 2 regarding the coverage provided by the GSL and AWL for various genres of text, it is likely that the majority of MIC students in the first three semesters fall short of both the 98% coverage required for unassisted comprehension or the 95% figure for assisted comprehension of authentic texts.

How might this situation be addressed? There has been considerable debate over whether lexical knowledge is better acquired through explicit instruction or implicit, exposure-based, learning (Han and Ellis, 1998; Krashen, 1989). Most researchers would now argue that a well-planned vocabulary learning component within a larger educational program would allow for both approaches, since each complements the other (Sökmen, 1997).

The argument for explicit attention to vocabulary in the classroom is that it will lead to greater noticing, and therefore uptake, among learners (Schmitt, 2010a). It has also been shown that programs which draw attention to, and have learners work on vocabulary knowledge in class lead to greater gains than those which rely solely on incidental learning (Folse, 2004; Laufer, 2005). Hunt and Beglar's (2005) framework for vocabulary development describes three forms of explicit lexical instruction. The first of these is studying decontextualized lexical items. This involves the learning of new vocabulary, consolidation activities for previously learnt words, and expansion activities to raise awareness of word families, affixation, collocational patterns and secondary or abstract meanings. The second form of explicit instruction is training in dictionary use. Nation (2008) suggests that dictionary training helps develop word knowledge by providing examples of natural usage and requiring learners to consider the appropriate senses of polysemous words. The final form of explicit instruction is training learners in inferring word meaning from context. While this is actually a compensation strategy to aid reading rather than a vocabulary learning skill per se, it reinforces knowledge of affixes and word families. It should be stated, however, that some studies have found low success rates for inferring word meaning and extremely low rates of retention for words learned in this fashion (for reviews, see Laufer, 2003, 2005). To have a reasonable chance of inferring word meaning correctly, learners may require knowledge of around 98% of the contextual words (Hunt & Beglar, 2005), and any words learnt in this fashion ought to be reinforced with other learning activities.

As for implicit instruction, the clearest advantage is that it will allow access to a far greater range of vocabulary than could ever be provided through explicit classroom instruction. Simple time constraints mean that learners must acquire much

of their vocabulary from exposure to discourse. Implicit instruction also provides access to contextual, particularly collocational, word knowledge that can be difficult to teach in the classroom. And finally, implicit learning occurs alongside other class activities, making it ideal for programs with multiple goals (Schmitt, 2010a).

Hunt and Beglar (2005) argue that while explicit instruction can lead to gains in lexical knowledge, the learning it provides is of limited value unless it can be activated in context. Extensive reading (ER) programs are designed to provide as much exposure as possible to contextualized comprehensible input. Both quantity and quality of exposure are crucial here: vocabulary learning is incremental, and without review of previously encountered words through regular reading, any gains may quickly disappear. Similarly, ER material must be at an appropriate level for learners. If the vocabulary demands are too high, it will not be possible for fluent reading to occur because learners will have little chance of understanding without using a dictionary. Likewise, material that is too easy will not provide students with frequent enough opportunities to recycle recently-learned words (Nation, 2001).

Automaticity, or speed of access, is one aspect of vocabulary knowledge that must be nurtured for fluent language use to occur (Meara, 1997), and implicit learning activities are ideally suited to promoting this. Through activities that entail repeated encounters with key vocabulary, learners are given the opportunity to comprehend and use targeted language without having to process new meanings afresh. Integrated tasks, in which students might first encounter a word in written or audio-visual material and are then required to use it in discussion or a written response about the content of the material, are one way to achieve this. Other possibilities include narrow reading, in which learners read several texts on related topics, and mixer activities, in

which learners complete the same spoken activity several times with different partners (Hunt & Beglar, 2005).

Finally, in addition to teaching activities, consideration should be given to training students to be good vocabulary learners. The concept of self-regulation has been drawn from educational psychology and applied specifically to vocabulary learning in a foreign language (Dörnyei, 2005; Tseng et al., 2006). Self-regulation describes learners' ability to maintain commitment to learning by establishing goals and incentives, to increase self-awareness of learning style preferences and tendencies to procrastinate, to manage disruptive emotional states, and to control their environment so as to take advantage of positive influences and avoid negative influences that might hinder their learning (Tseng et al., 2006). These are abilities that can be developed through instruction and guided experience. Nation (2008) and Thornbury (2002) describe the need to develop positive attitudes toward vocabulary learning through training students in strategies for understanding, recording, reviewing, and utilizing vocabulary, and by creating opportunities for them to experience success that will enhance motivation.

On a broader scale, there is clear value in establishing both short- and long-term goals for lexical development at the level of individual courses, language programs, and institutions (Laufer, 1992; Nation, 2001, Schmitt, 2008). With data provided by studies such as this, instructors can make informed decisions regarding the appropriateness of instructional materials, the need for explicit instruction of particular lexical forms, and the suitability of learning goals for individuals or classes.

Testing can also play an important role in course programming. The instrument described in this paper is intended to provide diagnostic feedback for learners and instructors over multiple semesters of study; however, given the strong

relationship between lexical knowledge and performance in the four main language skills, achievement tests, the results of which are included in course grading, should be considered both at regular intervals throughout courses and at their conclusions (Nation, 2001).

Future Research

The findings of this study raise several questions worthy of further investigation. First, it would be informative to analyze a representative sampling of the teaching and learning materials used during the first three semesters at MIC in order to ascertain the lexical burden they place on students. Specifically, these resources could be analyzed with lexical profiling software such as that available on the Lextutor website (http://www.lextutor.ca/vp/) to determine the percentage of coverage provided by the GSL and AWL. This information, combined with the data we now possess on our students' vocabulary knowledge, would enable us to estimate the percentage of words our students are familiar with in the materials they encounter.

Second, and more important, it would be useful to assess students' level of comprehension of the texts currently used during the first three semesters at MIC. The 95 and 98% coverage required for assisted and unassisted comprehension (respectively) are probabilistic: learners whose lexical knowledge falls short of these figures are sometimes able to compensate by making use of world knowledge and/or reading strategies such as previewing texts, using context clues or making inferences (Laufer & Ravenhorst-Kalovski, 2010). In the MIC environment, students receive significant amounts of support in the form of not only direct vocabulary instruction and glossing of low-frequency words but also active learning techniques that engage learners with the ideas in a text in multiple ways.

A third line of inquiry would be into the rate of overall vocabulary growth of MIC students. The present study indicates that on average our students gain knowledge of approximately 350 words from the GSL and AWL during the first three semesters, or slightly less than one new word for every two hours of classroom exposure to English. Due to the lack of longitudinal studies on vocabulary development in the Japanese context, this figure cannot easily be interpreted. It should also be noted that our students learn low-frequency words in addition to those on the GSL and AWL. Our feeling is that the present rate of vocabulary growth should serve as a benchmark against which further developments in the college program can be compared.

Limitations

This study has several limitations. First, the measure of vocabulary knowledge is restricted to the written receptive aspect of the construct. Though this approach is consistent with the majority of existing studies, the degree to which students are able to understand English vocabulary in aural contexts or use it productively is unknown. Second, the instrument assessed only breadth of knowledge, that is, how well learners could match a word with its definition. Depth of word knowledge (e.g., common collocates, part of speech, polysemy) was not investigated. Third, because we assessed familiarity with only the GSL and AWL, we cannot make informed statements regarding overall vocabulary size. Fourth, the lexical knowledge of examinees with small vocabulary sizes may have been underestimated because these learners are likely to have been unable to understand some of the words in the item stems or answer choices. Previous research has demonstrated that learners score higher on tests of vocabulary when answer choices are in the L1 (Ruegg, 2007), but the degree of difference in scores would be mediated by both the size of learners'

vocabulary knowledge and the actual words used in the test items. Finally, we have evidence of differential item functioning (DIF) for some test items when comparing responses of Japanese to Korean respondents (Stoeckel & Bennett, in press). That is, even when differences in ability were accounted for, some test items were easier for one group over the other. Though we have identified the cause of DIF in some items as related to construct-relevant factors such as loanword status of English words in the Korean and Japanese languages, because equivalency of the four test forms was determined with item difficulty estimates derived from responses of only Japanese students, the equivalency of test forms for our Korean students (and also those of other nationalities) is uncertain. Because these students comprise a small percentage of our student body, it is unlikely that this impacts cohort-level estimates of growth over time, but it may result in less precision in reports of growth for these individual students.

Concluding Thoughts

These limitations notwithstanding, it is our hope that this project has had and will continue to have beneficial outcomes in the MIC context. For the past two academic years, participants have been provided with regular reports of individual vocabulary knowledge and growth together with study lists targeting the first sizeable gap in their understanding of the GSL and AWL (for an example, see the Appendix). Though obtaining estimates of vocabulary development is an important first step, there is a need to develop methods of instruction and guidance that will help students to fulfill their learning potential. With a clearer understanding of students' level of lexical comprehension, it should be possible to provide learners with more individualized support that will aid them in comprehending course materials and becoming more involved in classroom discourse. Finally, considering the relationship

between vocabulary size and performance in the traditional four skills, this project could play a role in helping the institution achieve its broader aims of producing graduates with high levels of language ability who are informed of global issues and confident in expressing their views.

References

Aizawa, K. (2006). Rethinking frequency markers for English-Japanese dictionaries. In M. Murata, K. Minamide, Y. Tono, and S. Ishikawa, (Eds.), *English Lexicography in Japan* (pp. 108-119). Tokyo: Taishukan-shoten.

Anderson, R. C., & Freebody, P. (1982). *Reading comprehension and the assessment and acquisition of word knowledge* (Technical Report No. 249). Retrieved from University of Illinois at Urbana-Champaign, Illinois Digital Environment for Access to Learning and Scholarship website: https://www.ideals.illinois.edu/bitstream/handle/2142/17540 /ctrstreadtechrepv01982i00249_opt.pdf?sequence=1

Bauman, J., & Culligan, B. (1995). *About the General Service List*. Retrieved from http://jbauman.com/aboutgsl.html

Bear, R. M., & Odbert, H. S. (1941). Insight of older pupils into their knowledge of word meanings. *The School Review*, *49*(10), 754-760.

Beglar, D. (2009) A Rasch-based validation of the Vocabulary Size Test. *Language Testing*, *27*(1), 101-118. doi:10.1177/0265532209340194

Beglar, D., & Hunt, A. (1999). Revising and validating the 2000 Word Level and University Word Level Vocabulary Tests. *Language Testing*, *16*(2), 131-162.

Bennett, P., & Stoeckel, T. (2012). Variations in format and willingness to skip items in a multiple-choice vocabulary test. *Vocabulary Education and Research Bulletin, 1*(2), 2-3.

Bennett, P., and Stoeckel, T. (in press). Developing equivalent forms of a test of general and academic vocabulary. *JALT 2012 Making a Difference: Conference Proceedings*.

Bensoussan, M., Sim, D., & Weiss, R. (1984). The Effect of Dictionary Usage on

    EFL Test Performance Compared with Student and Teacher Attitudes and

    Expectations. Reading in a Foreign language, 2(2), 262-76.

Browne, C. (2008). A Research-based approach to developing Portable Technology

    Software for testing and teaching high frequency vocabulary. In I. Koch (Ed.)

    CamTESOL Conference on English Language Teaching: Selected Papers

    Volume 4, 2008 (pp. 8-18). Retrieved from http://www.camtesol.org

    /Download /Earlier_Publications/Selected_Papers_Vol.4_2008.pdf#page=15

Coxhead, A. (2000). A new academic word list. *TESOL Quarterly, 34*(2), 213-238.

    doi:10.2307/3587951

Daulton, F. E. (2008). *Japan's built-in lexicon of English-based loanwords*.

    Clevedon: Multilingual Matters.

de Groot, A. M. B. (2006). Effects of stimulus characteristics and background music

    on foreign language vocabulary learning and forgetting. *Language Learning,*

    56(3), 463-506.

Dörnyei, Z. (2001). *Teaching and researching motivation*. Harlow: Pearson.

Dörnyei, Z. (2005). *The psychology of the language learner: individual differences in*

    second language acquisition. New Jersey: Lawrence Erlbaum Associates.

Folse, K. (2004). *Vocabulary myths: applying second language research to classroom*

    teaching. Ann Arbor: University of Michigan Press.

Han, Y. & Ellis, R. (1998). Implicit knowledge, explicit knowledge and general

    language proficiency. *Language Teaching Research, 2*(1), 1-23.

Hayashi, Y. & Murphy, V. (2011). An investigation of morphological awareness in

    Japanese learners of English. *Language Learning Journal, 39*(1), 105-120.

Hoey, M. (2005). *Lexical Priming*. Abingdon: Routledge.

Hu. M., & Nation, I. S. P. (2000). Unknown vocabulary density and reading
comprehension. *Reading in a Foreign Language*, *13*(1), 403-430.

Hulstijn, J. H. (1993). When Do Foreign- Language Readers Look Up the Meaning of
Unfamiliar Words? The Influence of Task and Learner Variables. *The Modern
Language Journal*, *77*(2), 139-147.

Hunt, A. & Beglar, D. (2005). A framework for developing EFL reading vocabulary.
Reading in a Foreign Language, *17*(1), 23-59.

Knight, S. (1994). Dictionary use while reading: The effects on comprehension and
vocabulary acquisition for students of different verbal abilities. *The Modern
Language Journal*, *78*(3), 285-299.

Krashen, S. (1989). We acquire vocabulary and spelling by reading: additional
evidence for the input hypothesis. *The Modern Language Journal, 73*(4), 440-
464.

Laufer, B. (1992). How much lexis is necessary for reading comprehension? In P. J. L.
Arnaud and H. Béjoint (Eds.), *Vocabulary and applied linguistics* (pp. 126-
132). London: Macmillan.

Laufer, B. (1998). The development of passive and active vocabulary in a second
language: same of different? *Applied Linguistics, 19*(2), 255-271.

Laufer, B. (2003). Vocabulary acquisition in a second language: Do learners really
acquire more vocabulary by reading? *The Canadian Modern Language Review,
59,* 565-585.

Laufer, B. (2005). Focus on form in second language vocabulary learning. *EUROSLA
Yearbook, 5*, 223-250.

Laufer, B., & Ravenhorst-Kalovski, G. C. (2010). Lexical threshold revisited: Lexical
text coverage, learners' vocabulary size and reading comprehension. *Reading*

in a Foreign Language, 22(1), 15-30. Retrieved from http://nflrc.hawaii.edu
/rfl/

Meara, P. (1992). *EFL vocabulary tests* (1st ed.). Swansea: Centre for Applied
Language Studies, University College Swansea.

Meara, P. (1997). Towards a new approach to modelling vocabulary acquisition. In N.
Schmitt and M. McCarthy (Eds.), *Vocabulary: description, acquisition and
pedagogy* (pp. 237-257). Cambridge: Cambridge University Press.

Messick, S. (1995). Validity of psychological assessment: validation of inferences
from persons' responses and performances as scientific inquiry into score
meaning. *American Psychologist 50*(9), 741-749. doi:10.1037//0003-
066X.50.9.741

Milton, J. (2009). *Measuring second language vocabulary acquisition*. Bristol:
Multilingual Matters.

Milton, J., Wade, J., & Hopkins, N. (2010). Aural word recognition and oral
competence in English as a foreign language. In R. Chacón-Beltrán, C.
Abello-Contesse, and M. Torreblanca-López (Eds.), *Insights into non-native
vocabulary teaching and learning* (pp. 83-98). Bristol: Multilingual Matters.

Nation, I. S. P. (1983). Testing and teaching vocabulary. *Guidelines, 5*(1) 12-25.

Nation, I. S. P. (2001). *Learning vocabulary in another language*. Cambridge:
Cambridge University Press. doi:10.1017/CBO9781139524759

Nation, I. S. P. (2006). How large a vocabulary is needed for reading and listening?
The Canadian Modern Language Review, 63(1), 59-82.
doi:10.3138/cmlr.63.1.59

Nation, I. S. P. (2008) Lexical awareness in second language learning. In J. Cenoz
and N. H. Hornberger (Eds.), *Encyclopedia of Language and Education, 2nd*

Edition, Volume 6: Knowledge about Language (pp. 167-177). New York: Springer.

Nation, I. S. P., & Beglar, D. (2007). A vocabulary size test. *The Language Teacher, 31*(7), 9-13.

Nesi, H., & Meara, P. (1991). How using dictionaries affects performance in multiple-choice EFL tests. *Reading in a Foreign Language*, *8*, 631-631.

Qian, D. D. (2002). Investigating the relationship between vocabulary knowledge and academic reading performance: an assessment perspective. *Language Learning, 52*(3), 513-536.

Richards, B., & Malvern, D. (2007) Validity and threats to the validity of vocabulary measurement. In H. Daller, J. Milton, and J. Treffers-Daller (Eds.), *Modelling and assessing vocabulary knowledge* (pp. 79-92). Cambridge: Cambridge University Press.

Richards, J. C. (1976). The role of vocabulary teaching. *TESOL Quarterly, 10*(1), 77-89.

Ruegg, R. (2007). The English vocabulary level of Japanese junior high school students. In K. Bradford Watts, T. Muller, & M. Swanson (Eds.), *JALT2007 Conference Proceedings*. Tokyo: JALT. Retrieved from http://jalt-publications.org/archive/proceedings/2007/E013.pdf

Schmitt, N. (2008). Review article: Instructed second language vocabulary learning. *Language Teaching Research, 12*(3), 329-363.

Schmitt, N. (2010a). Key issues in teaching and learning vocabulary. In R. Chacón-Beltrán, C. Abello-Contesse, and M. Torreblanca-López (Eds.), *Insights into non-native vocabulary teaching and learning* (pp. 83-98). Bristol: Multilingual Matters.

Schmitt, N. (2010b). *Researching vocabulary: A vocabulary research manual*.

Basingstoke: Palgrave Macmillan. doi:10.1057/9780230293977

Schmitt, N., Jiang, X., & Grabe, W. (2011). The percentage of words known in a text

and reading comprehension. *The Modern Language Journal*, *95*(1), 26-43.

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.

Shieh, W., & Freiermuth, M. R. (2010). Using the DASH Method to Measure

Reading Comprehension. TESOL Quarterly, 44(1), 110-128.

Sinclair, J. (1991). *Corpus, Concordance, Collocation.* Oxford: Oxford University

Press.

Sökmen, A. J. (1997). Current trends in teaching second language vocabulary. In N.

Schmitt and M. McCarthy (Eds.), *Vocabulary: description, acquisition and

pedagogy* (pp. 237-257). Cambridge: Cambridge University Press.

Stæhr, L. S. (2008). Vocabulary size and the skills of listening, reading and writing.

Language Learning Journal 36(2), 139-152. doi:10.1080/09571730802389975

Stoeckel, T., & Bennett, P. (in press). Sources of differential item functioning

between Korean and Japanese examinees on a second language vocabulary

test. *Vocabulary Learning and Instruction, 2*(1).

Thornbury, S. (2002). *How to teach vocabulary*. Harlow: Longman.

Tseng, W., Dörnyei, Z., & Schmitt, N. (2006). A new approach to assessing strategic

learning: the case of self-regulation in vocabulary acquisition. *Applied

Linguistics*, *27*(1), 78-102.

