The Asian EFL Journal

Professional Teaching Articles

November 2016

Issue 96

Senior Editors:

Paul Robertson and John Adamson

Production Editor:

I-Chin Nonie Chiang

Table of Contents

A study on the phonology of Korean L1 and English L2: Phoneme and sub-syllabic structure differences[1]

Chaehee Park
Sun Moon University, Republic of Korea

Bioprofile:

Chaehee Park, Ph.D. is an assistant professor in the department of English at Sun Moon University, Republic of Korea. The research area mainly includes L2 spelling and vocabulary acquisition. Email: chpark@sunmoon.ac.kr

Abstract

The purpose of this study is to examine which aspect of L2 phonology, sub-syllabic structure or phonological congruence, is more related to English L2 spelling accuracy. This study investigated sub-syllabic structure and phonology of English L2 in comparison to Korean L1. Two groups of college students participated in Korean and English sound similarity tasks and English spelling tasks. The pseudowords for the tasks were created from real words by changing one letter. The pseudoword list was recorded by a native English speaker and participants listened to them in isolation and recorded the pseudowords manually. The results of the experiments indicated the body of syllable is a more salient part in Korean L1, Korean EFL learners are more sensitive to body-coda segment rather than onset-rhyme segment in English L2 as well, and Korean EFL learners were more sensitive to sub-syllabic structure (body-coda) rather than the phonemic congruence in both English sound judgment tasks and English spelling tasks. It is concluded that English L2 spelling accuracy of Korean EFL learners relates to Korean language linguistic properties, which is body-coda prominent rather than phonemic congruence between Korean L1 and English L2. It is suggested that in addition to phoneme differences between L1 and L2, sub-syllabic structure differences deserve much more attention in English education to help EFL learners develop literacy.

Keywords: English L2 spelling, phonology, sub-syllabic structure, phonological awareness

---

English Department
Sun Moon University, Tangjeong-Myeon,
Asan-Si, Chungnam, S. Korea

4

Introduction

Researchers have argued that, in an alphabetic writing system, phonological awareness (PA), the child's ability to reflect on the sound structure of oral language (syllables, onset-rimes, phonemes) and manipulation of these components separately from their meaning (McBride-Chang, 1995) is a critically important skill in early literacy development. The speech sounds usually include phonological units such as onset-nucleus-coda and syllable levels, and children's awareness of these phonological units develops sequentially in English (Stahl & Murray, 1998; Treiman, 1985; 1992; Treiman & Zukowski, 1991).The internal structure of syllable also relates to the development of PA. One of the important views on this syllable internal structure is that two-phoneme sequences are intrinsically more closely associated with each other than the other two-phoneme sequences in the syllable (Fowler, Treiman, & Gross, 1993). Specifically, some researchers claimed that English speakers parse CVC monosyllabic words into a C/VC partition (Fowler, Treiman, & Gross, 1993; Treiman, 1983; Treiman, Salasoo, Slowiaczek, & Pisoni, 1982) in which the major boundary lies between the onset and the nucleus of the syllable. However, other researchers of Korean language have found that Korean children have a preference for syllable-initial CV (body) units over syllable-final VC (rhyme) sequences in their early literacy development (Kim, 2007; Yoon & Derwing, 2001). These studies confirm that the onset-rhyme boundary prominent in English is not universally accessible across alphabetic languages (Ziegler & Goswami, 2005).

There has been continuous research on L2 spelling focusing on phonological differences (Wang & Geva, 2003; Yeong & Liow, 2010). For instance, Park (2012b) found that Korean learners of English had difficulty in spelling phonemes which do not exist in Korean but are present in English (e.g., /f/ and /v/). Similar results were also found in other L2 background learners (Bebout, 1985; Cook, 1997; Fashola, Drum, Mayer, & Kang, 1996; Ibrahim, 1978; Wang & Geva, 2003). Although contrastive analysis of L2 spelling performance focusing on phonological difference has been increased, studies emphasizing sub-syllabic structure differences have been ignored. The two goals of this study were to examine English L2 spelling performance focusing on sub-syllabic structure differences and phoneme differences between Korean L1 and English L2 and to understand which of these two differences plays a more significant role in English L2 spelling performance.

Research on Phonology and Orthography and L2 Spelling Skills

Reading and spelling acquisition involves the association between the letter and sound. In an alphabetic language, sound can be mapped onto the letter which consists of a vowel or consonant. For instance, one syllable word (sound) *cat* consists of phonemes /k/ (onset), /æ/ (nucleus), /t/ (coda) and children in a certain stage of phonological development are able to segment *cat* into /k/, /æ/, /t/. Some research pointed out that phonological knowledge is a crucial underlying component in reading and spelling development (Goswami & Bryant, 1990; Lundberg, Olofsson, & Wall, 1980; Mann & Liberman, 1984) and that early phonological processing skills, the ability to analyze and manipulate sound units of spoken language, predict later reading success (Goswami & Bryant, 1990; Lundberg, Olofsson, & Wall, 1980; Shankweiler, 1999).

The importance of phonology in spelling is also noted in several studies (Bryant & Bradley, 1980; Goswami & Bryant, 1990). They categorized the words into four groups on the basis of children's performance of reading and spelling words. Bryant and Bradley (1980) gave six and seven year old children lists of words to read and write and found empirical evidence for all four categories of words: 1) both read and spell, 2) neither read nor spell, 3) only read not spell, and 4) only spell not read. Categories 1 through 3 are normal occurrences which are common with views of children's reading and writing. Children may be able to read some familiar words and to spell them as well (*category1*) even at the earliest stage of learning, or may not be able to read and spell some words which are more difficult and less familiar (*category2*). However, they found that there was a clear and consistent difference in the types of words which children read but did not spell (*category3*) and those which they spelled but did not read (*category4*). They might be able to read some words which they were not able to spell (*category3)* and the most common words of this category include *light*, *school*, *train*, and *egg* because "these do not lend themselves well to a grapheme phoneme correspondence strategy" (p. 59). In contrast, they found the most frequent words in category 4 require a phonological code such as *bun*, *mat*, *leg* and *pat*. This result suggests that children depend on the strategy for letter sound relations and phonological code, in order to spell but not in order to read (Bryant & Bradley, 1980).

Knowing that phonological knowledge of language can be a useful strategy in order for children to spell words, orthography of language has been a major topic in

spelling research. Orthography is defined as conventional spellings including spelling rules and spelling patterns of the language (Varnhager, Boechler, & Steffler, 1999). The knowledge of orthography, which is defined as children's understanding of the conventions used in the writing system of their language (Treiman & Cassar, 1997), has been shown to facilitate children's spelling and reading acquisition. For instance, Cassar and Treiman (1997) found that children have acquired some knowledge of the acceptable form and position of consonant doublets by late kindergarten. These young children preferred spellings with acceptable doublets (e.g., *yill*) to those with unacceptable doublets (e.g., *bbaf*). One major theory for discussing processing differences in orthographies among alphabetic language systems is the Orthographic Depth Hypothesis (ODH) (Katz & Frost, 1992). According to ODH, in transparent orthographies such as Serbo-Croatian, Spanish, and Italian, there is a relatively simple and consistent one to one correspondence between grapheme and phoneme. For instance, Serbo-Croatian is a major language in the Balkan Peninsula, and its present alphabet was introduced based on the principle "Spell a word like it sounds and speak it the way it is spelled." In other words, each grapheme presents only one phoneme in the spoken word and each phoneme is represented by only one letter. The relation between graphemes and phonemes is also isomorphic and exhaustive, and the spelling system follows the phonemic structure of spoken words.