Wenden, A. L. (1986). What do second-language learners know about their language

learning? A look at retrospective accounts. *Applied Linguistics, 7*(2), 186-205.

West, M. (1953). *A general service list of English words*. London: Longman.

Wise, S. L., & Kingsbury, G. G. (2000). Practical issues in developing and maintaining a computerized adaptive testing program. *Psicológica, 21*, 135-155.

Wolfe, E. W. (2000). Equating and item banking with the Rasch model. *Journal of Applied Measurement*, 1(4), 409-434.

Zimmerman, D. W., & Williams, R. H. (1965). Chance success due to guessing and non-independence of true scores and error scores in multiple-choice tests: Computer trials with prepared distributions. *Psychological Reports, 17*(1), 159-165.

Appendix

Example Report of Vocabulary Knowledge

Figures A1 and A2 provide an example of the reports which are given to students after each test administration.

Figure A1. Example first page of vocabulary test reports for students. Estimated percentage of words known in the GSL1, GSL2, and AWL word bands is provided for the most recent and all previous test results.
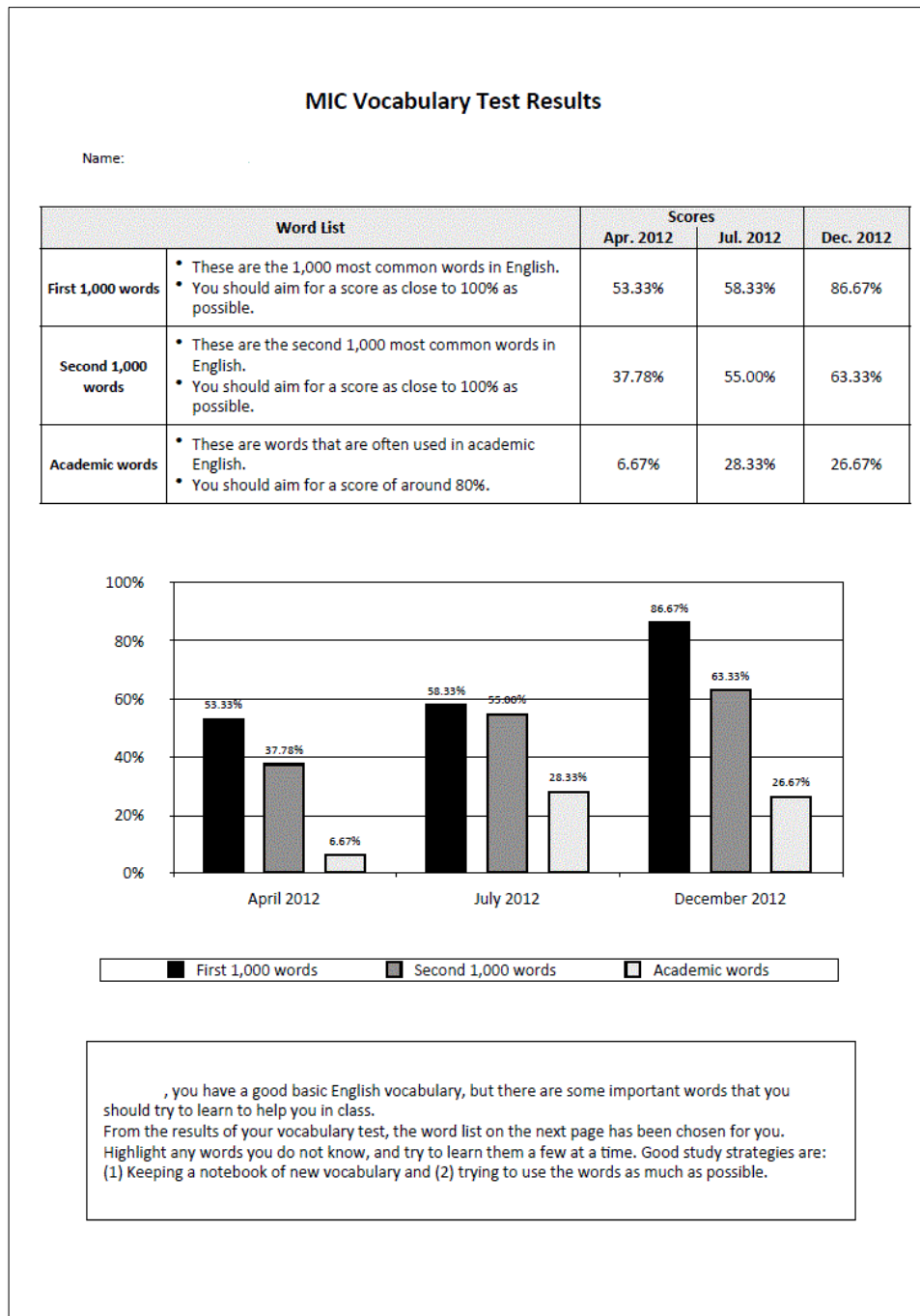

MIC Vocabulary Test Results

Name:

Word List		Scores		
		Apr. 2012	Jul. 2012	Dec. 2012
First 1,000 words	• These are the 1,000 most common words in English. • You should aim for a score as close to 100% as possible.	53.33%	58.33%	86.67%
Second 1,000 words	• These are the second 1,000 most common words in English. • You should aim for a score as close to 100% as possible.	37.78%	55.00%	63.33%
Academic words	• These are words that are often used in academic English. • You should aim for a score of around 80%.	6.67%	28.33%	26.67%

Legend: First 1,000 words ■ Second 1,000 words ■ Academic words □

_____, you have a good basic English vocabulary, but there are some important words that you should try to learn to help you in class.
From the results of your vocabulary test, the word list on the next page has been chosen for you.
Highlight any words you do not know, and try to learn them a few at a time. Good study strategies are:
(1) Keeping a notebook of new vocabulary and (2) trying to use the words as much as possible.

Figure A2. Example second page of vocabulary test reports for students. Page 2 provides students with a list of specific words to study based upon test results. The top of the page indicates to the student which words are included in the list (in this example it is part 2 of the second 1,000 words) and instructs students in how to use the information to address gaps in lexical knowledge.

Second 1,000 words (Part 2 of 2)

- These words are very common and would be useful for you to learn.
- Highlight any words you do not know and try to learn them.
- Ask your teacher for advice on building vocabulary.

quantity	weekend	merchant	cheap	shut	connect	cheer	straighten	objection	false	prevention	astonish
fun	treasury	coal	dish	confident	upset	nut	scarce	silver	fortune	urgent	shave
chicken	egg	mud	overcome	ruin	distant	split	lunch	tent	cap	aunt	feather
forgive	efficient	extension	cat	introduction	greet	melt	slavery	saddle	thread	zero	sauce
holy	grain	recognition	sacrifice	courage	flood	swear	creep	wrap	haste	idle	lid
wooden	calculate	kiss	complain	actor	excuse	sugar	sweat	nest	bare	fever	debt
prompt	drag	crop	elect	belt	insect	bury	gay	grind	shirt	Christmas	fade
crime	opposition	sail	roar	stir	ocean	wipe	stiff	spell	bargain	regret	confess
sorry	worship	attractive	sake	package	ceremony	faint	brave	plaster	leather	jaw	classification
republic	arrest	habit	temple	punish	decrease	creature	seize	arch	rail	soap	descend
anger	discipline	relieve	self	reflection	prize	tail	convenient	swell	butter	pronounce	cape
visitor	string	wisdom	compete	breathe	harm	wealth	horizon	friendship	dot	empire	mild
pile	harbor	persuade	nurse	anywhere	insure	earnest	moderate	bath	inquire	bowl	clever
violence	cow	certainty	stuff	amuse	verse	envelope	bleed	brass	obey	defendant	soup
steel	grand	cloth	stomach	dull	pot	invention	coin	taxi	hut	daylight	whistle
wing	funny	eager	peculiar	fate	sincere	sheep	fond	wax	axe	dip	scenery
delay	insurance	deserve	repair	net	cotton	splendid	autumn	duck	translation	suspicious	humble
gentleman	reduction	sympathetic	storm	fellowship	leaf	stamp	classify	button	collar	imaginary	fancy
pour	strict	cure	ton	fault	rub	float	omit	invent	delivery	ash	decay
medicine	translate	complicate	bundle	warmth	outline	brick	loyal	remedy	reproduce	carriage	polite
stroke	suspicion	dig	grateful	decisive	organ	rice	needle	bush	confession	educator	tribe
bite	noble	curse	crown	vessel	imitation	businessman	lessen	thunder	pan	saw	shield
lung	inquiry	weigh	boundary	pity	caution	backward	complaint	weaken	prejudice	stove	veil
lonely	journey	priest	nowhere	steam	mineral	qualification	pad	poverty	voyage	rubber	kneel
admission	hesitate	excessive	asleep	pin	disagree	artificial	steep	scrape	tobacco	rug	tray
stupid	extraordinary	quarrel	clock	bound	blade	attraction	skirt	arrow	simplicity	misery	explosive
scratch	borrow	widow	boil	companion	trick	lamp	curtain	tender	paste	awkward	log
composition	owe	modest	altogether	toe	treasure	curl	calculation	cruel	cake	rival	heap
broadcast	funeral	dine	lend	reward	immense	shower	laughter	soften	elephant	roast	suck
drum	ambition	politician	holiday	forbid	convenience	elder	solemn	mouse	ribbon	deed	ladder
resist	mixture	custom	precious	wherever	disapprove	bunch	grease	hay	harvest	preference	gap
neglect	slope	educate	wander	tower	destructive	bell	interfere	anyhow	ashamed	explosion	towel
absent	criminal	salesman	ugly	bathe	fork	steer	explode	alike	cave	theatrical	refresh
passenger	seldom	nail	reputation	lodge	noon	flavor	fasten	circular	customary	cultivate	essence
adventure	map	tap	ticket	swallow	ownership	spit	flag	juice	thief	collector	fur
beg	spin	eastern	pretend	multiply	tune	rob	resign	shelf	damp	miserable	ambitious
pipe	praise	possession	dismiss	bow	polish	cream	postpone	bake	sew	wrist	lump
beard	spare	satisfaction	delicate	kingdom	poison	interrupt	patience	hatred	rust	rabbit	annoy
bold	plow	behave	despair	garage	shame	pen	boast	cautious	separation	accustom	toy
meanwhile	telegraph	mercy	awake	permission	loyalty	weave	rope	basket	waiter	tide	heal
devil	barrel	scatter	tea	pump	cottage	orange	envy	wreck	pet	insult	shallow
rescue	airplane	width	straw	thumb	repetition	crush	rid				

An Analysis of Erich Fromm and Karen Horney through Cultural Psychology

Futoshi Kobayashi

Miyazaki International College

Address correspondence to Futoshi Kobayashi, Faculty of Comparative Culture, School of

International Liberal Arts, Miyazaki International College, 1405 Kano Kiyotake-cho

Miyazaki-shi Miyazaki-ken 889-1605 JAPAN; fkobayas@miyazaki-mic.ac.jp (e-mail).

Abstract

In this opinion paper, I analyzed Erich Fromm and Karen Horney, two of the famous Neo-Freudians from the perspective of cultural psychology. First, I described the common characteristics of Neo-Freudians. Second, I stated Fromm's and Honey's criticisms of Freud. Third, I summarized the philosophies of Fromm and Horney. Fourth, I elucidated why they chose specific values as the important ones for humans from their life experiences, cultural backgrounds, and the zeitgeist of their time. In conclusion, both scholars underscored the importance of achieving two opposing motivations (i.e., individuation from others and connectedness to others) for humans, yet their answers were ambiguous because the self-construct they used allowed only one of these two opposing motivations.

要旨

本論文では新フロイト派に分類される二人の巨人、エーリッヒ・フロムとカレン・ホーナイの思想が文化心理学の視点から論じられる。新フロイト派の思想の概略が説明された後、上記二人の学者が述べたフロイト批判と彼ら自身の思想が紹介される。続いて彼ら自身の思想が、各々の人生経験に加えて彼らの生きた時代や文化的背景から非常に強く影響を受けながら形成された事実を指摘する。結論として、フロムとホーナイが現代人に提言した生き方の処方箋は「自己確立」と「他者との関係性の確立」という二つの目標の達成を暗示しているが、各々の文化により異なる概念（例：自己意識）の存在が彼らの提唱した生き方の処方箋を不明確にしていると推論された。

An Analysis of Erich Fromm and Karen Horney through Cultural Psychology

In cultural psychology, psychological terms that are traditionally assumed as universal, such as self, mind, and emotion, are considered artifacts of culture (Kitayama, 1997; Miller, 1999). Therefore, individuals create their own subjective reality from their own unique experiences with other people, the societies they belong to, and the zeitgeist of their time. In this paper, Erich Fromm and Karen Horney, two of the famous Neo-Freudians, are analyzed from the viewpoint of cultural psychology. I discuss these two scholars in the following order. First, the characteristics of Neo-Freudians are explained. Second, Fromm's and Horney's criticisms of Freud are introduced. Third, Fromm's philosophy is outlined. Fourth, Horney's philosophy is summarized. Fifth, both Fromm and Horney's philosophies are culturally analyzed. At the end, conclusions are drawn.

1. Characteristics of Neo-Freudians

Historians of psychology assumed that Fromm and Horney belonged to the Neo-Freudian psychoanalytic school of thought (Schultz & Schultz, 1992). There were four major common assertions among Neo-Freudians. First, unlike Freud, sexuality was not a crucial factor for human behavior. Second, they thought environment, especially society and its own culture, played a much greater role in psychological development than Freud's developmental theory had suggested. Third, human relationships with other people were considered significant factors in psychological development. Fourth, the present was deemed more important than the past (Suzuki, 1992).

2. Criticism of Freud from Fromm and Horney

Fromm (1956/1989) explained his criticism of Freud as follows:

> My criticism of Freud's theory is not that he overemphasized sex, but his failure to understand sex deeply enough. He took the first step in discovering the significance of interpersonal passions; in accordance with his philosophic premises he explained them physiologically. In the further development of psychoanalysis it is necessary to correct and deepen Freud's concept by translating Freud's insights from the physiological into the biological and existential dimension (pp. 34-35).

Fromm focused on the relationship between an individual and society and emphasized the influence of modern capitalism upon the cognition and behavior of contemporary human beings. Horney also criticized Freud's theory as too physiological, deterministic, and mechanical and emphasized the influence of human relationships upon behavior and the importance of cultural factors (Suzuki, 1992). Nevertheless, Freud was also influenced by the zeitgeist of his time: faith in science (Schultz & Schultz, 1992). Freud tried to explain human beings from a mainly biological perspective (Sulloway, 1979) because most of the people in his time believed that science could solve all the problems they faced.

3. Fromm's Philosophy

For Freud (1930/1961), sexuality was the center of human life because he tried to understand human beings solely from their physiological functions.

> We said there that man's discovery that sexual (genital) love afforded him the strongest experiences of satisfaction, and in fact provided him with the prototype of all happiness, must have suggested to him that he should continue to seek the satisfaction of happiness in his life along the path of sexual relations and that he should make genital eroticism the central point of his life (p. 48)

Nevertheless, in Fromm's philosophy, the essential key for understanding human beings was not sexuality. Fromm (1941/1969) claimed that the most critical issue in

psychology was how an individual should relate to others, not simply through instinctual and physical gratification. According to Fromm (1941/1969, 1973), not only psychopathology but also violence and cruelty were the results of isolation, insecurity, and anxiety, which are by-products of the Renaissance and modern capitalism. Fromm (1941/1969) stated that in the Middle Ages, people believed that the universe was a single, giant organism in which each person had his or her own role under God's rule. Although human beings in this era had extremely limited individual freedom, people were very confident about their past, present, and future lives because everything was believed to be in God's almighty hands. However, since the Renaissance, modern science has developed so rapidly that it has taken over the absolute authority of God. Although modern human beings gained individual freedom and were liberated from the authority of the church, they lost social stability and a sense of belonging to community, nature, and God. In other words, people attained freedom but lost something on which they could depend. However, many of Fromm's critics complained that Fromm idealized human life in the Middle Ages (Knapp, 1989).

Additionally, modern capitalism makes a human being "a commodity, experiences his life force as an investment which must bring him the maximum profit obtainable under existing market conditions" (Fromm, 1956/1989, pp. 77-78). Therefore, many contemporary human beings are faced with feelings of isolation, insecurity and anxiety in their daily life. Thus, those who can not stand the feelings of isolation, insecurity, and anxiety are trying to escape from the freedom which promotes these feelings. The Germans, who were devoted to Nazism during World War II, might indicate that Fromm was correct because they followed Hitler without questioning his orders and threw away their cherished freedoms. Additionally, Fromm (1973) stated that the more people lose social belongingness and feel more alone, the more cruel, violent, and destructive they become.

Fromm's prescription for the disease of contemporary humans was an altruistic style of love and productive work for the sake of general welfare. Nevertheless, many of Fromm's critics complained that his explanations were unclear (Knapp, 1989).

In his later years, Fromm proposed that the ultimate aim of humans is self-realization: to attain the capacity of real love and productive work. His concept of "self-realization" was heavily influenced by Zen-Buddhism because of his longtime friend, Daisetz T. Suzuki: "the mediator of Zen in the West" (Funk, 1978/1982, p. 122).

4. Horney's Philosophy

Many psychological theories have been produced from the psychologist's own life experience. For example, Erikson felt excluded from any group of people throughout his life (i.e., he felt no sense of identity); as a result, he invented a new field of psychology: identity psychology (Tatara, 1990). Karen Horney's mother favored Horney's older brother more than Horney herself, and her father underestimated her figure and her mental ability (Schultz & Schultz, 1992). As a result, Horney herself developed a "basic anxiety," which she defined as a child's "feeling of being isolated and helpless in a world conceived as potentially hostile" (1950/1991, p. 18). This means that the "basic anxiety" was not innate but the result of an unsatisfactory relationship between the child and the parents. As for Freud, the major human drive was the satisfaction of libido. Nevertheless, Horney proposed that the major human drive was aiming "not at satisfaction but at safety" (1945, p. 13). For her, the utmost motivation of human behavior was to relieve "basic anxiety." Although Horney also admitted the magnitude of childhood experiences in Freud's theory, she did not accept a deterministic view of personality development. She wrote, "I believe that man can change and go on changing as long as he lives" (1945, p. 19). Rejecting a universal pattern of personality

development, Horney (1939) proposed that personality development was totally dependent upon cultural and social factors. Horney (1945) stated,

> And my impressions were confirmed when I came to the United States in 1932. I saw then that the attitudes and the neuroses of persons in this country differed in many ways from those I had observed in European countries, and that only the difference in civilizations could account for this (p. 12).

According to Horney, in attempting to relieve this "basic anxiety," people fall into three categories of maladjustment. First, "the compliant type, manifests all the traits that go with 'moving toward' people" (1945, p. 49). "In sum, this type needs to be liked, wanted, desired, loved; to feel accepted, welcomed, approved of, appreciated; to be needed, to be of importance to others, especially to one particular person; to be helped, protected, taken care of, guided" (1945, p. 51). The second category was the aggressive type of people who are always "moving against people" (1945, p. 63). They want to conquer everybody. The third category consisted of those who are "moving away from people" (1945, p. 72). Such people don't seek others' affection or fight against it, but they try to keep away from other people. They want to avoid any dependency. Nevertheless, none of these behavior types was the ideal way to resolve basic anxiety. Horney (1945) wrote,

> The neurotic must be helped to retrieve himself, to become aware of his real feelings and wants, to evolve his own set of values, and to relate himself to others on the basis of his feelings and convictions (p. 220).

Horney (1945) believed that these three neurotic types are by-products of a negative and unsatisfactory childhood environment. According to Horney (1945), these three neurotic types can be prevented by warmth, understanding, and loving care in the secure family environment.

<u>5. Cultural Analysis of Fromm's and Horney's Philosophies</u>

Both Fromm and Horney offered two layers of answers to contemporary human beings' conditions. The first layer of Fromm's answer was to accomplish non-possessive love and productive work. Fromm's second layer was self-realization. The first layer of Horney's answer was having love, warmth, understanding, and security. Horney's second layer was to find the real self, which is hidden under the idealized self-image, and to live honestly according to one's real self.

In their first layers, both Fromm and Horney picked up several values, but neither of them could explain why these values were more essential than other values. I assume that they chose these values for three reasons. First, their Western culture traditionally and historically had regarded these values as more significant than other values. Second, the zeitgeist of their time and their own life experiences and environments led them to choose these values unconsciously. Third, they were longtime friends who shared ideas until their friendship ended when Fromm diagnosed Horney's daughter with a psychological problem caused by her mother (Knapp, 1989).

In Western culture, the fundamental way of thinking originates from Judaism and Christianity. In Christianity, the highest value is God and God is love. Although the definition of love in English varies from person to person (Sternberg & Barns, 1988), love in Christianity only means non-possessive, altruistic love: loving other humans without expectation of any reward. Additionally, the Protestant ethic emphasizes productive work. Germany, where Fromm and Horney came from, used to be a Protestant country before the rise of Nazism.

A brief review of Fromm's life shows he was born in Germany in 1900. Fromm was raised in an Orthodox Jewish family (Fromm, 1994; Schaar, 1964). As an extraordinarily intelligent student at the Universities of Heidelberg and Munich, Fromm witnessed the

breakdown of the German Protestant cultural environment (Schaar, 1964). In addition, many

of Fromm's ancestors were Rabbi and scholars of the Talmud. He came to the United States

in the early 1930's in order to escape the Nazi persecution of the Jews. From that time on he

lived in the U. S. and Mexico and died in 1980 (Fromm, 1994; Schaar, 1964).

Schaar (1964) described Fromm's life as follows:

> War, cultural chaos, psychoanalytic explorations, homelessness, and totalitarianism - these are the epochal features of the world of Erich Fromm. He lives in a day when the sun has gone out of the human condition, and all his writing starts with the conviction that the life of Western man has gone desperately wrong. This sense of urgency which pervades Fromm's work has made of it an ambitious system of social criticism (p. 4).

Because of these life experiences, Fromm denied the fundamental core of Western

civilization: God. Therefore, when he had to choose the most important value for human

beings, he did not choose God. Yet, he chose the one and only synonym of God:

non-possessive love. Unconsciously, Fromm might have chosen the one and only synonym of

God because he claimed to be a humanist in his conscious sphere. Nevertheless, both his

family background as an Orthodox Jew and the cultural environment of the German

Protestant atmosphere made him choose non-possessive love and productive work as the core

values of human beings.

The development of Horney's theory is also similar to Fromm's situation. As

mentioned before, Horney did not obtain love, warmth, understanding, and security from her

parents in her childhood. When she turned fourteen, she began experiencing adolescent

crushes in order to gain these values. Schultz and Schultz (1992, pp. 479-480) described her

as follows.

> At seventeen she started a newspaper called 'a virginal organ for supervirgins' and took to walking the streets frequented by prostitutes. "In my own imagination," she

confided to [sic.] her diary, "there is no spot on me that has not been kissed by a burning mouth. In my own imagination there is no depravity I have not tasted, to the dregs" (Horney, 1980, p. 64, cited in Schultz & Schultz).

Although she married, her married life was a disaster. She was unable to enjoy her sexual life with her husband and was depressed and had frequent stomach aches. Both before and after her divorce, she engaged in several affairs (Schultz & Schultz, 1992).

Feeling starved for these values during childhood and continually pursuing them throughout her life, Horney chose love, warmth, understanding, and security as the fundamental values of the human being.

Although Fromm's answer in his second layer was self-realization, he "failed to find that indestructible core called the self. His failure is due partly to bad metaphysics and partly to bad logic" (Schaar, 1964, p. 67). But I can suppose his concept of self-realization is very similar to that of Zen-Buddhism because of the influence of his longtime friendship with Daisetz T. Suzuki. "For Suzuki, Zen is 'the quintessence and the spirit of Buddhism' and 'the teaching of the heart of Buddha'"(Funk, 1978/1982, p. 122). Horney also stated in her second layer that self-realization is the answer. She defined the real self as a "central inner force, common to all human beings and yet unique in each, which is the deep source of growth" (1950/1991, p. 17). Her description of the real self mirrors one of the main concepts of Buddhism: The one is whole and the whole is one. In Buddhism, self-realization means to understand this concept not only rationally but also through one's total existence (i.e., body, soul, and mind). This kind of understanding is termed as *satori* in Zen-Buddhism. It is a kind of "experience of enlightenment" (Funk, 1978/1982, p. 122).

Therefore, both Fromm's and Horney's self-realization accomplishments seemed similar to Buddhism. In attainment of *satori*, individuality and relatedness with others become one integrated accomplishment. In Zen-Buddhism, happiness can not be thought of from the individual perspective because we can not become happy in isolation from others.

Zen-Buddhism says that we can not become happy in confronting other people's agony because all living beings are connected and interdependent. This assertion is supported by some social scientists in the West. For example, Durkheim (1897/1966) studied suicide, an individual behavior, and found that religion, gender, and marital status influenced the probability of suicide in individuals. He concluded that as individuals lose meaningful connections to society, they begin to feel that life is meaningless, and society itself also begins to suffer from anomie, the feeling of despair and meaninglessness (Durkheim, 1897/1966).

Many terms in Zen-Buddhism are context dependent, losing all meaning when removed from their context. Likewise every human being needs their social environments to give meaning to their lives. Each person gains his or her uniqueness from one's relationships with others. Therefore, we need others in our life for our mental and physical well-being. Actually, Berscheid and Reis (1998) found that people consider high quality, close human relationships as more important than anything else for psychological and physical health.

Several scholars have pointed out that individualism is one of the fundamental ideologies in the United States (Bellah, Madsen, Sullivan, Swidler, & Tipton, 1985; Berscheid, 1999; Sampson, 1988; Triandis, 1995). In order to sustain individualism, two core values are implied in every aspect of American life: autonomy and freedom from any external force. For example, many schools of psychotherapy (Sue & Sue, 1990), the educational system (Tobin, Wu, & Davidson, 1989), and the institution of marriage (Dion & Dion, 1993) in the U.S. have attempted to adjust to the standard of autonomy and freedom. However, this ideology also risks isolation and the devaluation of human relationships, because conformity, obedience, and interdependence have been viewed as signs of weakness and helplessness in Western psychology (Markus & Kitayama, 1991, 1994). Furthermore, some have argued that psychopathology (Allen, Coyne, & Huntoon, 1998), obesity (Schumaker, Krejci, Small, &

Sargent, 1985), narcissism (Lasch, 1978; Mijuskovic, 1979), risky sexual behaviors (Miller & Paone, 1998), and violence (May, 1969/1989) in the U.S. are partially caused by the effects of isolation and the devaluation of human relationships.

Contemporary scholars have also proposed that every human being has two opposing motivations: individuation from others and connectedness to others, and that both motivations are essential to healthy human life (Guisinger & Blatt, 1994; Triandis, 1995).

Finally, both Fromm and Horney noticed a critical problem of contemporary individuals in modern societies. The individualism in Western society caused them to question the meaning of happiness of individuals out of social context. In addition, both of these scholars used the psychological term "self," which was referred to as "independent-self" by Markus and Kitayama (1991). This kind of self fundamentally assumes that one's self has clear boundaries and emphasizes the significance of independence and autonomy. Although both scholars' philosophies were similar to Zen-Buddhism, they were somehow critically different from Zen-Buddhism, because their psychological term for self is independent self, and psychological self in Zen-Buddhism may be interdependent-self (Markus & Kitayama, 1991). The interdependent-self assumes that uniqueness of an individual is created from one's relationships with others. Moreover, Suzuki (1967) underscored that in traditional Western thought, the existence of A is different from the existence of B or C. There is a clear boundary of individual unique existence. Yet in Zen-Buddhism, existence of A can become the existence of B or C within a particular context. Therefore, many Westerners have problems in utilizing Zen-Buddhism.

6. Conclusion

Both Fromm and Horney indicated that we need to obtain a balanced integration of individuality and relatedness for happiness long before excess individuality becomes a real

headache of contemporary American life. From a cultural psychologist's perspective, both Fromm and Horney thought and lived in their environments, the zeitgeist of their time, and their cultures. They chose their own unique values from their unique life experiences and searched for happiness from an individualistic perspective because the United States of America, the country to which they had recently immigrated, was the most individualistic country on Earth (Triandis, 1995). Their unique framework of zeitgeist and culture molded them to have their values and their philosophies of life. Both scholars searched for the answer of self-realization that mirrors Zen-Buddhism. However, their answers for attaining the two goals (i.e., individuation from others and connectedness to others) became vague because a crucial psychological construct, "self," they used allowed only one of these two motivations. The more independence and autonomy an individual with independent-self achieves, the more connections and the meaning of life he or she loses. Yet their unsatisfactory answers can be prevented. We can integrate some artifacts (e.g., psychological constructs, religion, formatted behavior patterns) of different cultures into our own psychological system. The crucial point is to acknowledge the critical differences among cultures (e.g., the construct of "self" in case of Fromm and Horney). We are likely to assume that the word means the same thing across cultures. However, the reality and the thoughts that are created by our languages might be significantly different to each person. The most fundamental point is to recognize the critical differences among cultures when we try to integrate artifacts of cultures interculturally.

References

Allen, J. G., Coyne, L., & Huntoon, J. (1998). Complex posttraumatic stress disorder in

    women from a psychometric perspective. *Journal of Personality Assessment, 70,*

    277-298.

Bellah, R. N., Madsen, R., Sullivan, W. M., Swidler, A., & Tipton, S. M. (1985). *Habits of*

    the heart: Individualism and commitment in American life. New York: Harper & Row.

Berscheid, E. (1999). The greening of relationship science. *American Psychologist, 54,*

    260-264.

Berscheid, E., & Reis, H. T. (1998). Attraction and close relationships. In D. T. Gilbert, S. T.

    Fiske, & G. Lindzey (Eds.), *The handbook of social psychology* (Vol. 2, 4th ed., pp.

    193-281). New York: McGraw-Hill.

Dion, K. K., & Dion, K. L. (1993). Individualistic and collectivistic perspectives on gender

    and the cultural context of love and intimacy. *Journal of Social Issues, 49,* 53-69.

Durkheim, E. (1966). *Suicide: A study in sociology* (J. A. Spaulding & G. Simpson, Trans.).

    New York: Free Press. (Original work published 1897)

Freud, S. (1961). *Civilization and its discontents* (J. Strachey, Trans.). New York: W.W.

    Norton. (Original work published 1930)

Fromm, E. (1989). *The art of loving.* New York: Harper & Row. (Original work published

    1956)

Fromm, E. (1969). *Escape from freedom.* New York: Avon (Original work published 1941)

Fromm, E. (1973). *The anatomy of human destructiveness.* New York: Henry Holt.

Fromm, E. (1994). *The Erich Fromm reader* (R. Funk, Ed). Atlantic Highlands, NJ:

    Humanities Press International.

Funk, R. (1982). *Erich Fromm: The courage to be human* (M. Shaw, Trans.). New York:

    Harper & Row. (Original work published 1978)

Guisinger, S., & Blatt, S. J. (1994). Individuality and relatedness: Evolution of a fundamental

    dialectic. *American Psychologist, 49,* 104-111.

Horney, K. (1939). *New ways in psychoanalysis.* New York: W.W. Norton.

Horney, K. (1945). *Our inner conflicts.* New York: W.W. Norton.

Horney, K. (1980). *The adolescent diaries of Karen, 1899-1911.* New York: Basic Books.

Horney, K. (1991). *Neurosis and human growth.* New York: W.W. Norton. (Original work

    published 1950)

Kitayama, S. (1997). Bunka shinrigaku towa nanika [What is cultural psychology?]. In K.

    Kashiwagi, S. Kitayama, & H. Azuma (Eds.), *Bunka shinrigaku: Riron to jitushou*

    [Cultural psychology: Its theory and verification] (pp. 17-43). Tokyo: Tokyo

    University Press.

Knapp, G. P. (1989). *The art of living: Erich Fromm's life and works.* New York: Peter Lang.

Lasch, C. (1978). *The culture of narcissism: American life in an age of diminishing*

    expectations. New York: Norton.

Markus, H. R., & Kitayama, S. (1991). Culture and the self: Implications for cognition,

    emotion, and motivation. *Psychological Review, 98,* 224-253.

Markus, H. R., & Kitayama, S. (1994). A collective fear of the collective: Implications for

    selves and theories of selves. *Personality and Social Psychology Bulletin, 20,*

    568-579.

May, R. (1989). *Love and will.* New York: Bantam Doubleday Dell. (Original work

    published 1969)

Mijuskovic, B. (1979). Loneliness and narcissism. *Psychoanalytic Review, 66,* 479-492.

Miller, J. G. (1999). Cultural psychology: Implications for basic psychological theory.

    Psychological Science, 10, 85-91.

Miller, M., & Paone, D. (1998). Social network characteristics as mediators in the relationship between sexual abuse and HIV risk. *Social Science and Medicine, 47,* 765-777.

Sampson, E. E. (1988). The debate on individualism: Indigenous psychologies of the individual and their role in personal and social functioning. *American Psychologist, 43,* 15-22.

Schaar, J. H. (1964). *Escape from authority.* New York: Harper & Row.

Schultz, D. P., & Schultz, S. E. (1992). *A history of modern psychology* (5th ed.). New York: Harcourt Brace Jovannovich.

Schumaker, J. F., Krejci, R. C., Small, L., & Sargent, R. G. (1985). Experience loneliness by obese individuals. *Psychological Reports, 57,* 1147-1154.

Sternberg, R. J., & Barnes, M. L. (Eds.). (1988). *The psychology of love.* New Haven, CT: Yale University Press.

Sue, D. W., & Sue, D. (1990). *Counseling the culturally different: Theory and practice.* (2nd. ed.). New York: Wiley

Sulloway, F. J. (1979). *Freud, biologist of the mind: Beyond the psychoanalytic legend.* New York: Basic Books.

Suzuki, D. (1967). *Toyo no kokoro*. [Mind of the East] Tokyo: Shunjuusha.

Suzuki, S. (1992). *Froit igo* [After Freud]. Tokyo: Kodansha.

Tatara, M. (1990). *Aidentity no sinrigaku* [Psychology of identity]. Tokyo: Kodansha.

Tobin, J. J., Wu, D. Y. H., & Davidson, D. H. (1989). *Preschool in three cultures.* New Haven, CT: Yale University Press.

Triandis, H. C. (1995). *Individualism and collectivism.* Boulder, CO: Westview.

Art, Heroism, and the Novels of Virginia Woolf

Eric Sandberg

The publication of Lytton Strachey's Eminent Victorians in 1918 represented a break with the Victorian conventions of memorially hagiographic biography. It was, in its clear Oedipal drive, an instance of what Hermione Lee has described as "striking the father dead" through a process of sly debunking (75). It was also, however, an equally clear rejection of the Victorian notion of the heroic individual. General 'Chinese' Gordon of Khartoum fame, for instance, is portrayed not as a prototypical imperial hero, but as a delusional religious maniac; Florence Nightingale not as a semi-divine angel of mercy and genius of compassion but as a power hungry, manipulative harridan. As such Eminent Victorians is typical of the intellectual ethos of the Bloomsbury group, and of at least important elements within post-war British society in general, in its persistent and searching questioning, challenging, and and ultimate rejection of versions of heroism which no longer seemed tenable in the post-war era.

Virginia Woolf was very much a participant in this anti-heroic programme. As a young woman she participated in the infamous Dreadnought Hoax of 1910, during which she and several other members of the Bloomsbury group dressed up, at the instigation of notorious prankster Horace de Vere Cole, as the Emperor of Abyssinia and his entourage and were given an official reception aboard the one of the navy's newest and most deadly warships, the H.M.S. Dreadnought (Wright 47). The ensuing scandal caused much outrage among right-thinking member of society. Woolf's delight with the prank indicates an early disdain for the pomp and mock heroism of the military, while in her writing she frequently ridiculed the heroic roles and trappings associated with Britain's young men and the Empire.

Her 1931 novel The Waves, the central character of Percival, around whom the six

narrative voices circle, is a good example of Woolf's mockery of heroic afflatus. From childhood Percival is a natural leader who inspires an almost religious awe in his school fellows:

> Look now, how everybody follows Percival. He is heavy. He walks clumsily down the field, through the long grass, to where the great elm trees stand. His magnificence is that of some mediaeval commander. A wake of light seems to lie on the grass behind him. Look at us trooping after him, his faithful servants, to be shot like sheep, for he will certainly attempt some forlorn enterprise and die in battle. (654)

The tone of this passage is carefully balanced between the undoubted sincerity of the narrator - here a young man named Louis - and the irony implicit in the scene: while the love he engenders is real, Percival is not a military commander but a cricket player, and and he leads not troops into battle but children into play. This simultaneous appreciation of the allure of the hero and satiric attack on heroic pretension is ultimately realised in Percival's fate. He dies an early death, but not as another narrator, Neville, imagines, riding "alone at the head of troops," denouncing "some monstrous tyranny," not that is to say in the approved heroic manner involving bravery, self-sacrifice, and the quest for the good of the community, but in a horse-riding accident: "His horse stumbled; he was thrown" (709). While it is not difficult to imagine a less heroic death, the offhand mundaneness of this death effectively undercuts Percival's heroic pretensions. His death is, from the heroic perspective, trivial and meaningless. While in The Waves "it is Percival who inspires poetry," who is the hero around which the other characters circulate, he is a hero made up of other people's fantasies about heroism, and ultimately his story is at least as much a "farce" as it is a tragedy (709).

<center>***</center>

This is an example, then, of one of the ways Woolf participates in Bloomsbury's, and very generally speaking modernism's, questioning and reformulation of the traditional figure of the hero, which occurred, as one would expect, in a number of different ways. Ford Maddox Ford's ironic treatment of Edward Ashburnham in The Good Soldier, for example, gradually and delicately, although quite thoroughly, undermines his moral stature. James Joyce's Leopold Bloom, on the other hand, gains enormously in moral stature as he moves through the novel, but is in many other respects decidedly unheroic, being, for instance, glutinous and excessively uxorious. However, one of the main ways that modernist fiction reexamined the very idea of heroism was though the figure of the artist-as-hero. This is perhaps only to be expected in an artistic movement so thoroughly fascinated by art as a "self-conscious practice" possessed of distinctive traits which differentiated it from other forms of social activity (Levenson 14).