However, in opaque orthographies such as English, there is a more complex and less consistent relationship between graphemes and phonemes. For instance, grapheme <a> can be pronounced differently in the words *cat*, *about*, *all*, and *lady* in English. In addition to the differences in grapheme-phoneme correspondence at the phonemic level, English morphophonemic spelling makes use of this principle even more extensively and in even more diverse ways. For instance, a morpheme can be represented by different sounds, but these are written in the same way; the plural morpheme in English is represented by three different sounds as in *cars* [-z], *cats* [-s], *cases* [-Iz], but they can be written with the same letter <s>. Orthographic awareness refers to "a visually mediated ability to analyze and recognize letter and letter strings" (Katzir, Kim, Wolf, Kennedy, Morris, & Lovett, 2006, p. 846) and is very important in spelling words correctly in a relatively opaque orthography due to this inconsistent correspondence between grapheme and phoneme. It has been found that children need to learn spelling rules explicitly in an opaque orthography such as English (Bryant, 2002; Kessler & Treiman, 2003). They must learn alternative spelling rules for the

same sound as well and utilize these alternative rules to spell words correctly. Researchers also suggest that phonological and orthographic knowledge mutually facilitate each other and that grapheme-phoneme knowledge provides young learners with a powerful tool to bind the spelling patterns of individual and multiple letters with their pronunciation in words (Ehri, 1998).

Knowing that phonology and orthography play an important role in spelling skills, it is reasonable that phonology and orthographic knowledge of L2 is quite important in L2 spelling skills. First, it seems that L1's phonology and orthographic rules are transferred to L2 spelling performance. Fashola and his colleagues (1996) investigated the English spelling errors of Spanish speaking children on the basis of phoneme differences between two languages. It was found that Spanish children committed more predicted errors (caused by applying Spanish phonological and orthographical rules) than the English speaking children. For example, a Spanish-speaking child might spell the English word *soccer* as *soker* in which *cc* can be replaced with *k* in Spanish phonology. It implies that Spanish children are applying their L1 spelling rules to the English L2 spelling. However, it is difficult to assume that these errors were made due to an influence of Spanish orthography since they did not provide information on the Spanish-speaking children's prior literacy experience in Spanish. With more carefully designed experiments, Sun-Alperin and Wang (2008) also found that native Spanish-speaking children committed significantly more vowel spelling errors that were consistent with Spanish orthography. That is, error types were more phonologically legitimate in Spanish (e.g., *mit* for meat and *mun* for moon), which suggests that orthographic properties of Spanish L1 influence children learning to spell in English L2.

Similar results were also found in non-Roman alphabetic L1 background. Park (2012a) investigated English pseudoword spelling performance comparing native English children and Korean English L2 learners. He found that the Korean group significantly made more errors in the less consistent pseudowords than in the more consistent pseudowords, suggesting that the learners' L1 orthographic transparency influences L2 spelling accuracy. Vowels in Korean have direct one-to-one grapheme-phoneme correspondence whereas vowels in English usually have multiple spellings. Learners with rather transparent L1 background (direct one-to-one grapheme phoneme correspondence) had difficulty in spelling words of less

transparent L2 language. These findings support the idea that orthographic transparency between L1 and L2 affects spelling development in English L2.

In the same vein, spelling errors also occur if there are incongruent phonemes between L1 and English L2. When English L2 learners replace the English phoneme that does not exist in L1 with the other phonemes that exist in their L1, errors were more likely to occur (Bebout, 1985; Cook, 1997; Fashola et al., 1996; Ibrahim, 1978; Park, 2012b; Wang & Geva, 2003; Yeong & Liow, 2010). Ibrahim (1978) collected samples of undergraduate student writing (examinations, homework assignments, papers, reports), and found that students tended to make errors on the phonemes which do not exist in their Arabic L1. For instance, students replaced the phoneme /p/ with the phoneme /b/. Spelling errors such as the following are fairly frequent in Arabic; *bicture* (picture), *compination* (combination), *distripution* (distribution), *clup* (club), *hapit* (habit), and *pit* (bit) due to the absence of /p/ in Arabic. Because many Arabic dialects do not permit final consonant clusters in the coda position, students also made more errors in consonant clusters. Research comparing English L1 speakers and English L2 speakers has shown that English L2 phoneme awareness, especially the phonemes found only in English L2, need to receive special attention for ESL learners (Park, 2012b; Wang & Geva, 2003; Yeong & Liow, 2010). For instance, Park (2012b) conducted experiments on English L2 spelling performance using psudowords and found that Korean learners of English made more errors on the phonemes found only in English such as /f/ and /v/. He could not find any statistically significant spelling errors of those phonemes among native English speakers, suggesting that phonemes found only in English can be challenging to Korean Learners of English. Concerning the analysis of spelling errors, it was found that phonemes found only in English were replaced by the articulatory adjacent sounds such as bilabial /p/, forward position and alveo-dental /t/ and /s/, backward position from the target /f/; /b/ (34.5%), forward position, alveo-dental /d/ (11.1%), backward position from the target /v/, implying that phonological properties of Korean L1 influence children's learning to spell in English L2.

The link between phonology and orthography is central to the Psycholinguistic Grain Size Theory (PGST) (Ziegler & Goswami, 2005). PGST fundamentally relates to reading development. Because spelling involves phonological and orthographic process as the reading process does, introducing PGST helps to understand the spelling process. The PGST assumes that grain size along with

orthographic consistency plays an important role in learning to read. According to Ziegler and Goswami (2005), the phonological contribution seems to be made at larger grain size units (e.g., the whole syllable *grasp* or onset-rhyme *gr-asp*) at the early stages of reading acquisition. When children begin learning to read and spell in later stages, smaller grain size units such as onset-nucleus-coda (*gr-a-sp*) play a more significant role. For instance, the substructure of the one syllable word *cat* consists of /k/ (onset), /æ/ (nucleus), /t/ (coda). Initially, they seem to recognize the sound of the word *cat* as a whole and begin segmenting the sound to a rather larger grain size; onset-rhymes (/k/ - /æt/). However, when children start learning to read and spell, they seem to recognize smaller grain size units of the word such as /k/, /æ/, /t/. The smaller grain size units (letters or phonemes) are most important in transparent language systems such as Greek, German, and Spanish because the phonemes, the smaller unit, are consistent and reliable to correspond to the grapheme. In contrast, larger grain size units are considered more important in reading in an opaque orthography language system. Phonemes (the smaller unit) are more inconsistent and unreliable than larger units (syllables and rhymes) in English: inconsistent grapheme phoneme correspondence. According to this hypothesis, in less transparent orthography, the larger grain size units (e.g., syllable, rhyme) are easier to recognize than the smaller grain size units (e.g., phonemes).

Given the fact that grain size units play an important role in literacy development, English and Korean show quite different sub-syllabic patterns in which particular segments in a syllable are associated with each other. Onset is the initial consonant or a consonant cluster of a syllable, while rhyme refers to the syllable-final vowel or vowel-consonant. Studies in the field of phonology generally report that there exist sub-syllabic patterns in which particular segments in a syllable are associated with each other. For instance, boundary between onset and rhyme is natural in English (Treiman, 1992; Treiman & Zukowski, 1991). English speakers parse CVC monosyllabic words into a C/VC partition (Fowler, Treiman, & Gross, 1993; Treiman, 1983) in which the major boundary lies between the onset and the nucleus of the syllable. This finding is also supported by the prominence of the rhyme unit in reading English (Goswami, 1988;1993;1998; Kessler & Treiman, 1997; Treiman, Mullenniz, Bijeljac-Babic, & Richmond-Welty, 1995).

However, as for the Korean sub-syllabic structure, it has been found that CV/C partition is more salient (body-coda prominent) than C/VC partition in

children's literacy development (Kim, 2007; Yoon, Bolger, Kwon, Perfetti, 2002; Yoon & Derwing, 2001). For instance, Kim (2007) used blending and segmenting tasks to examine four different phonological units: syllable, onset-rhyme, body-coda, and phoneme. She found that the onset-rhyme unit tasks were the most difficult and followed by phoneme task, body-coda task. The syllable tasks were easiest for the children, indicating that body–coda boundary (e.g., ca-t) is more salient than onset-rhyme boundary (e.g., c-at) for Korean children. This finding suggests that Korean children's knowledge of the similarity of the spelling sequence in the body facilitates word reading and spelling development in Korean.

Investigative Framework

The insights gained from the above literature review can be summarized as follows:

(1) Spelling ability requires knowledge of orthography, phonology, and orthography-phonology (letter-sound) mapping.