The most obvious figure here is perhaps James Joyce's Stephen Dedalus whose commitment to "forge in the smithy of my soul the uncreated conscience of my race" represents a sort of rallying cry for the modernist self-imagination of the artist figure (Joyce 217). This is not to say that Joyce's Dedalus (who also appeared in the posthumously published Stephen Hero) is in any way a simple or unmodulated heroic figure, untouched by irony. One has only to think of his hydrophobia to realise that Joyce's artist hero is not without certain obvious flaws; heroes are not normally filthy and malodorous. Similar points could be made with reference to other artist-hero figures of the modernist canon, such as Lewis's Tarr, an unpleasant misogynist and more, after all, of an anti-hero than a hero. However, there remains the general tendency in early twentieth-century fiction to relocate heroic virtues and postures to the artist figure, be that figure a writer, a painter, a musician or an intellectual. In addition, it is possible to identify a set of values embodied in this modernist artist hero. Lee Lemon, for instance, argues in Portraits of the

Artist in Contemporary Fiction that pre WWII fiction represented artists as "isolated rebels" showing "Byronic contempt" for the commonplace world around them (Lemon ix, xii). Like Victorian or Edwardian heroes, who so often operated outside the boundaries of the society they sprang from and protected, be it H. Rider Haggard's African adventurers or Arthur Conan Doyle's eccentric and drug addicted detective, modernist artist heroes stand outside and above an uncomprehending society and bestow upon it the products of their creative labour.

To a certain extent, Virginia Woolf also participates in this modernist shifting of the locus of heroic value from the barrel of a gun to the barrel of a pen or paintbrush. Many of her novels focus on, or at least contain, an artist figure, and her representations of these artists sometimes employ a rhetoric of heroism. However, it is my contention that Woolf's vision of the artist as hero balances between a genuine belief in the heroic nature of the artist's vocation and a lively sense of the absurdity of the heroic posturing also associated with artistic work. The reading of Percival offered previously thus acts as a sort of model for my reading of not just Woolf's response to the heroic figure in general, but to the artist-hero in particular. Heroism is for Woolf at once vitally magnificent and faintly ridiculous.

<p style="text-align:center">***</p>

While it is of course impossible in a paper of this length to examine all of Woolf's artist-heroes, it is worth spending a few moments looking at how this role develops in her early fiction before moving on to the more canonical novels of her high-modernist period. In Woolf's first novel, the role of the artist hero is only slightly developed, and is shared between two characters: the female protagonist, Rachel Vinrace, and her romantic opposite, Terrence Hewet. Rachel is a pianist, whose passionate love of music coexists uneasily with her social and gender role - her aunts, for instance worry that she "will spoil" her "arms if" she insists "upon so much practising"

(15). Yet her music lies at the core of her character, and is valorised by the narrative. Rachel's improvisations at the piano lead the dance which creates a "gigantic circle" amongst the visitors to the hotel in Santa Marina, one of the text's primary representations of the potential for human unity, and her performance of Bach provides her audience with a sensation of meaning and order: "They sat very still as if they saw a building with spaces and columns succeeding each other rising in the empty space. Then they began to see themselves and their lives, and the whole of human life advancing very nobly under the direction of the music. They felt themselves ennobled [. . .]" (186 - 187). Similarly, Terrence Hewet is a writer whose ambition, if not necessarily his product, has heroic implications:

> "I want to write a novel about Silence," he said; "the things people don't say. But the difficulty is immense." He sighed. "However, you don't care," he continued. He looked at her almost severely. "Nobody cares. [. . .] As for the novel itself, the whole conception, the way one's seen the thing, felt about it, made it stand in relation to other things, not one in a million cares for that. And yet I sometimes wonder whether there's anything else in the whole world worth doing. (249)

Thus both of the artist figures in Woolf's first novel contain elements of the figure of the artist hero: isolation from the society around them, a sense of the importance of the act of artistic creation or performance, the role of art as a sacramental or unifying force, and a recognition of the extreme difficulty, even danger, of what might be called the artistic quest.

However, overall, it must be recognized that the emphasis of The Voyage Out falls not on the heroic aspects of art, but on the ways in which social and gender roles can interfere with the success of the artistic endeavour. Engaged to be married, Rachel and Hewet recognise the essential likeness of their different mediums - "What I want to do in writing novels is very much

what you want to do when you play the piano, I expect," as Hewet says - but find that there is a conflict between their respective artistic impulses (252). While Rachel, "the best musician in South America" plays "up and up the steep spiral of a very late Beethoven sonata [. . .] like a person ascending a ruined staircase, energetically at first, then more laboriously advancing her feet with effort until she could go no higher [. . .]" Hewet cannot write, and his response reveals the essential incompatibility of a conventionally structured and gendered relationship and genuine artistic endeavour: "I've no objection," Hewet says, "to nice simple tunes - indeed, I find them very helpful to my literary composition, but that kind of thing is merely like an unfortunate old dog going round on its hind legs in the rain" (340). As Ann Ronchetti has argued, "Woolf's earliest novels seem to suggest that heterosexual involvement takes a high toll on aspiring artists of both sexes" (13). To this I would add that homosexual involvement is never clearly presented as an option in Woolf's first novels - although there are certainly critics who would disagree with me here - and that the main conflict or incompatibility lies between the alternate impulses towards an artistic heroism and the need for human relationship.

<p style="text-align:center">***</p>

The Voyage Out was published in 1915, with the bulk of the composition and revision completed by 1913 (DeSalvo 104 - 106). By 1927, when Woolf published To the Lighthouse, her vision of the artist hero had shifted considerably. While the incompatibility of artistic and social life remains a theme, and indeed a richly developed one, the novel is also interested in the dual nature of artistic heroism, its daring bravery and its essential futility. Both of these ideas are developed largely through the figure of Lily Briscoe, an amateur painter, but they also rely on a series of contrasts which the narrative gradually developes between Lily the amateur woman painter and Mr Ramsay, the successful male philosopher. By taking Mr Ramsay as an example

here, I am implicitly endorsing a broad and expansive view of 'art' that would include many if not all acts of human intellectual endeavour, but this is a view that seems clearly supported by Woolf's novels, in which non-traditional creative activities such as mathematics or philosophy are placed alongside more usual types of art such as painting and writing.

In many ways Lily makes a peculiar artist hero: as Mrs Ramsay thinks, "one could not take her painting very seriously" (267). Lily herself agrees with this assessment of her work - she realises that "it would be hung in the

attics [. . .]; it would be rolled up and flung under a sofa" (372). Like Rachel, Lily embodies a conflict between the artistic life and the social life of marriage and procreation - she remains single in spite of Mrs Ramsay's attempts to marry her off, and dedicates herself to her art, and this in the face of the world's insistence, as voiced by the egregious Charles Tansley, that women "can't paint, can't write" (314). Mr Ramsay is also an odd character to embody the values of artistic heroism: he is petulant, prone to bullying his wife and children, curmudgeonly and eccentric. He "would," for instance, "talk by the hour about his boots," surely not a topic generally associated with heroic discourse (322).

On the other hand, both of these figures are strongly associated throughout To the Lighthouse with a wide range of heroic imagery. The scenes in which Lily is actively engaged in painting supply a great deal of the drama of the novel, and her artistic activity is described in terms of adventure and heroism. Her attempt to paint in what is recognisably a post-impressionist, rather than the more socially acceptable and fashionable impressionist idiom, is a struggle "against terrific odds to maintain her courage; to say: 'But this is what I see; this is what I see,' and so to clasp some

miserable remnant of her vision to her breast, which a thousand forces did their best to pluck

from her" (269). The act of painting, even as a neglected amateur, is or at least can be heroic, in spite of the fate that awaits the product of this heroism. Heroism is, it seems, the attempt, not the product. For Mr Ramsay, on the other hand, the relationship between heroic artistic production and posterity is considerably more fraught, and this is another place where gender difference seems to come into play.

For Mr Ramsay, the great question concerns the persistence of artistic production - here, as I have indicated, loosely defined to include intellectual endeavour in general: "And [. . .] fame lasts how long?" Mr Ramsay asks. "It is permissible even for a dying hero to think before he dies how men will speak of him hereafter. His fame lasts perhaps two thousand years. And what are two thousand years? [. . .] The very stone one kicks with one's boot will outlast Shakespeare" (Woolf, Lighthouse 279). This is a primary difference then between Lily Briscoe and Mr Ramsay; the former embraces, or at least accepts without complaint the inevitable prospect of neglect, for she has known little else throughout her life, while for Mr Ramsay the question of how he will be remembered is of vital importance.

This difference is textually framed within a series of images of heroism that both assert the heroic nature of artistic production and ironically counter-point it with the bombastic self-importance of what is here gendered as male self-aggrandisement. Mr Ramsay is shown struggling to think, to surpass his earlier philosophical achievements in a process of thinking, and this is represented as a performance on a keyboard divided into the 26 letters of the alphabet. Mr Ramsay can reach Q, and as "very few people in the whole of England ever reach Q" this is not an achievement to be despised (278). While Mr Ramsay remains stuck at Q, unable to reach R, the effort he makes is characterised in unabashedly heroic terms. He has, for instance "qualities that would have saved a ship's company exposed on a broiling sea with six biscuits and a flask of

water[. . .]" (278). This is only one of a series of comparisons which unite intellectual endeavour to the traditional Victorian notion of the imperial explorer hero. He is also likened to a member of "a desolate expedition across the icy solitudes of the Polar region" and "a leader who, now that the snow has begun to fall and the mountain top is covered in mist, knows that he must lay himself down and die before morning comes" (278 - 279). This is of course on one level utterly absurd and mere self-dramatisation, apiece with much of Mr Ramsay's behaviour throughout the novel. But the heroic impulse, the drive of the narrative towards a recognition of the heroism of the artistic act is also, I would argue, real if here radically undercut by the narrative's satiric approach. A brief glance at Lily's achievement as a painter will make this clearer.

Lily's work is as we have seen destined to be forgotten, and part of her heroism lies in her quiet acceptance of this fate, a resignation which is strongly contrasted to Mr Ramsay's futile yearning for an unachievable immortality. Yet Lily is as heroic in the practice of her work as Mr Ramsay is in the self-image he derives from his:

> One line placed on the canvas committed her to innumerable risks, to frequent and irrevocable decisions. All that in idea seemed simple became in practice immediately complex; as the waves shape themselves symmetrically from the cliff top, but to the swimmer among them are divided by steep gulfs, and foaming crests. Still the risk must be run; the mark made. (358)

The act itself of painting, and by implication artistic production in general - as she paints Lily murmurs to herself "like a work of art," as if to make the broader applicability of her heroic artistry apparent - is firstly one of active process rather than finished product, and secondly one of enormous and persistent danger (360). It might also be noted that the essential artistically heroic virtue is one of ordering, or bringing pattern to that which is apparently without form.

As I have tried to indicate earlier in this paper, this vision of the artist hero is hardly unique to Virginia Woolf. Rather, it is part of the broader modernist movement which relocated the heroic virtues - and Mr Ramsay's list of the heroic qualities might do as well as any other here, including "endurance and justice, foresight, devotion, skill" - from the man of action to the man of creation (Woolf, Lighthouse 278). In Woolf, this movement is complicated both by a recognition of the difficulty in reconciling the figure of the hero (be it the man of action or the man of creation) with the figure of the mother, wife, or lover, and by an awareness of the inherent fatuousness of the typically male heroic posture, be it associated with a warrior figure or a creative one.

Yet it is important when reading Woolf not to move too far towards an ironic or satirical reading, for the impulse towards heroism presented in her novels is extremely powerful, and to a large extent resists, and thus achieves an equilibrium with, their equally apparent debunking impulses. I will end my paper with the ending of To the Lighthouse, and the ending of Lily Briscoe's artistically heroic endeavour: "With a sudden intensity, as if she saw it clear for a second, she drew a line there, in the centre. It was done; it was finished. Yes, she thought, laying down her brush in extreme fatigue, I have had my vision" (Woolf, Lighthouse 391). All of the poignancy of Woolf's balanced vision of the artist hero lies in this famous passage. The intensity of artistic insight into the nature of the world, the pattern that lies behind the seeming randomness of life and death, is the heroic achievement, but it is at once momentary, and, grammatically, a fiction: Lily Briscoe does not see it clear, but is able to act as if she saw it clear - and that clarity of vision is the saving grace of artistic heroism.

Works Cited

DeSalvo, Louise A. Virginia Woolf's First Voyage: A Novel in the Making. London:

　　　Macmillan, 1980. Print.

Joyce, James. A Portrait of the Artist as a Young Man. New York: Penguin, 1999. Print.

Lee, Hermione. Biography: A Very Short Introduction. Oxford: Oxford UP, 2009. Print.

Lemon, Lee. Portraits of the Artist in Contemporary Fiction. Lincoln: U of Nebraska P,　1985.

Print.

Levenson, Michael. Modernism. Yale: Yale UP, 2011. Print.

Ronchetti, Ann. The Artist, Society & Sexuality in Virginia Woolf's Novels. New York:

　　　Routledge, 2004. Print.

Woolf, Virginia. To the Lighthouse. Selected Works of Virginia Woolf. Ware:　　Wordsworth,

2005. 253 - 391. Print.

---. The Voyage Out. Oxford: Oxford UP, 2009. Print.

---. The Waves. Selected Works of Virginia Woolf. Ware: Wordsworth, 2005. 635 - 779. Print.

Wright, E. H. Virginia Woolf. London: Hesperus, 2011. Print.

Toponymy & Environmental Identity

Jonathan Parker

Abstract

This article is the result of a study undertaken to compare indigenous toponyms in southern Chile (where I visited the UNESCO Cape Horn Biosphere Reserve with the Sub-Antarctic Biocultural Conservation Program) to toponyms in Texas (where I completed my dissertation research). Toponyms offer a means of understanding the relationship between people and the environments they inhabit. Toponymy as an expression of a cultural relationship to the environment has received little attention in the environmental philosophy literature, and it is the purpose of this article to help fill that void.[1] Due to the threats facing linguistic and cultural diversity, this paper also addresses the value of such diversity and offers some arguments for its preservation.

A study of toponymy can showcase the different sorts of environmental identities that are possible. Sharing the same root, *topos*, some toponyms offer topographical descriptors of the place where one finds oneself, calling attention to specific features of the landscape. For example, the city of Sapporo, Japan, derives its name from the Ainu language and can be translated as "river lined with large reed bed," or "large dried-up river" (Sapporo 2013). Some toponyms, however, tell us about the culture of the people who inhabit a place and where they came from while telling us nothing of the surrounding environment. For example, take the case of New York, whose name references historical connections to the English Duke of York. Toponyms can also serve a practical navigation

---

[1] David Abram engages language in his work and makes his case that the development of the alphabet, and alphabetic writing more precisely, have contributed towards the rupture of humans from nature. In his text at times he engages authors like Keith Basso (1996, 154-163), who have shown how American Indian toponyms are often rich descriptors of the natural environment, however, aside from a few references to naming Abram largely focuses on the function of the written word more generally.

function, while otherwise being relatively shallow in terms of conveying any other meaning, as in the case of alphabetic and numeric numbering common as a means of providing easy navigation around many cities.

Toponyms of the sort I am interested in for the purposes of this article encapsulate communal attitudes towards the natural environment; or, if not communal, a reflection of the attitudes of those in power. I argue that many non-Western indigenous toponyms are an outgrowth of a collaborative environmental identity whereby the inhabitants of an area are deeply informed in their personal identities by their natural surroundings. That is, personal and communal identities are a reflection and an extension of the surrounding environment. This, I contest, is not always the case with Western toponyms, which frequently indicate a rupture from the surrounding environment and a cultural imposition onto the environment that is not an extension from the surrounding environment.

A comparison of Western and non-Western toponyms can be informative and reveal the benefits of collaborative environmental identities for both the inhabitants and the environments of a given area. Attentiveness to toponyms is one aspect to focus on in the construction of more sustainable environmental identities for those of us in the West. That is, by examining how others use the power of language and naming to connect their identities to the surrounding environments, we may strive to create meaningful and place-specific toponyms for ourselves that can connect us in meaningful ways to the environments we inhabit. An outgrowth of this research has been a concern for linguistic and cultural diversity. The second part of this paper will discuss the threats to linguistic

and cultural diversity and present different arguments for the importance of preserving these forms of diversity.

Why Toponyms?

The question that motivated my study of toponyms was a question that Paul Shepard asked in an essay entitled "Itinerant Thoughts on Place:" What is the difference between discovering and finding place? Shepard claims that in the United States especially, "one finds place rather than discovers place…quality is not given, it is made" (1999, 185). I interpret him to mean "finds" in the sense that we use the word *found* as an act of creation—as in Denton was "founded" in 1857. Used in this way, we mean to say that something was established, that it was created anew. To discover, for Shepard, retains an element of wonder, surprise, and chance; there is a sense of encountering something already existing that does not originate from oneself. To discover place, then, is to be impressed by the quality of place in its given state, as opposed to creating and *finding* place in the sense of imposing one's own image upon a landscape.

There is something problematic about Shepard's use of the word discovery however, because phrases such as the "discovery of the new world" are commonplace in the average American vocabulary; when Americans are told as children that Columbus discovered the new world, we are indeed told that he chanced upon this new land. However, we all later learn that this discovery story ignores the presence of the indigenous inhabitants of the lands and paved the way for colonial conquests laying claim to that which had been "discovered." Using the sense of discovery as Shepard intends is more collaborative in that it is responsive to what is given, rather than an establishment upon a

perceived void. The latter has been more typical of the early exploration of America, whereby colonists sharing a mindset reflected in the philosophy of John Locke, perceive the landscape as a kind of nothingness—a waste of land waiting to be improved.[2] The same difference can be seen today in the way two different people can respond to the same tract of land. One person may respond positively to the land in its given state and wish it to continue in its given state. Another person may lament the plot of land in its given undeveloped state and see it as a waste of space until such time as someone with the means to do so could "improve" or develop that land for use as a super market, etc.

These varying types of relationships to environments are expressed via toponyms. The purpose of this article is to explore some of these different relationships through toponyms. I argue that one of the effects of colonization has been an attack on local toponyms in favor of imported ones.

The reason it is worth focusing on something like toponyms is to counter the largely technological focus of much of the discussions about contemporary environmental issues. Environmental issues are often conceived of as scientific/technological problems in need of scientific/technological solutions. Environmental philosophers contend that at the root of these problems are actually philosophical values and ideas. Technologies are not value free, and so even a discussion of environmental technologies requires discussion of the values underlying not only those technologies, but also our ideas about the environment and how we interact with it. Attentiveness to small things like toponyms is one way that we can begin to explore our attitudes towards the environment and our relationship with it. With

---

[2] For a good discussion of the influence of John Locke's philosophy on early American land-use attitudes, see Gene Hargrove's *Foundations of Environmental Ethics*, particularly chapter 2.

this understanding we can begin to search for ways to change our relationship and our attitudes if necessary.

Examination of Toponyms

There are at least three categories of toponyms: those that make historical/cultural references, those that make natural references, and those that make no meaningful references. I first provide some examples of the first two, as I think these are the more common types of toponyms found in the US; however the third has become increasingly common in areas with new development.

Let's begin with the example of Denton, Texas. This area was historically inhabited by the Wichita and Caddo American Indian tribes. Finding an original name for this city or this area is difficult because most available histories begin with the *founding* of Denton, which traces its origins back to 1857.

Returning to this distinction introduced by Shepard between *discovering* place and *finding* place; this act of finding or *establishing* place often violently displaces a previously existing place. To say that the city of Denton was established in 1857 can give the impression that nothing existed prior to that time; unfortunately this is rarely the case. In the case of Denton, as already mentioned, this area was Wichita and Caddo territory; John B. Denton was—in addition to being a prominent Methodist preacher and lawyer—a Texas Militia Captain and American Indian fighter. He died in the battle with the American Indian tribes of the region and the city was named in his honor.

There was a history and a place prior to Denton and it is important to realize that the establishment of Denton was simultaneously the displacement of this previously existing

place. That is, Denton was established *against* the indigenous place. The toponymy of newly established places often take their names from prominent figures involved in the act of settlement, or from places or people associated with the locations these settlers just left. Other prominent Texas cities and counties share this naming feature.

Fort Worth was named for Major William Jenkins Worth, who was a prominent soldier in the period of the Mexican-American War. Fort Worth is situated in Tarrant County, which is named for General Edward H. Tarrant, who similarly to Denton was an Indian fighter attempting to clear the area of Comanche and Kiowa settlements. Houston was of course named after the General Sam Houston, one time President of the Republic of Texas, who famously commanded the Texas Army to victory against the Mexican forces under Santa Anna. Austin, the state capital, was named for Stephen Fuller Austin, who, as a lieutenant colonel of the militia planned and sometimes led campaigns against local American Indian tribes earned the name the "Father of Texas." Finally, Austin is the county seat of Travis County, named after William Barrett Travis—commander at the battle of the Alamo.

For the most part then, the most prominent toponyms in Texas all refer to male military heroes who were significant in the establishment of what eventually became the state of Texas. These toponyms, rather than connecting us to the Texas landscape and its human and nonhuman inhabitants, connect us to a history of military conquest. In this respect then, they are more connected with Shepard's understanding of finding place, wherein those with the power to name overlook the naturally given and impose their own image upon the place. This fails to represent any element of collaboration with or

attunement to the naturally given; rather, it celebrates the overcoming of the naturally given.

To be fair, this is a somewhat selective listing of Texas toponyms, even though they are the most well-known Texas toponyms. There are Texas toponyms that retain a connection to the indigenous place, most notably the name Texas itself. It is speculated that name Texas comes from the Spanish—Tejas—which itself was adapted from the Caddo language word for friends, or allies—taysha.

While these toponyms may be retained, they do not perform the same function as indigenous toponyms. That is, they no longer maintain the same reference or function in the same manner for the people using those names. A more common tendency than retaining these original toponyms that cease to have meaning for the colonizers is to displace the indigenous place and import names that are disconnected from the local place.

Arguably the worst form of external imposition onto a given environment is the use of alphabetic or numeric names. New York City is a great example of alphabetic and numeric toponyms with its "alphabet city" in the Lower East Side of Manhattan and the expansive numeric grid starting from 14th street, going on into the hundreds and two hundreds all the way up to the Bronx, intersected by the long avenues, also numbered, and stretching from the East River to the Hudson. This mathematical grid makes navigating the city incredibly convenient, but is imposed abstractly upon the natural environment and tells us nothing of nothing of the environmental and cultural landscape[3].

---

[3] This has been made up for to some extent however by the common practice of affixing supplementary street signs underneath or above the regular alphabetic or numeric signs. These secondary signs often do tell passersby something of the history or culture of the area. This is the case, for example where signs

There is something in this reminiscent of Edmund Husserl's critique of the mathematicization of the world that he traces back to Galileo in *The Crisis of the European Sciences* (1970)—that is, we abstract from the world to create pure ideal forms and then bring them back to bear on the natural world, which may not always conform to this abstract mathematicization. This numeric grid makes navigating the city easy, but does not tell us anything about the landscape or region of New York; it rather imposes a human desire for order, convenience, and simplicity upon the land. This is arguably the worst sort of rejection of the agency of the given environment.

Perhaps even stranger than this alphanumeric system of naming is the practice of thematic or remote naming. Here, rather than the convenience of alphanumeric grids, entire neighborhoods are occasionally named, again without connection to the indigenous place, but according to the whim of a developer. This is especially noticeable in the new housing developments around the area where I lived in Denton, TX.

Before moving to Texas, I had never really witnessed the modern, suburban subdivision style housing developments so common in Texas and other areas before, where developers raze a large plot of land and in short order build similar style houses and put them up for sale. When I first moved to Denton, I was generously housed temporarily by a woman who lived in one of these thematically named subdivisions. Her house was located on Andalusian Drive. At the time I thought this was a reference to the area of southern Spain, which appeared to be a randomly chosen name with no connection to Denton, Texas.

underneath the "Ave C" sign read "Loisaida Ave." Loisaida provides testament to the Latino inhabitants of the area who referred to the "Lower East Side" area as "Loisaida" (Roberts 2012). Although not generally used, the 1945 designation of 6th avenue as the "Avenue of the Americas" also provides an interesting historical story of the avenue (Barry 2005).

I later realized it referred to the breed of horse, and the rest of the community was likewise equestrian themed; there were names like Bareback Lane, Saddleback Drive, English Saddle Lane, and Paddock Way. These were mixed in with other breed names such as Lippizan Drive, Clydesdale Drive, and Rocky Mountain Lane. Among these were names only tangentially related to equestrian culture such as Countess Lane, Butler Drive, and Spanish Lane.

This form of toponymy reflects back to the inhabitants something of their cultural and intellectual inheritance, but denies or at least obscures any natural referent; it tells us little to nothing about the surrounding nonhuman environment. In this sense, it may be even less desirable than the colonizing toponymy with its historical associations. At least the latter category of toponyms keeps alive some sense of the history of the place, for better or for worse.

Contrasting Indigenous and Colonial Toponymies

Indigenous names frequently are reflective of the local environment and serve as reminders, descriptions of, and connections to what is indigenously present. If it is true that the type of naming present in indigenous populations can help maintain and keep alive a connection to the natural environment, this might be the type of environmental identity-forming method that we can learn from indigenous cultures in constructing our own identities. For obviously, it would not make sense for us in Denton to name our towns or parks after aboriginal Australian places, flora, or fauna; so we should not appropriate indigenous names, but we can adopt that approach to toponymy and apply it locally. First let's turn to some examples of indigenous toponymy.

On a visit to the Cape Horn region of Chile, I stayed for a time in Puerto Williams, which is billed as the southernmost city in the world. Puerto Williams serves as an interesting case for examining toponyms. The city has been named Puerto Williams since the 1950's after Captain John Williams (or Juan Guillermos) who was influential in maintaining Chilean sovereignty. Prior to taking this name, the city was given the name Puerto Luisa by an Anglican missionary in honor of his daughter. Ricardo Rozzi of the Chile Sub-Antarctic Biocultural Conservation Program pointed out that these names ignore the prior history as well as the prior and current inhabitants (humans as well as local flora and fauna). The colonizing trend has been to displace the local in favor of the spatially and culturally removed toponyms.

The prior inhabitants of this area were the Yahgans, who often named places after bird or plant species that could be found in the areas they inhabited. The original name for this area perhaps was forgotten, but due to this known naming feature, and the fact that it there is a noticeable abundance of wild currant in the port vicinity, it was suggested to two of the last remaining speakers of the language at the time whether this area might once have been called Upushwaia—upush the name for the wild currant shrubs, and waia meaning bay. Rozzi reported that Ursula and Christiana Calderon, the two speakers in the discussion, agreed this might have been true, and the Yahgan community has since adopted the name. This case is interesting because we have a kind of reversal, a recapturing of the original way of relating to and inhabiting an area. This is not the universally adapted name, and the official name of the city remains Puerto Williams, but it is an interesting way of keeping the upush alive in the minds of those who hear the name, and perhaps can serve to

keep the species before the minds of the residents in a way that a name like Puerto
Williams cannot.

       Close to where I lived in Texas, there is a wonderful example of this in the town of
Flower Mound. I have several friends who live in the town of Flower Mound, and it has
been my experience that this name has generated more conversation than any other city
name in the area. The fact that this name generates questions and conversations about its
origins is further demonstrated by the following passage from local *D Magazine*:

> The first thing everyone wants to know about Flower Mound is the origin of its
> unusual name. "The Mound" really does exist, accompanied by much legend find
> lore, [sic] The most widely accepted explanation for the unusual 12 1/2-acre
> formation rising 50 feet above a prairie near am [sic] of the town's major
> intersections is that it was a sacred ceremonial ground of Wichita Indians in the
> early 1800s. The Mound is now a preserved historical site. (Michalski 1995)

       The name "Flower Mound" does, as *D Magazine* states, come from an actual flower
mound that exists in the town, which according to the town's website "rises 650 feet above
sea level, and it stands 50 feet above the surrounding countryside" (Town of Flower Mound
2013). The town, however, makes no reference to Wichita Indians in their account of the
name's origins; it simply states that there are many contradictory stories about the origins
of the mound before proceeding to outline the "unchallenged facts" about the mound
(which the authors admit are few). Referencing the Texas historian A.C. Greene, the site
contests the "mound" received its name in the 1840's because of the abundance of wild
flowers present on the mound; over 175 different species exist there according to a non-
profit organization devoted to the mound and its preservation and to ecological and
scientific outreach, called the Mound Foundation.

My interest is not so much in the accurate origins of the name, but in how the name functions for citizens of the town and surrounding areas. The name itself generates discussion of this mound, which is located beside a supermarket and other residential and retail developments—a piece of land that might otherwise go unnoticed as one drives through the town. The name itself calls attention to the mound and simultaneously to the native wildflowers that grow there. As a result, some interest has been made in this little spot that might otherwise have been leveled for more development, and it has been protected and serves as a site for environmental education and outreach.

I like this example of Flower Mound, because it functions in a similar way to the name Upushwaia in Cape Horn, but is perhaps more accessible on account of its geographical and cultural proximity. Furthermore, whereas Upushwaia sounds foreign and exotic, Flower Mound is actually quite mundane; as a result, it perhaps illustrates the point even better as to how toponyms can engender positive relationships and connections to the local environment.

If Flower Mound can focus attention on the mound of native wild flowers and assist in garnering attention on native wild flowers, I think this is beneficial for both the local environment and its inhabitants. Rozzi said of the name Upushwaia that such names are expressions of a "profound sense of living together with the plants and features of the landscape" (Rozzi et al 2008, 330). Flower Mound too engenders a sense of cohabitation with the flower mound itself. If the mound were razed to provide space for a new strip mall or a new supermarket, something important would be lost from the Town of Flower Mound. The name, rather than being a source of pride and a sign of cohabitation could

become a sore spot in the town's psyche by pointing to what was once there, and what is now lost.

I certainly don't want to overstate the power of toponyms. Surely many residents of Flower Mound have little to no interest in native wild flowers and would feel no psychic pain if the mound were razed to make way for a new development site and the toponym lost its reference. However, that does not mean that toponyms cannot be useful and powerful points of connection to the landscapes. Calling this town Flower Mound may well serve to inspire some of the citizens to interact with and care for this place that might otherwise go unnoticed. It can also serve as an opportunity for environmental education activities aimed at promoting a collaborative environmental identity.

Comparing non-western indigenous toponyms with American toponyms reveals a greater frequency of environmental referents present among the indigenous toponyms than American ones. I would go further to say these indigenous toponyms frequently indicate a closer relationship with and knowledge of the surrounding environment than is present in mainstream American communities. Any lessons to be learned from conducting such a study depend upon the presence of the indigenous toponyms as a source of comparison. In the course of my research I was alarmed at how frequently these toponyms vanish and are replaced with modern and often colonial toponyms. I would like to therefore devote the next part of this article to a brief treatment of the value of linguistic and cultural diversity and present some arguments in support of preservation of linguistic and cultural diversity.

Linguistic and Cultural Diversity

As mentioned earlier in the article, environmental problems are often framed as scientific/technological problems, as opposed to cultural problems. While greater attention has been given to the threats to biodiversity, the quality of our air, water, and food in the United States, less attention has been given to the similar dangers to linguistic or cultural diversity. I would like to devote some time here to elucidating why it is important to be concerned about linguistic and cultural loss and why we ought to be more proactive in promoting linguistic and cultural conservation.

There are two main reasons why we ought to be concerned about linguistic and cultural identity. The first has to do with the fact that languages and cultures are embedded with environmental knowledge. Different cultures have different ways of interacting with and inhabiting the natural world. Some of these ways are more sustainable than others. If we wish to learn more about how to inhabit our environments, we need these others with whom we may come into dialogue. Just as biological diversity can provide greater resilience in the face of environmental challenges and disruption, so too can cultural and linguistic diversity provide similar resilience. Focusing on the value of linguistic and cultural diversity for the purposes of resilience alone, however, can sound overly instrumental. The second reason has more to do with the people of these languages and cultures. We also ought to be concerned about linguistic and cultural diversity because in many cases indigenous communities desire to keep alive their customs and language, but find this a difficult tasks in the face of various pressures to assimilate to mainstream cultures and languages. We need to do more to support those who want to preserve their culture, language, and ways of life.

Regarding the importance of cultural and linguistic diversity, Michael Krauss opened up the eyes of many with his alarming forecast that, "at the rate things are going—the coming century will see either the death or the doom of 90% of mankind's languages" (Krauss 1992, 7). In a special report in 2007 titled "Language and Linguistics," the National Science Foundation estimated the impending loss at 50% (National Science Foundation 2007). Even this more conservative estimate is shockingly high. It is worth noting that this is twice the estimated loss of biological diversity as estimated by prominent scientists such as E.O. Wilson.

In a radio interview David Lightfoot, an assistant director at the NSF, said that experts estimate that on average, one language is lost every two weeks. The NSF has responded to this crisis by partnering with the National Endowment of the Humanities in creating the Documenting Endangered Languages (DEL)[4] program to fund research activities engaged in preserving endangered languages. This crisis however, and why it is important to protect linguistic diversity can be approached from a number of different ways. Below are three reasons taken from their special report why the NSF views the issue to be important:

- The enormous variety of these languages represents a vast, largely unmapped terrain on which linguists, cognitive scientists and philosophers can chart the full capabilities—and limits—of the human mind.
- Each endangered language embodies unique local knowledge of the cultures and natural systems in the region in which it is spoken.
- These languages are among our few sources of evidence for understanding human history. (National Science Foundation 2007)

---

[4] http://www.nsf.gov/funding/pgm_summ.jsp?pims_id=12816

What I am most interested in for this chapter is the second point. Each language and culture represents a unique perspective on living in and with the natural world. As we increasingly homogenize the planet on biological, cultural, and linguistic levels we are losing these different perspectives. These different perspectives are important because they can help us in the West in the creation of more sustainable environmental identities by presenting alternative modes in inhabiting and interacting with the natural environment.

Protecting Linguistic and Cultural Diversity

Luis Mizon (1997), who analyzed the colonial mode of appropriating the indigenous places around Cape Horn, Chile, quoted the woman said to be the last of the Wollaston islanders who said:

> I am the last survivor of the Wollaston islanders. There used to be five Yahgan tribes, each from a different place but all speaking the same language. Before I could walk, I had been as far as Cape Horn, strapped to my mother's back. Everybody called me Rosa, because that was how I was christened by the English missionaries, but my real name is Lakutia the kipa. Lakutia is the name of a bird and kipa means woman. All Yahgans are named after the place where they were born, and my mother gave birth to me near Lakutia Bay. That's how it is done among our people: we are given the name of the place that welcomes us to the world. (1997, 32-33)

Mizon argues that colonizers arrive at a place and treat it like a newborn baby waiting to be named, failing to acknowledge or respect that the place already has a name—naming, he claims, is an act of power, and is taken as a sign of possession. Prior to this colonial appropriation, Mizon claims that the names of this region of Chile "told the simple tale of people living in symbiosis with their environment," whereas now they "speak of the clash of two peoples and two cultures fighting over the same territory. Indigenous names and European names of coastal features and mountains are associated with different

memories. Today these names intermingle in a landscape into which the indigenous population wanted to melt while the Europeans simply wanted to take possession of it" (1997, 33).

As we saw with traditional naming however, it need not be an act of possession but rather of acknowledgment or cohabitation. However, a far more common trend is the movement not towards nature but away from it. This can be witnessed in the name change to Puerto Williams and other local toponymies; for example the channel separating Yahgan territory form neighboring Ona land was traditionally called Onashaga (Ona channel), similarly a channel within Yahgan land was named Yahgashaga (Yahgan Channel), those channels have respectively been renamed Beagle Channel in honor of the famous vessel carrying Charles Darwin, and Murray Channel.[5] When these names change the story that they tell changes as well. The former tell a story about the environment and its inhabitants, the latter tell the story about an expedition by an English scientist. I am not against celebrating the accomplishments made by Darwin; however I am concerned about what is lost when these indigenous toponyms are displaced.

By creating a new space, giving it a new name and new meaning, we can push away or at least overlook the endemic forms of life that had formed their own unique place. When we do so, we lose valuable forms of life, and in the process we exterminate or endanger cultures, we lose knowledge about that part of the world and how to inhabit it, this knowledge which can offer universalizable features of ways to be human in general and interacting with particular landscapes in particular.

---

[5] This history of naming in Chile was reported to me in conversations with Ricardo Rozzi, director of the Sub-Antarctic Biocultural Conservation Program.

When we lose indigenous toponyms that typically refer to local features of the surrounding environment or conspicuous species present in that area, this represents a significant, but difficult to quantify, loss of knowledge. I also suggest that this loss of a toponym affects our relationship to what was named. For instance, if upushwaia is the bay of wild currants, but then becomes Puerto Williams, the latter toponym does not serve to keep the wild currant present to us in our discourse.

Knowledge, culture, language, and the environment are all inextricably linked together. Do violence to one, and it threatens to destroy the others. Colonizer's toponymies often lead to anthropocentric relations with nature, and push out and do violence to the people and what had been indigenously present biologically and culturally.