(2) Experience in processing L1 orthography, phonology, and orthography-phonology (letter-sound) mapping transfers to L2 processing, which in turn influences L2 spelling acquisition.

(3) English and Korean differ in terms of their phonology, orthography, orthography-phonology (letter- sound) mapping, and sub-syllabic structure prominence.

Research Questions

The goal of this study was to extend the documentation of L2 spelling research by examining English L2 spelling performance in terms of phoneme and sub-syllabic structure differences between L1 and L2. The following research questions were posed:

(1) How does the sensitivity to syllable segment (structure) of native Korean speakers differ in their L1?

(2) How do Korean learners of English sense English L2 phonemes (onset-rhyme vs. body coda in their English L2)

(3) Do Korean L1 learners of English sense phonemes that exist only in English but not in Korean? If not, does their sensitivity to syllable segment relate to phoneme differences or (presence vs. absence) or syllable position differences (syllable initial or syllable final)?

(4) Do Korean L1 learners of English spell English words that include phonemes that

do not exist in Korean? If not, does their spelling accuracy relate to L1 phoneme differences (presence vs. absence) or sub-syllabic structure differences of English and Korean (syllable initial or syllable final)?

The present study

Participants

The participants were native speakers of Korean who were learning English as a foreign language. One group of students (18 students, 11 female and 7 male students) participated in Korean sound judgment task and English sound judgment task. Another group of students (27 students, 15 female and 12 male students) participated in English spelling task. The mean age of the participants was 18.6 years (SD = 0.85) and 20.5 years (SD = 0.91) respectively. All participants had received formal education in their L1 through high school, which ensured they were phonologically developed in their L1. The exposure to English L2 education was also taken into consideration. All subjects had studied English during their elementary, middle and high school years, which ensured they were also phonologically developed in English L2.

Measures

Korean sound similarity judgment task (KSSJ task)

The sound similarity judgment (SSJ) task was introduced as a tool for investigating phonological units (Greenberg & Jenkins, 1964). In SSJ experiment, similarity of overall sound is rated by participants who listen to pairs of words. More recently, Yoon and Derwing (2001) conducted a new SSJ experiment in Korean. They tested the relative weights of the subcomponents of CVC syllable to see whether there were position effects (body vs. rhyme). The results showed that the body is a more salient part of the Korean syllable than the rhyme in predicting global sound similarity judgments. Following a task used in previous syllable structure research in Korean (Yoon & Derwing, 2001; Yoon et al.,2002), a total of 15 test sets (three types of phoneme matches X five token of each type) were created. The stimulus set was derived from the four consonants /p/ <ㅂ>, /t/ <ㄷ>, /m/ <ㅁ>, /n/ <ㄴ> and the four vowels /a/ <ㅏ>, /ə/ <ㅓ>, /u/ <ㅜ>, /o/ <ㅗ>. Table 1 presents three different types of words with examples.

Table 1. Examples of stimulus pairs used in the KSSJ task

Number of matched phonemes	Word Types	Examples	Number of tokens
3	Same	tom-tom	5
2	B-sharing	pam-pan	5
2	R-sharing	tom-pom	5

*B-sharing: body sharing words.

*R-sharing: rhyme sharing words.

English sound similarity judgment task

English CVC monosyllabic pseudowords were created in an English sound similarity judgment (ESSJ) task. There were 15 test sets (five types of phoneme matches Ｘ three tokens of each type). The stimulus set includes phonemes which do not exist in Korean but are present in English such as /f/, /v/, /r/ in either onset or coda. Table 2 presents five different types of words with examples.

Table2. Examples of stimulus pairs used in the ESSJ task

Number of matched phonemes	Word types	Examples	Number of tokens
3	Same	pin-pin	3
2	B-sharing	mis-mit	3
2	R-sharing	mis-tis	3
2	O-A-B-sharing	fis-fin	3
2	C-A-R-sharing	pef-mef	3

*B-sharing: body sharing words.
*R-sharing: rhyme sharing words.
*O-A-B-sharing: onset absent phoneme but body sharing words
*C-A-R-Sharing: body absent phoneme but rhyme sharing words.

Spelling task

Following a task used in previous L2 spelling research (Sun-Alperin & Wang, 2008; Wang & Geva, 2003), a pseudoword spelling task was used in this test. Pseudowords are pronounceable combinations of letters that do not include semantic

meaning but can be decoded and pronounced via phonological processing and alphabetic (letter-sound) knowledge (Frederickson, Frith, & Reason, 1997). The reason for using pseudowords was to avoid lexical knowledge effect on spelling performance. Students can spell familiar real words by sight word memory without converting phonemes to letters. Twenty monosyllabic pseudowords were created. Each pseudoword item included target phonemes which do not exist in Korean (/f/, /v/, /θ/, /ð/, /ɹ/ as English only consonants) and phonemes which exist in both languages (/m/, /n/, /t/, /k/ /p/, /s/). Five items for the congruent phonemes, 5 for English only consonants in onset, 5 for English only consonants in coda, and 5 for the English only consonants both in onset and coda were created. Each target consonant was placed into either onset only (e.g., [fip]), coda only (e.g., [pif]), or onset and coda simultaneously (e.g., [fiv]). Table 3 presents five different types of pseudowords with examples.

Table 3. Examples of pseudowords used in the spelling task

Word types	examples	Number of tokens
all present	pim	5
coda absent	pif	5
onset absent	fip	5
onset-coda absent	fiv	5

A coding scheme was adapted for scoring the task to reflect word spelling performance more precisely. Following the procedure used in Treiman, Berch, and Weatherston (1993), the spelling the children produced was scored as either correct or incorrect according to the native speakers' answers. Only consonants (onset and coda) were considered to score points. For the initial consonant (onset), when participants spelled only one letter (consonant as in *p _ _* for [*pim*]), or consonant and vowel (e.g., *p i _* for [*pim*]) in order, it was considered as the initial consonant. For the final consonant, when the participants spelled vowel and following consonant (e.g., *_ i f* for [*pif*]), it was counted as the final consonant. When the participants spell two consonants without a vowel between them (e.g., *m _ p* for [*mip*]), the first was counted as initial <m> and the second as final <p>. Participants received two points for items

answered correctly, one point for items misspelled by one letter, and zero points for items misspelled by both consonants.

Procedure

The stimulus pairs of KSSJ and ESSJ tasks were recorded by native speakers of each language. Each pair was recorded in a random order, and played back to the participants with three second intervals. Before the task, participants were informed that the word pairs of KSSJ were made from Korean words, and the word pairs of ESSJ from the English words. Participants were instructed to focus on global impression of sound only and to rate the similarity in sound for each pair on a five point scale, ranging from 1 (completely different) to 5 (exactly the same). Five practice sessions for each task were given to check whether the participants understood and were following the instruction. The pseudowords for spelling task were recorded and played back to participants. Before listening to each pseudoword, participants were informed that the words are English monosyllabic nonsense words and instructed to focus on sound not the meaning of the words. Participants listened to each pseudoword twice with five second intervals between each word. They were encouraged to pay attention to each sound (pseudoword) as much as possible and write them using the English alphabet. Three practice sessions were given to make sure the participants followed the instructions.

Results

Table 4 illustrates summaries of mean scores and standard deviation of each measures. For the KSSJ task, the mean scores of the same words was 21 (SD = 2.6), body sharing words, 9.4 (SD = 4.1), and rhyme sharing words, 6.4 (SD = 2.9) respectively. For the ESSJ task, the mean scores of the same words 11.6 (SD = 2.3), body sharing words, 6 (SD = 2.3), and rhyme sharing words, 4 (SD = 1.7), onset-absent-body sharing words, 5.7 (SD = 2.7), and coda-absent-rhyme sharing words, 4 (SD = 2), respectively.

Table 4. Means and standard deviation of each measure.