As mentioned earlier, the loss of language and culture could be simultaneously be seen as a purely social issue that should be given attention regardless of environmental concerns. However, we have also seen how intimately connected language, culture, knowledge and place are, as well as how colonizing trends lead to the homogenization of biological, cultural, and linguistic diversity. The implication of this loss of diversity is a loss of knowledge—knowledge that is the result of human co-evolution with the land. This is a loss of knowledge not only about specific places but often a loss of traditional ecological knowledge. For the same reasons that it is dangerous to deplete the earth's biodiversity, it is also dangerous to deplete its cultural and linguistic diversity. These endangered languages and cultures offer different perspectives on the world and how to inhabit it, we need to incorporate as many of these different perspectives as we can to effectively confront our global environmental problems. This spirit can often be found in

indigenous toponymies, and this is why the recovery of the name upushwaia is refreshing,

Rozzi and his collaborators claim,

> By reincorporating a Yahgan name like Upushwaia, we recover the profound sense of living together with the plants and features of the landscape which are expressed by the indigenous language. By preserving an explicit reference to the bio-cultural diversity of the place, the Yahgan name helps to continue cultivating an indigenous environmental ethic that regards the place as belonging to the whole biotic community and not only to humans. (Rozzi et al 2008, 330)

<center>Toponyms as a Reflection of Environmental Identity</center>

Comparing toponyms from different areas shows that our toponyms reflect our relationship to the surrounding environment in both positive and negative ways. From the limited range of toponyms that I have examined, there appears to be trend among indigenous non-western communities of having toponyms that express a more healthy and collaborative relationship to the environment. In contrast, the Texas toponyms I looked at tended to displace what was given both biologically and culturally, and replace it with culturally and spatially remote toponyms. This can serve to disconnect those inhabitants from the surrounding environment, such that it is no longer present in their everyday discourse. When this happens, we may neglect the and overlook the environment with harmful consequences.

I am not advocating that we borrow foreign indigenous toponyms. For even if we wanted to, they would not serve the same function disconnected from their natural context. We can emulate that practice as described by Mizon and establish even mundane names like Flower Mound that fulfill the same function as a name like Upushwaia for us in our context. To come to this awareness, it is helpful to enter into dialogue with non-Western

communities. By doing so, we can learn different ways of expressing a connection to the natural world.

I do not want to be misunderstood as suggesting that we should eliminate anthropocentric toponyms in favor of non-anthropocentric ones. There would be a certain irony to that given that I express concern about language loss; eliminating historical toponyms would likely trouble historically minded individuals who would be concerned about the history being lost. However, I would like to see a recovery, or recreation of some toponyms with environmental referents, such that we can restore a connection with and knowledge of our surrounding environments. In and of themselves, these sorts of toponyms will not resolve our environmental problems. They are just one small way that we can foster a greater awareness of our environments in our everyday discourse. However it is by focusing on these small, but significant reflections of our relationship to the environment that we may eventually cultivate a healthier and more sustainable relationship.

References

Abram, David. 1996. *The Spell of the Sensuous*. New York: Vintage Books.

Barry, Dan. 2005. "About New York; No Way To Name An Avenue." *New York Times*, September 21. Accessed March 18, 2013. http://select.nytimes.com/gst/abstract.html?res=F10F10F73D540C728EDDA00894 DD404482&smid=pl-share.

Flower Mound Texas. 2013. "The Flower Mound." Accessed March 18, 2013. http://www.flower-mound.com/index.aspx?nid=104.

Hargrove, Eugene. 1989. *Foundations of Environmental Ethics*. Englewood Cliffs, NJ: Prentice Hall.

Husserl, Edmund. 1970. *The Crisis of the European Sciences and Transcendental Phenomenology*. Translated by David Carr. Originally published in 1954. Evanston, IL: Northwestern University Press.

Krauss, M. 1992. "The World's Languages in Crisis." *Language* 68 (1): 4-10.

Michalski, Dan. 1995. "The Best Places to Live." *D Magazine*, November. Accessed March 18, 2013. http://www.dmagazine.com/Home/1995/11/01/THE_BEST_PLACES_TO_LIVE.aspx.

National Science Foundation. 2007. *Language and Linguistics*. http://www.nsf.gov/news/special_reports/linguistics/index.jsp

Roberts, Sam. 2012. "A History of New York in 50 Objects." *New York Times*, September 2. Accessed March 18, 2013. http://www.nytimes.com/interactive/2012/09/02/nyregion/a-history-of-new-york-in-50-objects.html?utm_source=twitterfeed&utm_medium=twitter&_r=0#/?gridItem=all.

Rozzi, Ricardo, et al. 2008. "Field Environmental Philosophy and Biocultural Conservation: The Omora Ethnobotanical Park Education Program." *Environmental Ethics* 30: 325-36.

Sapporo. 2013. "About Sapporo." Accessed March 18, 2013. http://www.welcome.city.sapporo.jp/english/history.

Shepard, Paul. 1999. *Encounters with Nature*. Washington, D.C.: Island Press.

Multi-Language Implementation Framework
for Legacy and New Systems

Anderson Passos

ABSTRACT

Globalization is in the mind of business owners and managers all over the world. From small family-based companies in the suburb to big corporations, everyone knows that the world (for now) is the limit. For these people, using an IT system that speaks their language is something taken for granted. This paper presents a multi-language implementation framework that can be used in new as well as legacy systems. For new systems, it can be an integral part of the application's backbone while for legacy systems it can be used as a totally independent translation system. We will go through some implementation aspects of our proposed framework without going too deep into technical details, allowing the reader to understand the concept without any technical background.

Keywords: *language, system development, information technology, translation*

Introduction

Around 10 to 15 years ago, companies were struggling to come up with solutions for e-commerce portals. The word Globalization was in everyone's mind and the path to it, at least from a system's design point of view, was to provide as many languages as possible to users.

Multiple languages or multiple interface languages are not only overseen by Internet users but also by business owners. It is easy to guess why a Japanese company will not implement a software solution if such has no support for Japanese language for example. Some companies like Rakuten, a Japanese Internet shopping mall operator, decided to hold all internal meetings in English in an effort to become a more global company. We can assume

Anderson Passos teaches at Miyazaki International College. Correspondence may be sent to: MIC, 1405 Kano, Kiyotake-cho, Miyazaki-shi, Japan 889-1605, Tel: 0985-85-5931, Fax: 0985-84-3396, Email: apassos@sky.miyazaki-mic.ac.jp

that, for Rakuten, it would not be a problem to implement internal software in which the main menus and functions are written in English, but this is not feasible for every company.

Purpose of this work

The work introduced here tries not only to address the issue on providing multiple languages in the same interface for business applications, but also suggests a framework to be used when implementing such systems. In an attempt to keep the contents of this work reachable for people without any technical background, pseudo-code will be used instead of actual programming language code.

Learning through experiencing it

From this author's personal experience, most projects developed in the period of 1998 to 2006 had one thing in common: They were all trying to keep content separated from functionality. This is somehow understandable since designers and programmers were, and still are, usually not very friendly to each other and once a problem in the project is detected program managers always experience flames flying around.

In 2002, Microsoft released a new programing language/framework called .NET (pronounced dot net) and one characteristic very welcomed by the developing community was the code-behind model, in which the content provided by the designer and the functionality provided by the programmer sit in two (or more) different files. Most programing languages provide such mechanism direct or indirectly, and the code necessary to deal with multiple-languages can be put into one of those files providing functionality to the program.

Proposed Method

It is important to notice that we will be using a database system as a backbone to implement our framework. The first thing to define is the structure of our system tables; We will need a table to hold the available languages in the system (**tblSystemLanguage**), a second table that will hold the keys to be replaced in the program's interface (**tblTextKey**), and a third table that will connect both previous tables together (**tblSystemText**). Figure 1 shows our suggestion for such table layout.

The idea is that, once the program asks for a given text key (**textKey**) in a given language (**languageID**), our program is able to retrieve such information directly from **tblSystemText** if the pair language/key (**languageID**, **textKey**) is provided.

The second thing is to define a common function that will be called every time we need to replace a label in our application. Depending on the programing language chosen this can be a relatively easy task with a simple function for more procedural ones or classes for object oriented ones.
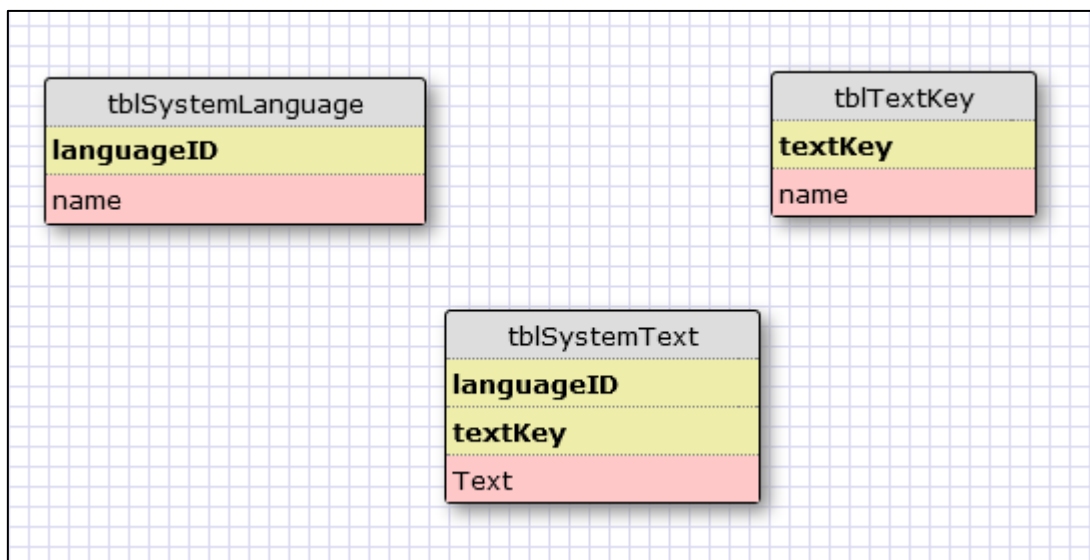
Figure 1. Tables necessary for the proposed implementation

It is important to mention that the **textKey** is not necessarily in a readable/understandable format. The idea of a **textKey** is that it would serve as a placeholder inside the text of a program or web page, making it possible to replace as many occurrences of the same **textKey** as possible. For example, in Figure 2 we have a login screen showing only the **textKey** in the placeholders.
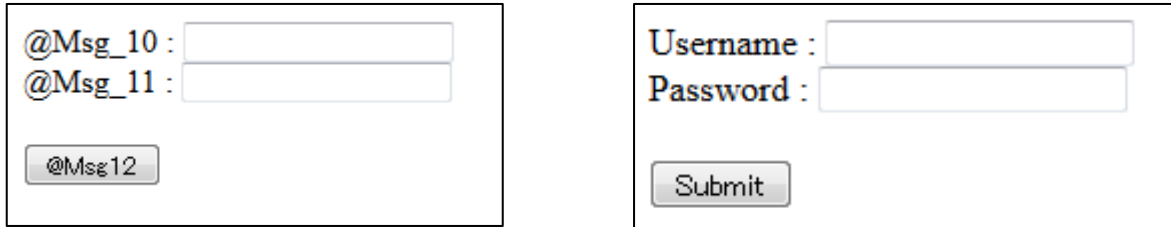
Figure 2. A login page showing only the placeholders for textKeys (left) and the translated text (right)

Our wonder function: getText

Many programing languages provide a function with the same name. It is not our intention to lead the reader into confusion here. As in the PHP scripting language, **getText** makes use of resources sometimes out of reach from some programmers. Things like operational system resources or even making use of definitions that are loaded in memory (all at once, what require a lot of system resources) are common. It does not need to be a genius to realize that more memory you use, slower your application will perform. Furthermore, the use of text files and operational system resources is heavily discouraged.

Our **getText** function makes use of a database connection to retrieve a text string based on its key and the language. Figure 3 shows the pseudo-code for our function. Note that the first line inside the function says to use an opened database connection. This is a small technical detail that can determine the performance of an application.

```
function getText(lang, Key){

        use an open database connection

        get the text string FROM tblSystemText

                            WHERE Key = textKey AND

                            lang = languageID

        return the string found

}
```

Figure 3. Pseudo-code for the getText function

Opening and closing a connection for each function call is unnecessary and resource intensive, resulting in slow performance. To avoid such bottle neck, application programmers usually open one database connection and reuse it as many times as possible inside their functions. Of course, implementation depends on the programing language used;

Updating text in an application

Once the tables and function are in place, the process of updating text strings in a program becomes a trivial thing. It is natural to assume that when paying for a development company to develop new software, a new interface to manage the languages and the text strings will be needed. Such cost can be compensated when the organization asking for the development (the client) has no more to request updates and/or fixes in the menus or labels. Of course it is a case-by-case scenario and the reader should evaluate which one is more beneficial.

The updates can be done in two different ways. The first possibility is to update all text strings in a batch mode as shown in Figure 4.

msgid	English	日本語
@Msg_10	This has already been assigned a categor	カテゴリに割当済みのため、削除できません
@Msg_100	Office hours have already been set for th	その曜日、時間には既にOfficeHourが設定
@Msg_101	End time must be later than start time.	終了時間は開始時間より後でなければなりま
@Msg_102	An advisor assignment function error has	アドバイザー割当機能でエラーが発生しました
@Msg_103		選択したアドバイザーとリーダーは同一人物で
@Msg_104	The data object has already been updatec	対象データは既に更新されています。再度処
@Msg_105	An error occurred while processing upda	更新処理中にエラーが発生しました。

Figure 4. One possible output for the administrative interface. Translation can easily be done in batches.

In case an application has too many languages, the above layout can prove itself difficult to visualize. As a suggestion, text can be updated in a one-by-one basis, always showing to the user what key he is updating and for which language like in Figure 5.

Figure 5. Different languages can be edited one by one and still use the proposed framework. Tabs with the different language allow the user to edit a given text in the corresponding language.

Additional functions

Clearly, additional functions are needed in order to manage the languages, texts and keys in the database. Table 1 lists the functions implemented during our prototype testing. Additional functions can be created to add new languages and keys as required.

Figure 6 shows the pseudo code for an optional function called **getDefaultLanguage**. This function could be called every time **getText** is unable to find a translated text. A workflow exemplifying how this can be achieved is shown in Figure 7.

Table 1. Minimum required functions to implement

Function	Parameters	Explanation
getDefaultLanguage	n/a	Returns the default languageID. Useful when getText returns no value for a requested language.
updateText	languageID, textKey, Text	Updates the text for the tuple textKey, Text

```
function getDefaultLanguage(){

    return languageID for English

}
```

Figure 6. Pseudo-code for the getDefaultLanguage function in case English is defines as the default system language
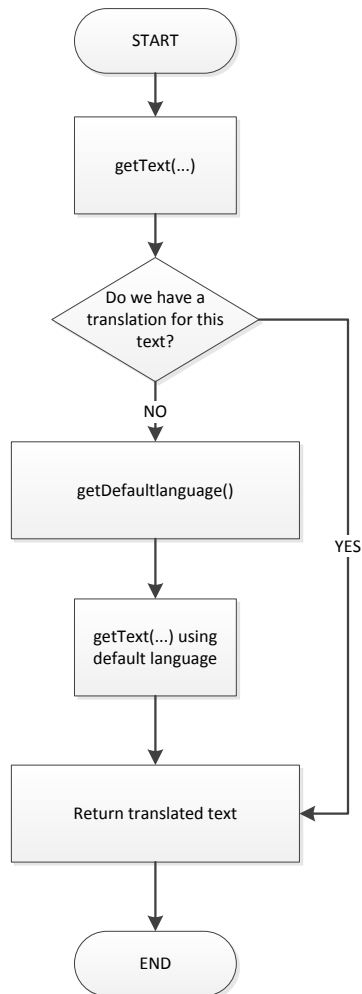
Figure 7. Optionally, the system can return a text in its default language in case a translated one is not found

Dealing with Legacy Systems

Legacy systems will, inevitably, require changes in the source code to be made. Depending on the programing language used and the programmer's skill modules can be created and minor impact will be felt by end users. For illustration purposes only, an updated HTML code is shown in Figure 7. This is a very simplistic case where text that was previously displayed in the page from hard-coded strings is now loaded from a database using our suggested framework.

Figure 7. Old HTML code (left) and updated PHP script using the framework proposed in this study

Final Considerations

Still today, many software developers rely on setting an application interface language to the user's system settings even knowing how cumbersome it is. GIMP, an open source image manipulation tool had it changed only in version 2.7 (not yet released at the time of this writing). This suggests that not only implementation of background methodologies on how to change language of a program in run-time has not evolved but also that such feature has been neglected until today.

The work showed here brings to the table an easy to implement and to understand framework for developing Multilanguage systems. New systems can have the set of tables and functions included in its blueprint while legacy systems will require a little bit more work. Such

work can be compensated later with the relative cut in time necessary to update a website or

program language strings. For the final user, having a system that he can control by himself can

end up cutting maintenance costs down the road.

References

Minoru Matsutani (2010). "Rakuten to hold all formal internal meetings in English". *The Japan Times Online, May 18, 2010.*

Richard Anderson et. Al. (2002). "Professional ASP.NET 1.0", Wrox Press

PHP Manual (2012). *Control Statements,* Retrieved from http://php.net December 2012.

PHP Manual (2012). *Gettext functons,* Retrieved from http://php.net December 2012.

GNU Image Manipulation Program, *GIMP 2.7 Release Notes*, Retrieved from www.gimp.org/release-notes/gimp-2.7.html December 2012

The Pedagogical Benefits of CRS use in Small-Group and Culturally-specific Settings:
A Pilot use and Study of Student Perceptions

Scot Hamilton
Anderson Passos

ABSTRACT

Electronic classroom response systems (or CRS) are tools used to engage students in material; promote discussion; and offer teachers an immediate way to measure lesson and learning effectiveness. Their benefits are well understood for large class sizes and lectures. However, we have been interested to study how they might perform in small classroom settings and to assist educators and students to overcome social and pedagogical constraints peculiar to Japan and ESL contexts. It's mainly been our wish to find ways to circumvent a certain reticence on the part of students to participate more dynamically in classroom discussions and to offer feedback as to their comprehension of concepts. In this paper we begin by describing the general rationales for using CRS and some description of practices employed. This includes some anecdotal reflections. We then describe our very simple (and small sample size) survey results on student perceptions of the system(s). This generally revealed CRS to be positively viewed on self-assessment of performance and engagement; and as a continuous assessment tool (and welcome alternative to traditional mid-term and final exam formulae). Lastly, the authors discuss CRS juxtaposed with traditional and new constraints in teaching and urge educators to reckon with the cognitive and wider culturally affected and adapted attributes of today's students.

Keywords: *Computer-supported collaborative learning, education support, Clicker, enhancing teaching and learning*

Introduction

Classroom response systems (CRS), also referred to as "audience response systems," or "clickers," have become increasingly popular in American Universities (Bruff, 2009; Hoekstra, 2008), and also more recently in Asia (Atkinson, McBeath, Soong, & Cheers, 2007). On first glance, these systems appear 'high-tech,' but they are relatively simple in components, easy to

use, and relatively inexpensive to buy. In terms of fixtures, they tend to complement existing classroom facilities, such as a projector and a teacher's in-class computer, rather than requiring any new or elaborate wiring or infrastructure.

A typical CRS includes a teacher base station (with a small antenna and USB port/connector) and each student is equipped with a wireless handheld device similar to a television remote control (but more basic in fact). The systems' components communicate by radio frequency and each handheld device is pre-registered and recognized by serial number/bar code to a student (who is identified as using that remote, each time, and is designated to it by student name or student ID number).

In large courses (with enrollment over 100, for example), the commercial distributors of these systems usually provide the professor with a base station and several spare remotes free of charge. This package includes the modest software required for organizing and collecting student polling data and for producing graded (cumulative session-by-session) performance reports and output. The polling sessions are generally conducted over or upon simple Power Point slide presentations of questions. Multiple-choice is usually the preferred method for these prompt items or questions and their 'sessions' - though second generation clicker devices do offer the opportunity for students to respond by giving short, text-based, responses as an alternative.