Measures	Word types	Examples	Mean (Max)	Std. Deviation	N
KSSJ task	Same	tom-tom	21.06 (25)	2.689	18
	B-sharing	pam-pan	9.44 (25)	4.119	18
	R-sharing	tom-pom	6.44 (25)	2.915	18
ESSJ task	Same	pin-pin	11.67 (15)	2.301	18
	B-sharing	mis-mit	6.06 (15)	2.235	18
	R-sharing	mis-tis	4.06 (15)	1.798	18
	O-A-B-sharing	fis-fin	5.78 (15)	2.713	18
	C-A-R-sharing	pef-mef	4.00 (15)	2.086	18

For KSSJ, the mean scores of the same words were higher than the mean scores of body sharing words and rhyme sharing words. The results also show that the mean scores of the body sharing words was higher than the mean scores of rhyme sharing words. For the ESSJ task, the mean scores of the same words was higher than the mean scores of body sharing words, rhyme sharing words, onset-absent-body sharing words, and coda-absent rhyme sharing words. The mean scores of body sharing words was also higher than the mean scores of rhyme sharing words, and the mean scores of onset-absent-body sharing words was higher than the mean scores of the coda-absent-rhyme sharing words as well. For the KSSJ task, a one-way repeated measures ANOVA was performed with word types (all presents vs. body sharing vs. onset sharing) as the within-word type factor to verify whether mean differences are statistically significant. A one-way repeated measures ANOVA on the mean differences of each word type was significant, $F(2, 34) = 126.463$, $p<.01$, indicating that the mean differences among word types are significantly large. Given the main effect for the word type was significant, it was necessary to compare each mean score to obtain more detailed results on mean differences. Table 5 illustrates mean differences, standard error, significance value, and confidence interval for mean differences of the KSSJ task. Significance values show that the comparison of mean scores between the same words and body sharing words was significant (p<.01), the comparison of mean scores between the same words and rhyme sharing words was significant (p<.01), and the comparison of mean scores between the rhyme sharing words and body sharing words was also significant (p<.01).

Table 5. Mean comparisons of KSSJ task in each word type

Word types		Mean Difference	Std. Error	Sig.	95% Confidence Interval for Difference	
					Lower Bound	Upper Bound
Same	B-sharing	11.611[*]	1.118	.000	9.253	13.969
	R-sharing	14.611[*]	1.061	.000	12.373	16.849
B-sharing	Same	-11.611[*]	1.118	.000	-13.969	-9.253
	R-sharing	3.000[*]	.672	.000	1.583	4.417
R-sharing	Same	-14.611[*]	1.061	.000	-16.849	-12.373
	B-sharing	-3.000[*]	.672	.000	-4.417	-1.583

For the ESSJ task, a one-way repeated measures ANOVA was performed with word types (same words vs. body sharing vs. rhyme sharing vs. onset-absent body sharing vs. coda-absent rhyme sharing) as the within-word type factor to verify whether mean differences among them are statistically significant. A one-way repeated measures ANOVA on the mean differences of each word type was significant, $F (4, 68) = 40.614$, $p<.01$, indicating that the mean differences among word types were significant. Because the main effect for the word type was significant, it was necessary to compare each mean score to obtain more detailed results on mean differences. Table 6 illustrates mean differences, standard error, significance value, and confidence interval for mean differences of the ESSJ task. Significance values show that the comparison of mean scores between the same words and body sharing words, rhyme sharing words, onset-absent body sharing words and coda-absent rhyme sharing words was significant ($p<.01$). The comparison of mean scores between the body sharing words and rhyme sharing words, and coda-absent rhyme sharing words was significant respectively ($p<.01$), the comparison of mean scores between the rhyme sharing words and onset-absent body sharing words were also significant ($p<.01$), as were the comparison of mean scores between the onset-absent body sharing words and coda-absent rhyme sharing words ($p<.01$). However, the comparison between the mean scores of body sharing words and onset-absent body sharing words was not significant ($p>.05$), nor were the mean scores of rhyme sharing words and coda-absent rhyme sharing words ($p>.05$).

Table 6. Mean comparisons of ESSJ task in each word type

Word types		Mean Difference	Std. Error	Sig.	95% Confidence Interval for Difference	
					Lower Bound	Upper Bound
Same	B-sharing	5.611*	.780	.000	3.965	7.257
	R-sharing	7.611*	.705	.000	6.123	9.099
	O-A-B-sharing	5.889*	.867	.000	4.061	7.717
	C-A-R-sharing	7.667*	.878	.000	5.814	9.520
B-sharing	Same	-5.611*	.780	.000	-7.257	-3.965
	R-Sharing	2.000*	.542	.002	.856	3.144
	O-A-B-sharing	.278	.808	.735	-1.426	1.982
	C-A-R-sharing	2.056*	.551	.002	.893	3.218
R-sharing	Same	-7.611*	.705	.000	-9.099	-6.123
	B-sharing	-2.000*	.542	.002	-3.144	-.856
	O-A-B-sharing	-1.722*	.547	.006	-2.877	-.568
	C-A-R-sharing	.056	.495	.912	-.989	1.100
O-A-B-sharing	Same	-5.889*	.867	.000	-7.717	-4.061
	B-sharing	-.278	.808	.735	-1.982	1.426
	R-sharing	1.722*	.547	.006	.568	2.877
	C-A-R-sharing	1.778*	.660	.015	.386	3.169
C-A-R-sharing	Same	-7.667*	.878	.000	-9.520	-5.814
	B-sharing	-2.056*	.551	.002	-3.218	-.893
	R-sharing	-.056	.495	.912	-1.100	.989
	O-A-B-sharing	-1.778*	.660	.015	-3.169	-.386

In regard to English spelling words, the mean scores of all present words were higher than the mean scores of coda absent, onset absent, and all onset-coda absent words. The mean scores of coda absent words were also higher than onset absent words. The mean scores of onset-coda words were the lowest among the word types (Table 7).

Table 7. Means and standard deviation of English Spelling Task.

Measures	Word types	Examples	Mean (Max)	Std. Deviation	N
English spelling task	All present	pim	9.11	.937	27
	coda absent	pef	7.37	1.334	27
	onset absent	fip	6.15	1.406	27
	onset-coda absent	fiv	5.67	1.441	27

A one-way repeated measures ANOVA was performed with word types (all presents vs. onset absent vs. coda absent vs. onset-coda absent) as the within-word type factor to verify whether mean differences are statistically significant. A one-way repeated measures ANOVA on the mean differences of each word type was significant, $F(3, 78) = 46.006$, $p<.01$, indicating that the mean differences among word types are significantly large. It is necessary to compare each mean score to obtain more detailed results of mean differences. Table 8 illustrates mean differences, standard error, significance value, and confidence interval for mean differences of English spelling tasks. Significance values show that the comparison of mean scores between all present words and the rest word types (coda absent, onset absent, onset-coda absent) were significant (p<.01). The comparison of mean scores between coda absent and the rest types (all present, onset absent and onset-coda absent) were also significant (p<.01). However, the mean scores of between onset absent and onset-coda absent was not significant (p>.05).

Table 8. Mean Comparisons of English Spelling Task in Each Word Type

Word types		Mean Difference	Std. Error	Sig.	95% Confidence Interval for Difference	
					Lower Bound	Upper Bound
All present	coda absent	1.741*	.285	.000	1.154	2.327
	onset absent	2.693*	.313	.000	2.319	3.607
	Onset-coda absent	3.444*	.367	.000	2.690	4.199
coda absent	All present	-1.741*	.285	.000	-2.327	-1.154
	onset absent	1.222*	.289	.000	.628	1.816
	onset-coda absent	1.704*	.349	.000	.986	2.422
onset absent	All present	-2.963*	.313	.000	-3.607	-2.319
	coda absent	-1.222*	.289	.000	-1.816	-.628
	onset-coda absent	.481	.308	.130	-.152	1.115
onset-coda absent	All present	-3.444*	.367	.000	-4.199	-2.690
	coda absent	-1.704*	.349	.000	-2.422	-.968
	onset absent	-.481	.308	.130	-1.115	.152

Discussion

The main purpose of this study was to investigate English L2 spelling performance with its phoneme and sub-syllabic structure differences. It has been well established that the salient intrasyllabic units are distinguished in English and Korean; onset-rhyme in English is preferred segmentation (Fowler et al., 1993; Treiman et al., 1982; Treiman, 1983) and body-coda in Korean (Kim, 2007; Yoon & Derwing, 2001). There is also support that in English L2 spelling, the English only phoneme and the orthographic differences between L1 and English L2 causes more errors affected by learners' L1 (Bebout, 1985; Fashola et al., 1996; Park, 2012a; 2012b; Sun-Alperin & Wang, 2008). This study was designed to examine which aspect of these two differences (phoneme differences and sub-syllabic structure differences) needs to receive more consideration in English L2 spelling instruction.

Before addressing the primary research questions, it was necessary to establish the consistency of the body-coda preference in Korean children as shown in other studies (Kim, 2007; Yoon, et, al., 2002). The first research question was posed to confirm this issue. The results of the KSSJ task confirmed that the body is a more salient part of the Korean syllable than is the rhyme in predicting global sound similarity judgments (Kim, 2007; 2009; Yoon, et, al., 2002). It showed that the body sharing words were judged more similar (9.4) compared to the rhyme sharing words (6.4). The comparison of mean scores between the rhyme sharing words and body sharing words was also significant (p<.01). Similar results were found in a few studies with slightly different design. First, Kim (2007) found from monolingual Korean-speaking children that a salient sub-syllabic phonological unit in Korean is the body-coda unit. Kim suggested that the body-coda boundary is more salient than the onset-rhyme boundary for Korean children and it is an important predictor of word decoding and spelling in Korean. Kim (2009) also found that children's awareness of larger phonological units such as syllable and body is associated with letter-name knowledge and further mediates literacy acquisition in Korean. Similar results were also found in Yoon and Derwing's (1994) study in which both native English and native Korean speakers judged the sound similarity of the same set of Korean CVC words, and the results were different in both groups. The rhyme was more important than the body for the native English speakers in predicting similarity scores, while the opposite was true for the native Korean speakers. Considering previous research and the results of this study, Korean speakers are likely more sensitive to the body part rather than the rhyme part in recognizing one syllable word.

Concerning the English phoneme, it seems that Korean speakers are more sensitive to body-coda rather than onset-rhyme in English L2 as in their Korean L1. The second research question was posed to investigate the degree of sensitivity to English phonemes among Korean L1 learners of English. In ESSJ tasks, the results showed that interactions among the mean scores of the same words (pin-pin, 11.67), body sharing words (mis-mit, 6.06), rhyme sharing words (mis-tis, 4.06), onset-absent-body sharing words (fis-fin, 5.78), and coda-absent rhyme sharing words (pef-pef, 4) were significant. In particular, a significant interaction between body sharing words and rhyme sharing words is one of the interesting results. It indicates that Korean speakers easily sense body sharing words rather than the rhyme sharing words in English L2, which is similar in their Korean L1 task. For example, they more easily

sense the words, *mis-mit* than the words *mis-tis* as a similar sound. This indicates that L2 learners retain their L1 language specific preference for segmenting the English L2 syllable. This result is not surprising since similar results can be found in other studies (Kim, 2007; Goswami, 1993; Yoon et al., 2002).

The main purpose of this experiment was to examine English L2 learner's sensitivity to sub-syllable structure of monosyllabic words with a phoneme absent in Korean L1 but present in English L2. In ESSJ tasks, the interaction of mean scores between body sharing words (m=6.06, mis-mit) and onset-absent body sharing words (m=5.78, fis-fin) was not significant (p>.05). These results indicate that there would not be much difference in recognizing the words although the body part contains phonemes absent in English L2. However, the interaction between the mean scores of onset-absent body sharing words (m=5.78, fis-fin) and rhyme sharing words (m=4.06, mis-tis) was significant (p<.01) and the interaction between body sharing words (m=6.06, mis-mit) and coda-absent rhyme sharing words (m=4.0, pef-mef) was also significant (p<.01). These results indicate that Korean L1 speakers easily recognize English L2 words sharing body parts even if the body contains English only phonemes, compared to the words sharing rhyme parts with both conditions: all present phonemes in Korean L1 and English L2 and English only phonemes in the coda. As mentioned earlier, a great deal of research (Kim, 2007; Kessler & Treiman, 1997; Treiman et al., 1995; Yoon & Derwing, 2001) supports the idea that native English speakers prefer rhyme parts rather than body parts in segmenting monosyllabic words and vice versa for native Korean speakers. The results of this task added an important finding to the previous studies that whether or not there is an English only phoneme in body parts, Korean L1 speakers more easily recognizes body parts rather than the rhyme parts. That is, body parts for Korean L1 speakers can be unmarked and vice versa for rhyme parts regardless of phoneme congruence between L1 and L2.

The interaction between onset-absent body sharing words (m=5.78, fis-fin) and coda-absent rhyme sharing words (m=4, pef-mef) was significant and the interaction between rhyme sharing words (m=4.06, mis-tis) and coda-absent rhyme sharing words (m=4.0, pef-mef) was not significant (p>.05). This suggests that in the same condition (one absent phoneme in different position), Korean EFL learners still recognize the body sharing words easier than the rhyme sharing words even though there are non-existent phonemes in their L1 in each onset or coda. There is no

difference for Korean EFL learners to recognize English L2 words sharing rhyme parts. There would not be much difference in recognizing the rhyme parts although the rhyme parts contain phonemes absent in Korean L1. Korean EFL learners experienced a similar degree of difficulty in recognizing those two types of words. The mean scores for rhyme sharing words were 4.06 and 4.0 for coda-absent rhyme sharing words. The results of this task added an important finding to previous studies that suggest that, Korean L1 speakers experience similar degrees of difficulty in recognizing rhyme parts whether or not there is L1 absent phoneme in the rhyme part. That is, English L2 rhyme parts for Korean EFL learners can be marked regardless of phoneme congruence between L1 and L2. The results of ESSJ tasks suggest that for Korean EFL learners, the sensitive to sub-syllabic structure of English L2 words relates to the Korean language specific preference, which is body-coda preference rather than the phonemic congruence between L1 and L2.

English spelling task was designed to test whether the spelling accuracy relates to phoneme sensitivity or sub-syllabic structure preference. First, the interaction of mean scores between all present (m = 9.11, pim) and onset absent (m = 6.15, fip) was significant (p<.01), the interaction between all present and coda absent (m=7.37, pef) was significant (p<.01), and the interaction between all present and all absent (m = 5.67, fiv) was also significant (p<.01). These results show that Korean EFL learners were better at spelling words with all present phonemes than the words with either oneset absent, coda absent, or both absent, confirming Park's (2012b) study in which Korean L1 students made more errors on phonologically absent phonemes compared when all phonemes were present. Second, the interaction between coda absent (m=7.37, pef) and onset absent (m=6.15, fip) was significant (p<01), which confirms that Korean EFL learners made more errors on onset-absent words than coda-absent words (Park, 2014). Third, the primary goal of this experiment was test differences between onset absent, coda absent, and all absent. The results show that the interaction between onset absent (m=6.15, fip) and onset-coda absent (m=5.67, fiv) was not significant (p>.05). However, the interaction between coda absent (m=7.37, pef) and onset-coda absent (m=5.67, fiv) was significant (p<.01), indicating that Korean EFL learners had much more difficulty in spelling all absent words than coda absent words, but they did not show statistical mean differences between onset absent and all absent. It is concluded that Korean EFL learners experienced similar degree of difficulty in spelling both all absent words and

onset absent words of English L2. All of the results show that regardless of phonemic congruence, Korean L1 learners of English had difficulty in spelling English onset parts. The results of spelling tasks suggest that for Korean EFL learners, the accuracy of English L2 spelling words relates to the Korean language specific preference, which is body-coda preference rather than the phonemic congruence between L1 and L2.