CRS practicalities, use-viability, and cultural placement

In the United States, and in the larger colleges, it's common for virtually all students to own their 'clicker' control device (purchased during their first year and costing in the region around

30-50 US dollars per unit). And these devices are routinely sold in college bookstores, often bundled with a prescribed textbook. It is now certainly the case that the vast majority of American students today will have used, or will now be likely to use CRS at some point during their college careers. Students may use their devices for different courses, and if the entire equipment set is controlled and/or owned by the instructor, then devices can also be shared between sections of the same small-enrollment course. This is how it occurs at our own institution where the devices are simply handed out each day of class when 'clickers' are scheduled or have been decided to be used.

In short, CRS are ubiquitous (in some parts of the world); they're flexible; simple to use, and part of a simple set-up. And they are affordable.

There is also at this time a growing body of research that has critically investigated the use of these systems (e.g., Sprague & Dahl, 2009; Lucas, 2009; Duncan, 2006). This literature provides useful insight into their effectiveness (or less-effective use) for certain populations and contexts, certain new ideas for employment, and information about how various constraints concerning implementation have been handled elsewhere.

One practical and economical question for us, for example (should we wish to making CRS more of a feature at our institution), is whether to provide all incoming students with a device, or to place the burden on the student him or herself to become equipped as a matter of course. If the faculty, however, are unfamiliar with the systems, or resistant to the perceived learning curve regarding the technology, or simply prefer their own techniques, then CRS course offerings will be few. Then it may be sensible to avoid blanket expenditure in either direction (by the institution or by the student intake).

CRS, class sizes, and methods of polling options

Miyazaki International College (MIC) has a policy limiting class-sizes to 25 students. We believe that, with small classes, professors and students not only get to know each other better, but these modest class sizes have the potential to foster conditions that are conducive to student-centered education, active learning, and the development of a critical thinking. These are elements that are integral to the stated institutional missions and our particular approach to pedagogy.

So we wondered how such an educational tool, albeit one with known use-value in other cultural contexts and larger class sizes (see Mayer et al, 2009) - could be applied to smaller groups, here in southern Kyushu, Japan, and aligned with these teaching and learning objectives.

An immediate impression was that there is considerable flexibility inherent to these systems. For example, teachers may instantiate a question for responding immediately after a piece of lecture material or a demonstration has taken place – this in order to ascertain how well the teacher has communicated the concepts or argument; and/or, at the same time, how well the class has understood the material (with each proposition here, largely following the other, or being interchangeable, of course).

Post-polling bar chart information is made instantly available and projected overhead. This offers visual feedback (for all), a pause for further teaching or explanation, and (quite often) a few moments for (moderately stunned or quietly satisfied) critical engagement among

the students themselves – as they not only get a chance to see the right answer but also witness the break-down of the class per answer available (usually an A through E of options).

Figure 1 – Prof. Hamilton during a CRS session

Alternatively, as a method for incorporating clicker, it is an option to pose perhaps 3-5 questions at the end of each and every class (and for this regimen to be understood and pre-advised in the syllabus). Again, with this practice, both students and teacher(s) are able to gain immediate information not merely towards the correctness or non-correctness of responses, but also towards arriving at some explanation as to why a given response may have been correct or otherwise (or indeed why so many or so few students seemed to opt for one choice or another, viewable in blocks on the bar chart). Similarly, follow-up explanations can be then given by the teacher or offered-up by one or several of the students who may have opted correctly (and who are willing to share their logic or basis for making their responses among choices available).

Perhaps one specific benefit to this method choice is that it allows for more uninterrupted or less truncated lessons and lectures. To some degree it also induces students to take notes and devote their attention to the types of things that may be quizzed for at the end of the period. However, the instructor's decision to adopt this might be influenced by the particular teaching subject matter involved (for example, technology and artificial intelligence, or abnormal psychology, or chemistry and sciences), as these may lend themselves better to one type of CRS incorporation or another. Moreover, with longer classes it is not only that 'clicker' might offer useful enlivening changes in tempo that can help sustain students' attention and engagement, the strategic placement of questions – and perhaps hybrid placement approaches – can contribute in gaining retention of the more important class material, themes and concepts.

Yet another method, and our most commonly employed here, is to designate an entire period for clicker questions (usually 8-10 questions, every 4th class or so) and using the opportunity to teach and coach pre-polling while doing it – generally by defining concepts in question slides, reminding students about arguments and theories from previous class notes and allowing them to think about and consult texts. This is truly a hybridized method, though in another fashion. Not quite a test and not quite a lesson either, these sessions do proffer sustained attention for a whole class period as the prompts appear sequentially (and students accrue grade points, and want to do well). They also cap off course modules to some degree, and they simply give us all a break from the normal class routines. In fact, we also change room location to do these sessions, just as a matter of choice and convenience (we go to where a instructor's laptop terminal happens to be already installed; the excursion adds just a little formality and change).

On several later occasions, in the more advanced weeks of courses, we chose to remove any nominal teaching and coaching over clicker questions (as sessions were happening). And for these special clicker sessions the overall emphasis (at instructor side) was placed upon students' **reading and understanding of ideas and arguments from within quite difficult texts** (for example, on the subject of the psychological benefits of adult human play). For this, the formula for quiz taking was altered significantly (with some interesting effect).
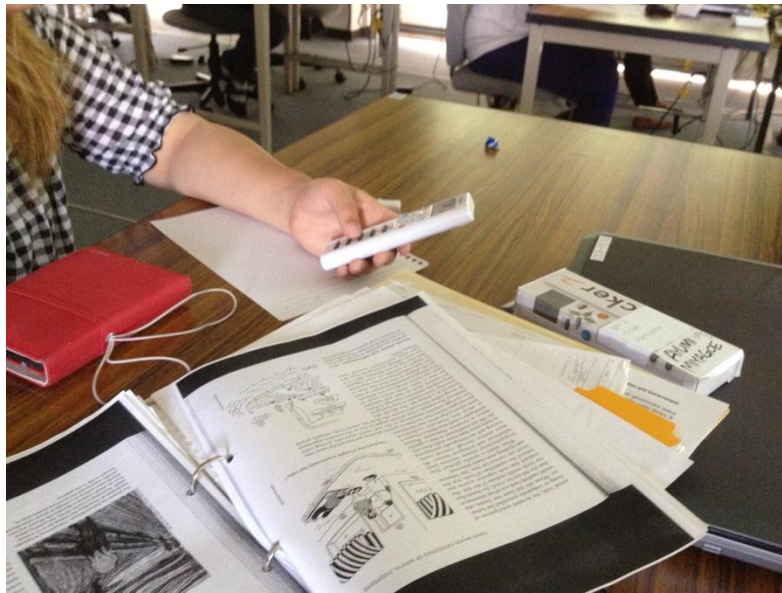
Figure 2 – Student reviewing class material before answering a clicker question

The task, then, for the clicker method was to induce students to make all (their best) intellectual efforts towards comprehending the nuances of designed (in contrast to merely posed) questions. And to read material before answering. The questions or prompts were what one would describe as the more "tricky" styles of multiple choice technique – designed, not for testing for the one (and the only one) correct option among choices. Instead, students were advised to look for, and choose, the best possible choice among alternatives that were (manifestly) somewhat correct or at least relevant to the question at issue. It was felt that this

more calibrated style of question and alternatives could elicit a discerning attention not so much to the details of question and supporting text (however this too, if necessary), but by default or necessity foster the type of mindset (albeit temporary and task-specific) that we might equate with, or equate as, deeper conceptual learning and critical thinking.

The set-up was as follows: students were working on the day with a targeted text (i.e., numbered paragraphs corresponding to clicker question numbers, e.g., 12 questions wrought-over a 12 paragraph sequential argument of a class material piece). Students could consult their personal electronic dictionaries, and of our specialist vocabularies, and importantly, they could also consult and work alongside another person.

This was the idea to conduct CRS in **student pairs** – with 2 students sharing one clicker/remote and responding jointly, giving one answer, with that one pair-assigned device. Ample time was allotted after question presentation before polling is switched on. This allowed each person within all pairs (by now, "teams,") to read independently and confer, and settle on an answer that the pairing felt to be correct or felt to be the best among choices.

In our particular context this (pairing and team) technique proved to be popular but (by our reflection) symbiotic in some culturally specific (and mini-culturally institutional) ways. Firstly, from a molar or general theoretical perspective, the situation was suspected to engender some beneficial social facilitation effect(s). Namely, the mere presence of 'competition,' as it were – in respect to the presence of the other working persons and 'team' pairs – would or could prime for some stimulus effect (on focused attention, arousal, performance, and so on).

But there is also the motivation factor of the notion of obligation as this might pertain to, or impact upon, a partner. We are a very modest sized private college, set in rural southern

Japan, and the students share many classes with each other. They are good friends, may travel to college, study, and may also work together in part time jobs. It is not uncommon for them also to have attended high school and middle school together. They are respectful young people, and of course, cultured in a society in which respect (and pride in respect) plays an important cultural role. So perhaps the diligence and sense of obligation observed between team members (in this small teaching and testing experiment) was different from what one might see, or expect to witness and exist elsewhere – for example, in larger institutional settings, larger cities, or in other regions and cultures.

Furthermore, in the relative absence of other (significant levels of) intra and extra-mural opportunities for competition, then the friendly "team clicker" – with competition itself talked-up to some minor extent by foreigner-instructors seeking to deflect at least some of the ambiance of serious test taking away (leaving intact the advantages of clicker as a teaching and learning rather than purely evaluation and testing tool), then the social production of this competition factor may have offered a learning and experiential bonus.

So the team clicker allowed for a momentary or rare competitive and inter-group team-challenge spirit to take place (for all). The way the students appeared to tackle the reading, intra-team conference, and response-giving was intriguing and impressive. And we also generally had fun with this format. Though purely anecdotal at this point, the overall accurate response performance for the class as a whole was good (again, immediately visible in the polling results/bar charts on projector). The adage that "two heads are better than one" may simply apply, of course. However, very difficult reading was assigned here, and with little or no teaching around it. The observed performance gave the hint of synergy, that possibility that the intra-team conference (after independent reading) may have produced something new. That

is, there is at least the possibly of some level of emergent understanding having taken place. This understanding or 'learning' may not be reducible (always, on occasion) to the merely increased chance of a correct answer residing *somewhere* in two (rather than one) respondents. Certainly, this may have to be measured more carefully within a systematic study (focused on CRS performance outcomes). However, this brief report here, its reminded, is set around a pilot use of CRS and casual variation and experimentation of techniques; the gaining of a modicum of students' perceptions; and offering of some reflected opinion and observation on our own use of the systems (within these cultural and institutional settings). Lastly, it should be noted that sessions of this type (the team sessions) were few in number. Therefore, our ideas might also be refined by sustained practice of CRS within this formula, or through more regularly scheduled occurrences of the technique.

Asking the students what they thought (an explanation of the questions only)

Our small data sample and the descriptive charts (of results) are derived from 36 third year students taking a brief questionnaire on the final day of class. Table 1 lists the five questions presented in our brief survey.

Table 1 - Questions used in the student survey

Q1	The interactive clicker system is useful for assessment and grading.
Q2	Clicker use - with instant feedback - is useful for making actual lessons, for learning and for the discussion/presentation of class content itself.
Q3	I prefer team clicker (allowing me to discuss questions with my partner) rather than responding on my own.
Q4	At the moment, we have approximately one clicker quiz every 4th class. However, some instructors prefer to use clicker during each and every class. For example, they might conduct a short 2-3 question session during, or at the end of each period. Do you think this would be more useful to you than having a (less-often occurring) larger clicker

	session covering larger areas of the course?
Q5	In this class, you were given less readings and more whiteboard explanations of the important themes and concepts (for you to take notes from). Would you have preferred more readings assigned, and for clicker to be based on these (take-home) readings? So that you could prepare for clicker in this way?

Additionally, students had a section on their survey where they could comment and/or give some suggestions more freely in an unstructured way. This merely asked them to comment about their class and use of CRS ("Do you have any further comments about the class and clicker-use?"). Not many further comments were given but what was given proved helpful for our interpretations of the more structured responding.

Question 1 (Q1)

Q1 relates to, or implies the contrast between CRS and the evaluation system customarily used in our institution, and the formula for most colleges and universities. Students typically receive a mid-term – partly providing some evaluation indicating (to student and professor) the level of performance at the half-way points. It also conveniently breaks-up the course, and course content to a degree. The final exam is, of course, administered at courses.' However, CRS use took place throughout the course and students accrued points towards their grade continuously and virtually from the beginning. The CRS component was set to account for some 60 per cent of the overall final grade. CRS points were also set to encourage attendance (with simple participation gaining points). It's projected that even bad performance on a given day (and while points were nevertheless available) nevertheless laid the basis for some learning. After-the-fact discussions and explanations are usually given or take place, and non-attendance was essentially a complete loss (in comparison).

Question 2 (Q2)

Q2 asks students how well they perceived the use of clickers to promote discussion in the classroom; about the delivery of class content; and the effectiveness of CRS (in students' own perceptions) of their learning experience.

Question 3 (Q3)

Some sessions were conducted in pairs and these were called "team clicker" sessions. As discussed at length in body of this paper/report, the main objective of these sessions was to elicit and socially 'produce' the understanding of critical themes and arguments (through reading and student conference) over difficult technical and theoretical material. Students were given an opportunity, read, to discuss thoughts with a partner (from a passage of highlighted reading), come to a consensus, and submit one answer for their "team." The prompt questions themselves were designed as conforming to best answer as opposed to correct answer formula (see main text here for fuller explanation).

Question 4 (Q4)

We opted to add a question here that concerned the possible change to a different approach towards the frequency for clicker to be utilized. Our clicker sessions (set at 1 every 4th class) took up the entire class period, and the instructor(s) felt that this method had some advantages and utility (again, please consult text here for greater reflection on this).

Question 5 (Q5)

As an experiment, and on occasion, students were given a section of material to read – in class itself – and the clicker session was based on that material (usually for team clicker). Such sessions were different from the rest of the course where whiteboard teaching was conducted

with students taking notes (while provided with reading from which the content was based, or as background). We have found it difficult for students to undertake significant pre-class reading, and for them to be able to understand many of the highly technical concepts (particularly in Western-derived and often medico-scientific psychology course content). Embracing such material without intensive hands-on assistance and explanation is very difficult, but students generally comprehend the content well with "on-the-fly" explanation. So in Q5, we tried to confront both possible clicker styles and to see which method students are, or may have thought to be, more comfortable with.

What the students did in fact think

In this section we are going to illustrate the results (from the questions advised in Table 1) and offer some simple interpretations.
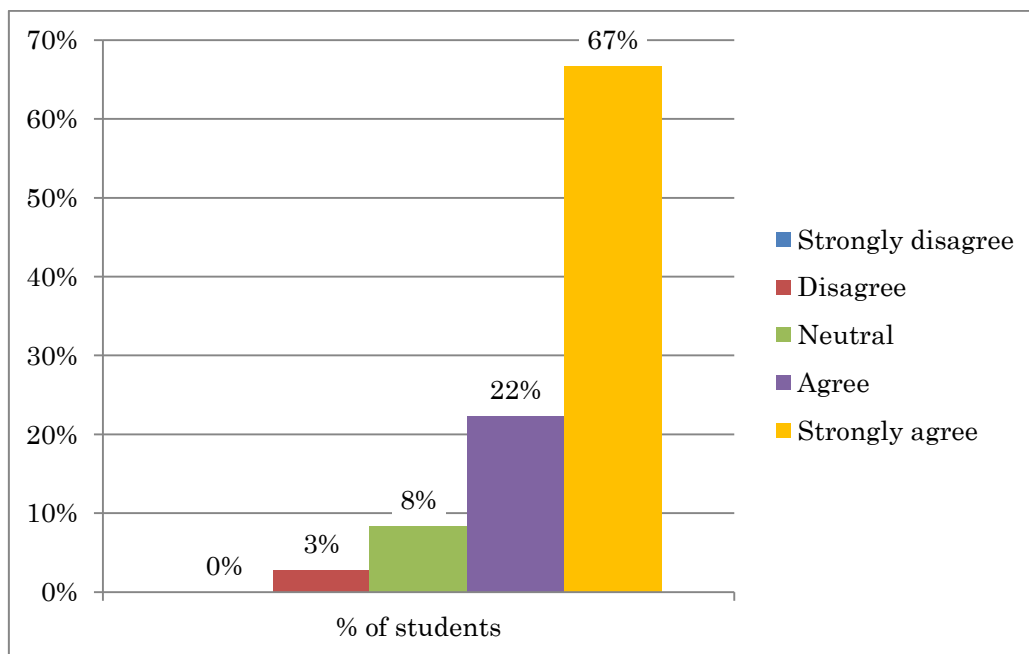

Figure 3- Summary of students' answers for Q1.

As expected, students' answers on question 1 (Figure 3) show that the majority of students agree with the benefits of (and find favor with) the continual assessment offered through CRS. This was somewhat expected after hearing, (and hearing of students' casually offered) remarks and opinions over the duration of CRS use in the courses (2 sections, of a singular course).

When asked about the effectiveness of CRS (Figure 4) as it pertained to the perception of substantive learning and discussion, only a few (6%) of our student sample did not agree that CRS was effective on the point of these central classroom objectives. Unfortunately these respondents did not provide much in the way of additional comments that would perhaps clarify their position. So it may be difficult for us to address this adequately. But it is not unimportant (even though the opinions in this direction were somewhat minimal). Perhaps we would, in a future questionnaire, however, ask for more information immediately below the Likert scale question prompt, rather than at the end of the whole survey.
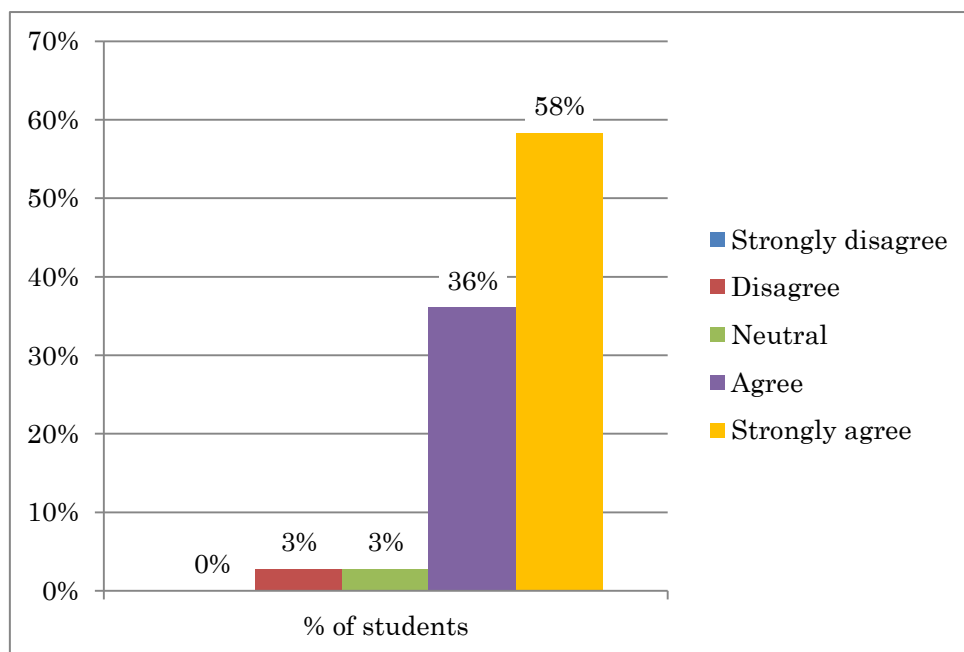

Figure 4 - Summary of students' answers for Q2

The results gained from question 3 (Figure 5, see below) are, more or less, according to expectations. We assumed that while some students would like to work in groups, others may prefer to work alone. And of course, some students in all classes prefer to sit alone.