Conclusion

We have discussed English L2 spelling performance in terms of phoneme and sub-syllabic structure differences. Drawing from previous findings in English L2 spelling performance (Bebout, 1985; Fashola et al., 1996; Park, 2012a; 2012b; Sun-Alperin & Wang, 2008) and sub-syllabic structure differences (Fowler, Treiman, & Gross, 1993; Kim, 2007; Treiman, 1983; Treiman et al.,1982; Yoon & Derwing, 2001), this study tested the degree of two differences to which English L2 spelling performance might be affected and our findings contribute important considerations regarding English L2 spelling instruction. Within the current framework, the two experiments focused on whether Korean EFL learners would sense English L2 phonemes, whether the sensitivity to syllable segment of Korean EFL learners would relate to L1 phonemes (presence vs. absence) or syllable position (onset and coda), and whether Korean EFL learners' spelling accuracy would relate to L1 phonemes (presence vs. absence) or sensitivity to syllable segments. The data demonstrated the following: (a) Korean speakers were more sensitive to body-coda rather than onset-rhyme in English L2; (b) English L2 sub-syllabic structure of Korean speakers related to syllable position rather than the presence or absence of phonemes in their L1; (c) English L2 spelling accuracy of Korean speakers related to sensitivity to syllable segments (sub-syllabic structure preference; body-coda and onset-rhyme), rather than the presence or absence of phonemes in their Ll.

Taken together, these results indicate that in addition to phoneme difference between L1 and L2, sub-syllabic structure differences should receive much more attention in English education in order to enhance L2 spelling performance. This study suggests that English L2 spelling instruction should be emphasized, particularly focusing on linguistic differences such as phonology and sub-syllabic structure differences. Furthermore, since a great deal of research discussed that learners' early spelling competence correlates to later literacy development (Ehri, 1989; 1991;

Treiman, 1998), it is also necessary to consider incorporating spelling instruction as one of the main skills EFL learners need to practice for linguistic performance.

It should be noted that only monosyllabic words were considered for these experiments; however, future studies should investigate errors in multi-syllable words in order to confirm the results of this study. In addition, while this study used data only from English nonsense words to avoid participants' vocabulary knowledge in their spelling and sound judgment task, experiments using real words are needed in order to verify the results of the study.

References

Bebout, L. (1985). An error analysis of misspellings made by learners of English as a first and as a second language. *Journal of Psycholinguistic Research*, 14, 569-593.

Bryant, P. (2002). Children's thoughts about reading and spelling. *Scientific Studies of Reading,* 6(2), 199-216.

Bryant, P., & Bradley, L. (1980). Why children sometimes write words which they do not read. In U. Frith (ed.) *Cognitive Processes in Spelling* (pp.355-370). London: Academic Press.

Cassar, M., & Treiman, R. (1997). The beginnings of orthographic knowledge: Children's knowledge of double letters in words. *Journal of Educational Psychology,* 89(4), 631-644.

Cook, V. (1997). L2 users and English spelling. *Journal of Multilingual and Multicultural Development,* 18(6), 474-489.

Ehri, L. (1989). The development of spelling knowledge and its role in reading acquisition and reading disability. *Journal of Learning Disabilities,* 22(6), 356-365.

Ehri, L. (1991). The development of reading and spelling in children; An overview. In M. Snowling & M. Thomson (Eds.), *Dyslexia: Integrating Theory and Practice* (pp. 63- 79). London, England: Whurr Publishers.

Ehri, L. (1998). Grapheme-phoneme knowledge is essential for learning to read words in English. In J. L. Metsala, & L. C. Ehri (Eds.), *Word recognition in beginning literarcy*, (pp. 3-40). Mahwah, NJ: Lawrence Erlbaum.

Fashola, O., Drum, P., Mayer, R., & Kang, J-S. (1996). A cognitive theory of

orthographic transitioning: Predictable errors in how Spanish-speaking children spell English words. *American Educational Research Journal,* 33, 825-843.

Fowler, C., Treiman, R., & Gross, J. (1993). The structure of English syllables and polysyllables. *Journal of Memory and Language,* 32, 115-140.

Frederickson, N., Frith, U., & Reason, R. (1997). *Phonological Assessment Battery:Standardisation Edition*. Windsor: NFER-Nelson.

Goswami, U. (1993). Toward an interactive analogy model of reading development: Decoding vowel graphemes in beginning reading. *Journal of Experimental Child Psychology,* 56, 2443-2475.

Goswami, U. (1988). Orthographic analogies and reading development. *Quarterly Journal of Experimental Psychology,* 40, 239-268.

Goswami, U. (1998). The role of analogies in the development of word recognition. In J. Metsala & L. Ehri (Eds.)*, Word recognition in beginning literacy.* Mahwah, NJ: Erlbaum.

Goswami, U., & Bryant, P. (1990). *Phonological skills and learning to read*. Hillsdale, NJ, England: Lawrence Erlbaum Associates, Inc.

Greenbert, H., & Jenkins, J. (1964). Studies in the psychological correlates of the Sound system of American English I and II. *Word* 20, 157-177.

Ibrahim, M. (1978). Patterns in spelling errors. *English Language Teaching Journal, 32*, 207-   212.

Katz, L., & Frost, R. (1992). The reading process is different for different orthographies: The orthographic depth hypothesis. In R. Frost & L. Katz (Eds.), *Orthography, phonology, morphology, and meaning* (pp. 67-84). Oxford, English: North-Holland.

Katzir, T., Kim, Y-S., Wolf, M., Kennedy, B., Morris, R., & Lovett, M. (2006). How do you spell read? An investigation of the relationship among orthographic recognition and reading related processes in children with and without dyslexia. *Reading and Writing: An Interdisciplinary Journal,* 19, 845-872.

Kessler, B., & Treiman, R. (1997). Syllable structure and the distribution of phonemes in English syllables. *Journal of Memory and Cognition,* 37, 295-311.

Kessler, B., & Treiman, R. (2003). Is English spelling chaotic? Misconceptions concerning its irregularity. *Reading Psychology,* 24(3-4), 267-289.

Kim, Y-S. (2007). Phonological awareness and literacy skills in Korean: An examination of the unique role of body-coda units. *Applied Linguistics,* 28, 67-93.

Kim, Y-S. (2009). The foundation of literacy skills in Korean: the relationship between letter-name knowledge and phonological awareness and their relative contribution to literacy skills. *Read Writ,* 22, 907-931.

Lundberg, I., Olofsson, A., & Wall, S. (1980). Reading and spelling skills in the first School years predicted from phonemic awareness skills in kindergarten. *Scandinavian Journal of Psychology,* 21, 159-173.

Mann, V., & Liberman, I. (1984). Phonological awareness and verbal short-term memory. *Journal of Learning Disabilities,* 17(10), 592-599.

McBride-Chang, C. (1995). What is phonological awareness? *Journal of Educational Psychology,* 87, 179–192

Park, C. (2012a). Orthographic differences and English L2 spelling errors. *Korean Journal of English Language and Linguistics*, 12(2), 285-302.

Park, C. (2012b). English L2 spelling errors phonologically absent in L2 but present in L2. *The New Korean Association of English Language and Literature,* 54(4), 271-288.

Park, C. (2014). Body-coda vs. onset-rhyme: based on English L2 spelling performance. *Studies in Linguistics*, 32, 125-142.

Shankweiler, D. (1999). Words to meanings. *Scientific Studies of Reading,* 3, 113-127

Share, D. L. & Blum, P. (2005). Syllable splitting in literate and preliterate Hebrew speakers: Onset and rimes or bodies and codas? *Journal of Experimental Child Psychology,* 92, 182-202.

Stahl, S., & Murray, B. (1998). Issues involved in defining phonological awareness and its relation to early reading. In J. L. Metsala, L. C. Ehri, J. L. Metsala, L. C. Ehri (Eds.), *Word recognition in beginning literacy* (pp. 65-87). Mahwah, NJ US: Lawrence Erlbaum Associates Publishers.

Sun-Alperin, M., & Wang, M. (2008). Spanish-speaking children's spelling errors with English vowel sounds that are represented by different graphemes in English and Spanish words. *Contemporary Educational Psychology,* 33, 932-948.

Treiman, R. (1983). The structure of spoken syllable: Evidence from novel word games. *Cognition,* 15, 49-74.

Treiman, R. (1985). Onsets and rimes as units of spoken syllables: Evidence from children. *Journal of Experimental Child Psychology,* 39, 161-181.

Treiman, R. (1992). The role of intrasyllabic units in learning to read and spell. In P. B. Gough, L. C. Ehri & R. Treiman (Eds.), *Reading acquisition* (pp. 65-106). Hillsdale, NJ: Erlbaum.

Treiman, R. (1998). Why spelling? The benefits of incorporating spelling into beginning reading instruction. In J. L. Metsala & L. C. Ehri (Eds.), *Word Recognition in Beginning Literacy* (pp. 289-313). Mahwah, US: Lawrence Erlbaum Associates Publishers.

Treiman, R., Berch, D., & Weatherston, S. (1993). Children's use of phoneme-grapheme correspondences in spelling: Roles of position and stress. *Journal of EducationalPsychology,* 85(3), 466-477.

Trieman, R., & Cassar, M. (1997). Can children and adults focus on sound as opposed to spelling in a phoneme counting task? *Developmental Psychology,* 33, 771-780.

Treiman, R., Mulleniz, J., Bijeljac-Babic, R., & Richmond-Welty, E. (1995). The special role of rimes in the description, use, and acquisition of English orthography. *Journal of Experimental Psychology: General,* 124, 107-136

Treiman, R., Salasoo, A., Slowiaczek, L., & Pisoni, D. (1982). Effects of syllable structure on adults' phoneme monitoring performance. *Research on speech Perception* (pp. 63-81). Progress Report No. 8, Department of Psychology, Indiana University, Bloomington, IN.

Treiman, R., & Zukowski, A. (1991). Levels of phonological awareness. In S. Brady, D. P. Shankweiler, S. Brady, D. P. Shankweiler (Eds.), *Phonological processes in literacy: A tribute to Isabelle Y. Liberman* (pp. 67-83). Hillsdale, NJ England: Lawrence Erlbaum Associates, Inc.

Verhoeven, C. K., Boechler, P. M., & Steffler, D. J. (1999). Phonological and orthographic influences on children's vowel spelling. *Scientific Studies of Reading,* 3(4), 363-379.

Wang, M., & Geva, E. (2003). Spelling acquisition of novel English phonemes in Chinese children. *Reading and Writing,* 16(4), 325-348.

Yeong, S. & Liow, S. (2010). Phonemic representation and early spelling errors in bilingual children. *Scientific Studies of Reading,* 14(5), 387-406

Yoon, H-K., Bogler, D., Kwon, O-S., & Perfetti, C. (2002). Subsyllabic units in

reading. In L. Verhoeven, C. Elbro, & P. Reitsma (Eds.). *Precursors of Functional Literacy* (Vol.11, pp. 139-163). Amsterdam: Benjamins.

Yoon, Y- B., & Derwing, B. (1994). Sound similarity judgments for Korean words by Korean and English speakers. In P. Koskinen (Ed.), *Proceedings of the 1994 Annual Conference of the Canadian Linguistic Association* (pp. 655-657). Toronto: Toronto Working papers in Linguistics, Linguistic Graduate Course Union.

Yoon, Y-B., & Derwing, B. (2001). A language without a rhyme: Syllable structure experiments in Korean. *Canadian Journal of Linguistics,* 46, 187-237.

Ziegler, J., & Goswami, U. (2005). Reading acquisition, developmental dyslexia and skilled reading across languages: A psychological grain size theory. *Psychological Bulletin,* 131, 3-29.

Arabic EFL University Students' Use of Source Texts in Argumentative Writing

Jingjing Qin
Department of English and Writing Studies, Zayed University
Abu Dhabi, United Arab Emirates

Bioprofile:
Jingjing Qin is an assistant professor in the Department of English and Writing Studies at Zayed University, UAE. Her recent publications appear in *System*, *Language Teaching Research*, and *TESOL publications*. She received her Ph.D. in applied linguistics from Northern Arizona University, USA. She taught in the US, China, and Turkey. Email: jingjingqin79@gmail.com

Abstract

Writing from sources is a common requirement in academic writing in the higher education contexts. Much research exists on the quality of source use in student writing, which often deals with textual appropriation (i.e., plagiarism). Very few studies, however, focus on students' selection of source text information to include in their essays. To this end, this study examined 20 Arabic EFL university students' selection of source text information and also explored its relationship with the quality of their argumentative essays written based on two source texts. The essays were analyzed sentence by sentence regarding the use of information. If the information was taken from the reading texts, then the function of the information was examined in relation to the argumentative essay, based on the adapted Toulmin model (1958, 2003) (i.e., claim, data, warrant, counterargument claim, counterargument data, rebuttal claim, and rebuttal data). In addition, the essays were scored using an analytic rubric, and students' inclusion of source text information was correlated with their overall essay quality. It was found that the students used the source text information to various degrees, and no relationship was observed between the use of source text information and the overall quality of the essays.

Keywords: Argumentative writing; Use of source texts; The Toulmin model; Arabic-speaking EFL university students

Department of English and Writing Studies,
Zayed University P.O. Box 144534, Abu Dhabi,
United Arab Emirates.

Introduction

The reading-to-write tasks are undeniably prevalent writing assignments in English-medium universities across disciplines or within specific disciplines. (e.g., Bridgeman & Carlson, 1984; Carson 2001; Hale, Taylor, Bridgeman, & Carson et al., 1996; Horowitz, 1986; Johns, 1993; Leki & Carson, 1994, 1997; Zhu, 2004). As early as 1980s, Bridgeman and Carlson (1983) found that argument writing based on several sources, or writing from sources was among the most frequently assigned writing tasks required of graduate and undergraduate international students in the United States. The importance of these integrated writing tasks has also been reflected in the adoption of these tasks in many internationally well-recognized standardized tests, such as the Test of English as a Foreign Language (TOEFL), the International English Language Testing System (IELTS), the Canadian Academic English Language (CAEL) and other institutional proficiency tests (Weigle & Parker, 2012).

A valid and crucial question in relation to the reading-to-write task is how students use source texts in their writing. Much of the recent literature in this line of inquiry has examined the practices of textual borrowing by both L1 and L2 university students (Cumming, Kantor, Baba, Eouanzoui, Erdosy, & James, 2006; Gebril & Plakans, 2009; Plakans & Gebril, 2012; Shi, 2004; Weigle & Parker, 2012). They generally categorized the textual borrowing into two broad types: exact copying or called verbatim, and paraphrase with different degrees of alternation of the words from the source texts. The issue of what is considered an appropriate textual borrowing without committing plagiarism has always been controversial, thus receiving much recent research interest in the field of EAP (e.g., see recent special issues on textual borrowing in *Journal of Second Language Writing* and *Journal of English for Academic Purposes*). As seen, most of the research investigating the use of source texts so far has focused on the formal mechanism of using the textual information in an integrated task (i.e., exact copying, paraphrasing, summarizing, or quoting). Very little has been done yet to investigate what ideas from source texts have been used in an integrated reading-to-write task by student writers. Since in this type of integrated task, writers are responsible for the content of readings and expected to use the readings appropriately in the writing task, it implies that to understand how student writers use the texts in terms of content or ideas is imperative for expanding this line of research on reading-to-write task (Grabe, 2003). The present

study examines how Arabic EFL university writers use ideas from source texts to write their argumentative essays.

A related and equally important question which deserves further research is how the use of source texts contributes to the quality of integrated writing task. The answer to this question can help shed light on another aspect of the complex, multifaceted integrated task. In L1 contexts, a series of studies conducted by Spivey and her colleagues (Spivey, 1983; Spivey & King, 1989) has found that more effective readers selected significantly more important content from source texts and produced more unified and connected synthesis papers than less effective readers. Little is known, however, about writing from sources with the rhetorical purpose of presenting an argument. Plakans and Gebril (2013) call for further research to understand distinct levels of source use and what key descriptors define these in terms of integrated writing assessment. The study can shed light on this less examined area of source use by having a closer look at how a group of L2 university writers use source texts in an argumentative writing, specifically, how they employ source texts for the purpose of arguing their position in a controversial issue.