Additional comments provided by some students indicated that for some, "team clicker" was good because they could consider and understand their partner's opinion. However, other comment providers indicated that being able to answer autonomously (with own assigned device) and later listening to teacher's explanation was more effective and preferred by them. Again in future surveys, it may be instructive to ask whether students felt any pressure to agree with a response that a partner and perhaps good had arrived at.

As discussed elsewhere in this paper, our students are a very cohesive community at Miyazaki. And therefore, it's an interesting question whether in this team condition, and in this micro-cultural context of friends and familiarity (and among young Japanese persons), interpersonal closeness either helps or hinders the assertive resolve to persuade a partner on the matter of a particular choice (based on a strongly felt correctness). Anecdotally for us, the interactions appeared quiet-though-involved and (nebulously) 'impressive,' as was simple overall performance over difficult material. But just how this plays out between paired partners here (in instances of deferring and asserting and in wrong or correct responding, perhaps) remains a fascinating social and cultural and psychological question (in itself).
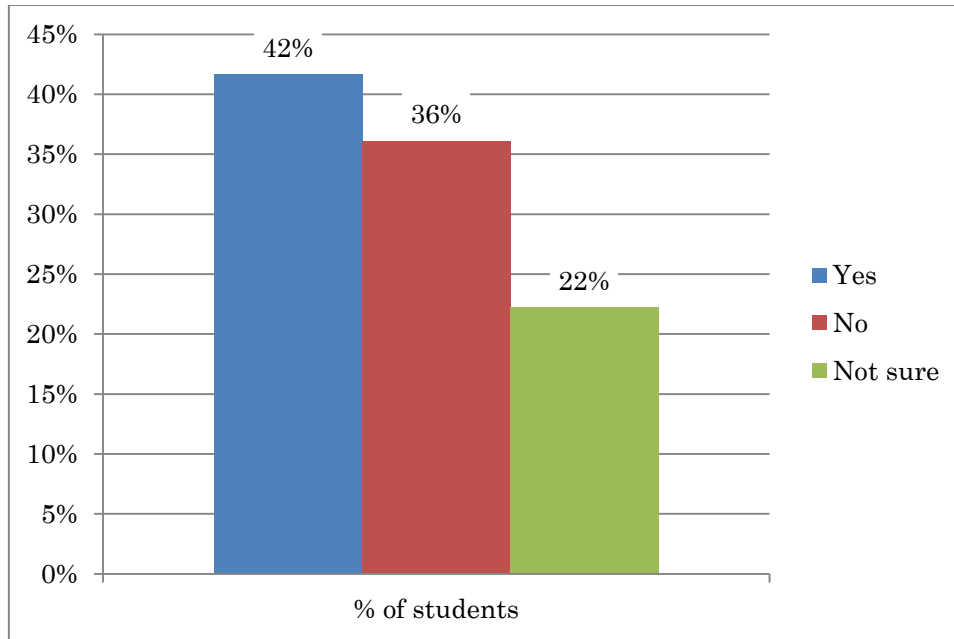
Figure 5 - Summary of students' answers for Q3

This may be overall difficult to interpret, however. Notwithstanding the general ambiance of fun and engagement of the pair/team sessions - and the fact that good, involved reading was appearing to take place (that may have not have been accomplished, or so effectively at home, otherwise) - it is obviously the case again that certain social psychological parameters may have been at play. For example, social facilitation [within and between pairs] as an enabler for increased concentration and as a spur to concentrated focus; but also again, a certain level of social desirability or conformity in acquiescing towards a given response that one may feel not to be what one wished to give, or feel to be correct.
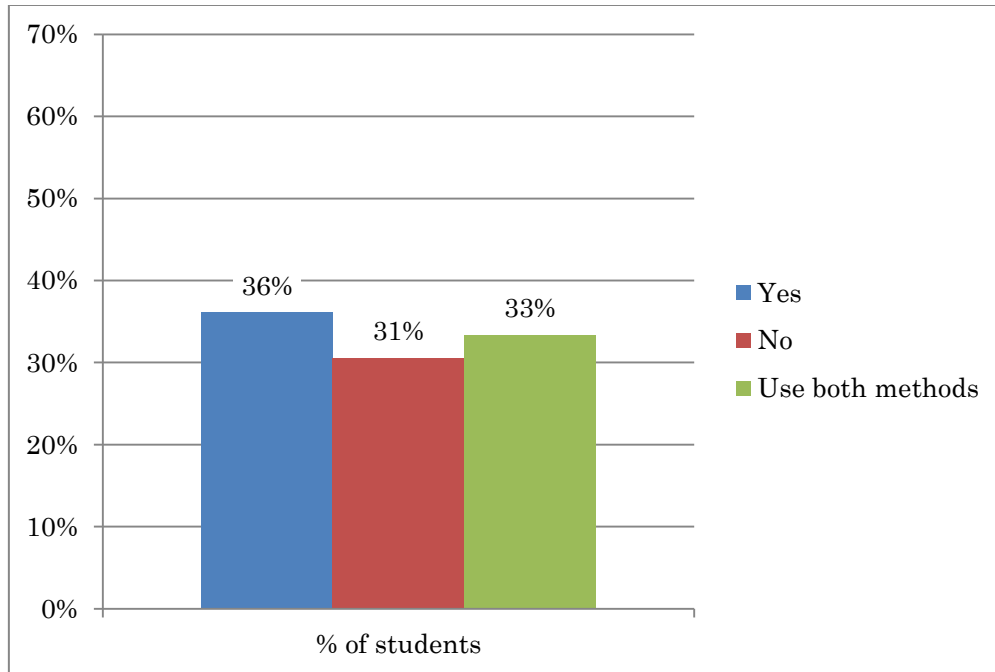
Figure 6 - Summary of students' answers for Q4

In question 4, the results were close to each other (Figure 6). With around 1/3 of our student sample preferring the larger CRS sessions (that were customary) this ran counter to our expectations to some degree. We expected students to prefer to be evaluated over small parts of content, and indeed within each and every class period. But of course, the projective use-methods were just that (projective, or hypothetical). However, the result indicates that these formats (or other formats) are viable at least in the minds of the students, who now are somewhat practiced and familiar with clicker as a technique and teaching/learning technology. And for us these tools and this result open up a space for several user-approaches or for certain hybridized methods of employment. These ideas (of mixing it up, and experimenting, and using the full range of clicker flexibility) are clearly encouraged by this virtual three-way split in survey data here.
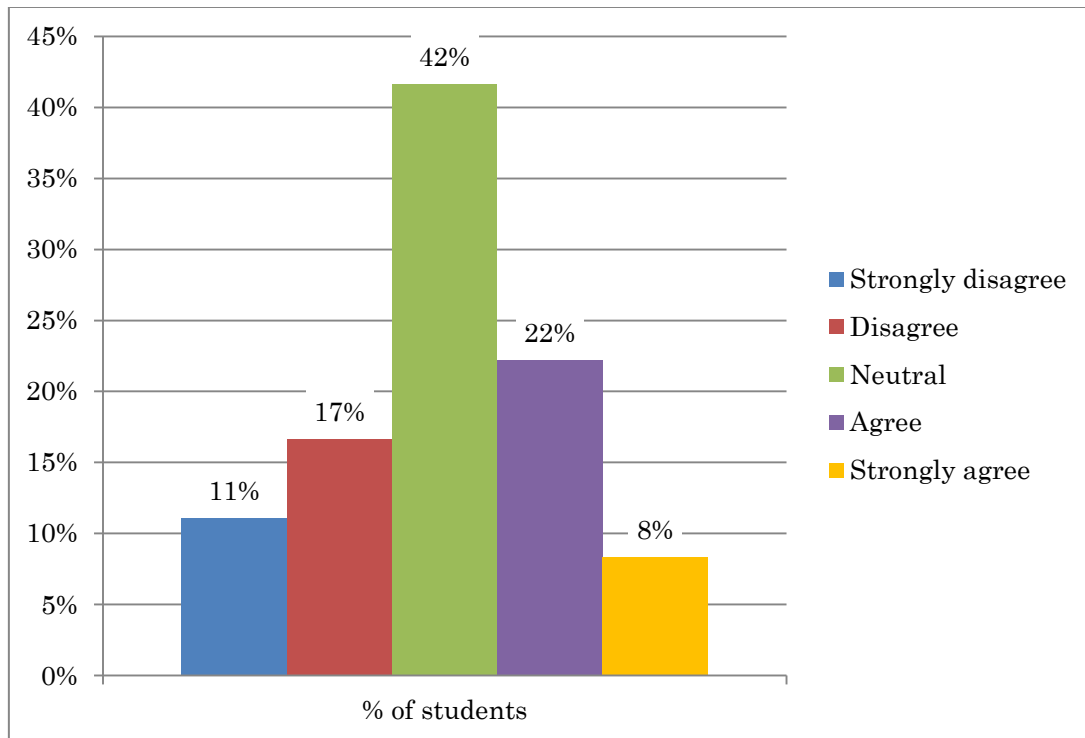
Figure 7 - Summary of students' answers for Q5

Figure 7 shows that again about one third of our small students sample here would prefer more readings and CRS sessions based on those readings. As our students are native Japanese - yet taught, nevertheless entirely in English - it is logical then perhaps that some have stated a preference for texts because it is may be easier for them to acquire, among other things, a valuable vocabulary (and more time and purpose to acquire it). And indeed this point was explicitly included in several comments.

The authors feel that this is a serious consideration, though, in the end analysis, this may be a trade-off or balance to be struck between our stated college missions (a greater focus on ideas and critical thinking) and our obvious constraints over non-native language teaching and instruction. Therefore, elements such as vocabulary-learning, particularly in technical areas,

may need to remain largely a matter of on-the-fly delivery during instruction (and perhaps in CRS question-defining pre-polling moments) or that it is emphasized more as a student's own responsibility (if the teacher pre-advises students, of course, over what concepts will appear, and need to be learned for overall course themes and elements.

The number of "neutral" responses given by students here may be more interesting and challenging to interpret. To conjecture on possible explanations for it, what immediately comes to mind is the notion that students simply do not know whether it would have been better for them, and in their interests, to have had more reading. And in such cases, "neutral" is as good and as valid a response as any. In this sense, another adage can be called forth, namely, 'that what one does not have, one cannot really miss.'

But yet another interpretation can be put forth. And this is purely one that stems from the perspective of student-as-consumer of education, and/or from a purely personal developmental point of view. Firstly, at time of survey, students simply did not have a great deal of time to establish, in their mind what indeed it was that they retained (and can feel for themselves for what they have taken from their learning experience). So in the absence of such a felt opinion, and not having the time in which such a perception or opinion may crystallize, then the more tangible index (of "having-learned it," learned a body of something) - is a textbook. But for us, this is pure speculation. But it's based upon personal reflection of our own biography of student years.

Additional (freer) comments provided by students were instructive, constructive and encouraging. Apart from comments already discussed as dispersed in the preceding passages, three main points appearing with an interpreted level of recurrence were: a) CRS helped

students in following the class and understanding its content; b) Students had a lot of fun using the CRS; and, c) listening to teachers' explanation right after answering questions was a 'consolidating advantage.

Future directions and some further discussion

In large part, and as educators, we feel that the CRS system gave us some culturally beneficial opportunities to gauge Japanese students' comprehension when these students may have been otherwise reluctant to speak and communicate with us in class. This is an omnipresent challenge for foreigners teaching in Japan (see Hammond, 2007). And yet arguably it is one that almost by default (should) encourage better and creative teaching and the adoption of more adaptive and humble attitudes.

Not the least of these is cultivating an ability to not take too much for granted, and to realize that while we lecturers may like to talk and hear ourselves for extended periods, others may not. And in any case the effectiveness of this has to be questioned – everywhere. And it is being seriously questioned in western higher education and in the Commonwealth. These education systems are coming to recognize and deal with their need to move away from teaching viewed as a one-way and a 'one-to-the many' (information transfer) process and enterprise (see, Wlodkowski & Ginsberg, 2009). It is perceived that things must move towards a mixed, more engaging and participatory learning framework. And as such, much of what we do here, of necessity in the ESL and culturally unique Japanese context, is already globally-oriented and progressive. For its part, CRS is just another tool in the tool chest that

foregrounds these overall philosophies, exigencies, and what's essentially in (or of) these local methods that are already practiced here.

However, there is something quite particular about CRS (here and at this overall cultural juncture). And this picks out the instantaneousness, and, (we have to say) the measure of screen presence and involvement with wireless device. Both of these aspects (i.e., instantaneousness, and "the screen") are consistent with one relatively new and one old or tradition pedagogical constraint.

Firstly, these highly visual, instant-result, and technologically interactive elements are – whether we like it or not – culturally, technologically and intuitively/cognitively in sync with today's student (everywhere in the industrialized world). And that young person is, him or herself, highly adapted (in fact, born into) a very different and dynamic world than the old classroom and the teachers that inhabit it (and still control it). One of the authors has recently joked with his class that one the buttons on the CRS/clicker remote, the button [E] – not often used as an option for multiple choice (A-D, in our usage) – can be used, instead, to mute the teacher if it is felt that he is becoming boring for any protracted period. Joking aside, while we should not have to make-merry with students and our classes (and of course, there is work to do) if these young people are jaded, bored, and not engaged – we are not really educating and they are not really being educated and educating themselves.

Secondly, and relatedly, it is the case for teachers (again anywhere but at any historical time and with any type of student) that it may not be apparent in an ongoing class or 'lecture' period (even if it broadly seems student centered or active) – whether they "are getting it." Paper and pencil tests, quizzes, or exercises all take days if not weeks to grade and to establish

this. And if the class/students has manifestly and essentially not been getting it – as may be now clear from laborious (teacher-take-home) grading, then revisiting the topics and arguments now possibly two class sessions later, well after the fact, is an obvious problem. And it is so in several respects. Plainly, the class is now likely to be working on a different topic or indeed module altogether. Furthermore, if the success of course in question really hinges upon a continued following or learning of a building-block/foundational and sequential conversation, then failed teaching and learning (from 2 classes prior) will not just involve going back to amend-learn and literally re-cover certain distinct ideas. It will involve a very much larger wheel to re-invent, and this might just not be practically possible. There may have occurred a collapse of the overarching and critical point of (or that underpinned) the whole course offering. And this may never be re-coverable.

So the provision of real-time descriptive statistics derived from CRS in-class polling is more than a mere gimmick. Though it does nevertheless (depending on presentation) quite often entertain – though it is always engaging and a useful and re-occurring trigger for modulating alertness (again) after the intellectual effort of responding, while perhaps waiting for others to 'vote' and for polling periods to close. Substantively, the timely elaboration and explanation of question-contained points, as they are given (post-polling) within the temporal aspects of its teacher-planned and sequential context, is simply invaluable.

It has been for these given (primary) reasons that we have made some use of CRS over the past 2 years. Incidentally, and for future plans, it is worth noting that virtually all our students here at Miyazaki travel abroad for study outside of Japan. This is integrated into their degree (and if we may also remind the reader that these students undertake their entire higher education, at home, also in a difficult-to-learn second language - English). Further and future

surveys might then possibly tap on pre- and post-travel juxtapositions of the differing pedagogical systems that these young people will experience. And these (surveys) may also involve investigation of the traveling students' likely experiences with (larger-class) CRS use abroad.

But for us also, and administratively, it may be worth pointing out (the often-overlooked feature) that these systems automatically record class attendance without sign-in sheets (at least for those class days for which CRS is used). Presumably, the option (mentioned earlier) to build in smaller clicker sessions of 2-3 questions into every class period would make this provision more attractive (by allowing teachers to dispense with sign-in sheets altogether).

And of course, the more substantive aspects of grading-work and course administration are considerably reduced for instructors as a whole with CRS. We like also that a CRS can be integrated within the course grading policy to any degree desired by the teacher. Student teaching reviews on the subject of CRS (as shown here), and overall teacher/class reviews illustrate favor towards the general idea of continuous assessment. Students may feel that too much weight is placed on final exams, and the ongoing assessment may be seen as a break from this to an extent. This might be a useful consideration in institutions or cultures and curriculums where a heavy of full class load is built into the structure of degree requirements (as is the case here). Freed from the need to "cram" at semester's mid-point and end – for perhaps one or two courses in their semester's course diet – students may feel less stress and pressure. This reduced apprehensiveness may (or may not) be conducive to greater retention of themes and concept.

Lastly, we would like to thank the students for their participation in survey and in CRS use. We are most grateful for the feedback.

References

Atkinson, R., McBeath, C., Soong, S., & Cheers, C. (n.d.). ICT: Providing choices for learners and learning.. *Proceedings Ascilite Singapore 2007*. Retrieved June 17, 2013, from http://www.ascilite.org.au/conferences/singapore07/procs/

Bruff, D. (2009). *Teaching with classroom response systems: creating active learning environments*. San Francisco: Jossey-Bass.

Duncan, D. (2006). Clickers: A New Teaching Aid With Exceptional Promise. *Astronomy Education Review*, *5*(1), 70.

Hammond, C. (2007). Culturally responsive teaching in the Japanese classroom: a comparative analysis of cultural teaching and learning styles in Japan and in the United States. *Journal of the Faculty of Economics*, *17*, 41-50.

Hoekstra, A. (2008). Vibrant Student Voices: Exploring Effects Of The Use Of Clickers In Large College Courses. *Learning, Media and Technology*, *33*(4), 329-341.

Laurillard, D. (2002). *Rethinking university teaching a conversational framework for the effective use of learning technologies* (2nd ed.). London: Routledge/Falmer.

Lucas, A. (2009). Using Peer Instruction And I-Clickers To Enhance Student Participation In Calculus. *PRIMUS*, *19*(3), 219-231.

Mayer, R., Knight, A., Campbell, J., Bulger, M., Chun, D., Bimber, B., et al. (2009). Clickers

    In College Classrooms: Fostering Learning With Questioning Methods In Large

    Lecture Classes. *Contemporary Educational Psychology*, *34*(1), 51-57.

Sprague, E., & Dahl, D. (2009). Learning to click: an evaluation of the personal response

    system clicker technology in introductory marketing courses. *Journal of Marketing*

    Education, *32*(1), 93-103.

Wlodkowski, R. J., & Ginsberg, M. B. (2009). *Diversity and motivation: culturally responsive*

    teaching (2 ed.). San Francisco: Jossey-Bass Publishers.