Literature Review

The importance of integrating information from outside sources in successful academic writing has been documented in Campbell (1990), who argues that "even the most original academic paper integrates facts, ideas, concepts, and theories from other sources by means of quotations, paraphrase, summaries, and brief references" (p. 211). It should be mentioned that according to Guo (2011), there are three different types of reading-to-write tasks depending on the expected role of source texts in the writing task that are commonly used in EAP tests. Firstly, text-based writing tasks requires using reading texts as the primary source of content for the written response as in summarizing; secondly, situation-based writing tasks refers to more casual tasks writing emails or letters based on certain prompts such as conversations or written notes; thirdly, thematically related writing tasks are often argumentative tasks in which writers are asked to present their own point of view on a controversial issue, supported by information from source texts as well as their own experience and knowledge. As seen, the third type of writing tasks involves more in-depth processing and integration of reading texts.

The earlier studies on the use of source texts in reading-to-write tasks mostly focused on summary tasks and revealed that students employed a limited inventory of strategies in using information from source readings, such as near copy, direct quotations, and paraphrase; more advanced strategies valued in the English-medium academic communities, such as summary and integration, had apparently not been fully grasped. For example, Campbell (1990) compared 20 ESL and 10 native English speakers' use of source information and concluded that copying was a major strategy used by all the students, who demonstrated little meta-awareness of the need for acknowledging outside sources. Interestingly, the ESL students in the study credited the source text or the author in their quotations and paraphrases more often than the native English speakers. Studying the summary skills of 70 Korean EFL freshmen, Kim (2001) found that the common strategy of summarizing used by the students remained at the low-level deletion technique, although they seemed to be able to select the most important information from the source text. She also found that text difficulties affected students' summarization skill in the following respects: a larger proportion of idea units were included in the summary of the easier text, with less use of selection and transformation rules and more inaccurate ideas units included in the summary of the more difficult text.

The more recent studies on textual borrowing have demonstrated more interest in the third integrated writing task, that is, thematically-related tasks. Shi (2004) investigated whether L1 and L2 university students used source texts differently in summary and opinion tasks. Forty-eight Chinese EFL university students and 39 native English-speaking university students were asked to write either a summary or an opinion essay after reading two preselected expository texts. It was found that summary tasks prompted both L1 and L2 writers to use more words from source texts or select more information from source texts than opinion essays; moreover, the Chinese students borrowed significantly more words from the source texts than their native English-speaking counterparts. Shi calls for more studies to explore writers' use of source texts across variables such as language background, language proficiency, and writing task type. Adopting Shi's (2004) coding categories, Weigle and Parker (2012) analyzed the use of source texts in 63 argumentative papers written by L2 undergraduate and graduate students on two different topics. The results showed that most of the students did not have problems with textual borrowing and

the cases of textual borrowing were restricted to short excerpts; textual borrowing practices did not differ much across topics, academic levels, and proficiency levels.

Very few studies , so far, have examined how L2 university students use source texts in terms of content and ideas in integrated writing task. Although Plakans and Gebril (2012) examined how L2 university students used source texts regarding the content both through think-aloud with a small group of students and then a questionnaire with a large group of students, the findings were general, such as gaining ideas about the topic, using as evidence, and also adopting the organizational pattern.

Furthermore, till now, relatively little research have investigated the relationship between the use of source texts in terms of content and the quality of integrated writing. The few existing studies examining the relationship between patterns of textual borrowing and quality of writing indicated that less effective papers tended to borrow heavily from source texts. For example, based on a few students' papers, Weigle and Parker (2012) seemed to suggest that lower-score writing used longer quotes from source texts than higher-score writing. Similarly, Cumming, Kantor, Baba, Eouanzoui, Erdosy, and James (2006) observed a tendency for the most proficient writers to summarize or synthesize ideas coherently from source materials, and for middle-range writers to paraphrase or employ verbatim piecemeal phrases from source materials. The least proficient writers were not able to comprehend the source materials sufficiently well to be able to summarize or to paraphrase competently. These findings are also consistent with Plakans and Gebril (2013), who observed that writers with high scores were able to select important ideas from the source texts, whereas writers with low scores depended heavily on reading texts with many instances of direct copying of words and phrases.

To extend this line of research on the use of source texts, the present study was intended to conduct a detailed written discourse analysis of the use of source texts in students' argumentative papers in terms of the function that particular use of source texts serve in an argumentative paper. Furthermore, the study was intended to correlate the use of source texts regarding the functions and the quality of argumentative writing. Specifically, the present study addresses the following two important issues related to the use of sources texts in an integrated writing task:

1.   How do L2 writers use source texts in an argumentative task in terms of functions?

2.   Is there a relationship between the use of source texts and the quality of integrated writing task?

Methodology

Participants

All participants were enrolled in an academic writing class where essays were collected for the present study and had completed two academic writing classes prior to this one. They were mostly second-year university students around the similar age, 18-21 years old, majoring in a variety of disciplines ranging from natural sciences to social sciences. In their previous English academic writing instruction, they had learned to write in different genres, such as narrative, rhetorical analysis, and persuasive essays.

Writing Task

Students were asked to read two excerpts on the controversial issue of whether *the future of English as the global language will be assured or not*. The two excerpts present opposing views on the issue. After reading the source texts, they were asked to write a clear and organized essay supporting their opinion, with at least two citations from the sources. The reading and writing tasks lasted 80 minutes in total.

Data Analysis

The analysis of students' papers took several steps. The first step was to ascertain the places where students borrowed information from reading texts. Then, the second step was to categorize the functions that each borrowed string served in the argumentative paper. To identify textual borrowing was not a straightforward and clear-cut task, as evidenced in previous studies on this issue. After examining the coding categories used in some previous studies (e.g., Cumming et al. 2006; Shi, 2004), several decisions were made. First, students' papers were analyzed based on the unit of orthographic sentence (Keck, 2006; Shi, 2004). Second, the use of source information was put into three categories: (a) verbatim copy of source texts, (b) partial paraphrase from source texts, and (c) full paraphrase from source texts. The first category was defined as exact copies from source texts with or without the use of quotation marks. Sentences in this category use strings of three words or more that are extracted from source texts word-for-word. The second category is called partial

paraphrase. Partial paraphrases are ideas drawn from source texts with slight changes of wording from source texts. Unlike the verbatim category, this category was slightly more ambiguous, thus demanding more subjective judgment during the coding process. The identification of information drawn from source texts was mostly based on the use of some key content words found in the paraphrases, which were also observed in source texts, or in some cases, synonyms of those key content words were used. Different from Shi (2004) and Weigle and Parker (2012) where the complete paraphrases were ignored, the present study also analyzed this kind of paraphrases in which no words were directly extracted from the source texts. The rationale for including this kind of paraphrase was due to the focus of the study, which examined the use of source texts in terms of ideas and content. During the analysis, it was found that such complete paraphrases, although rare in general, did exist. Since the focus of the study was on the use of source texts in terms of the content, not the textual borrowing practices as in Shi (2004) and Weigle and Parker (2012), it is sufficient to analyze the use of source texts into these three general categories.

Then after the textual borrowing was identified, the next step was to further categorize these textual borrowings in terms of the functions they served in an argumentative paper. The present study adapted the Toulmin model of argument structure (1958, 2003) as a framework for analysis, which consists of seven elements: *claim*, *data*, *warrant*, *counterargument claim*, *counterargument data, rebuttal claim*, and *rebuttal data*. In other words, each textual borrowing was put into these different categories of functions. During the analysis, two more categories of functions emerged, that is *introduction* and *thesis*. In other words, some papers used information from source texts in the introduction and thesis parts of the papers.

The second research question was concerned with the relationship between the use of source texts and the quality of integrated writing task. To assess the overall quality of English argumentative papers, a holistic rubric was needed. To this end, a 5-scale scoring rubric was developed based on the rubrics used by McCann (1989) and Nussbaum and Kardash (2005). This scoring rubric drew on the 6-scale rubric used by McCann (1989) for its clear description of organization and language use, and the 7-scale rubric used by Nussbaum and Kardash (2005) for its thorough delineation of the effectiveness of argument. This holistic rubric included the following three dimensions: the overall effectiveness of argument including the presence or absence of the possible opposing views (also called counterarguments), overall organization,

and language use in general. This rubric had been validated and used in previous studies (e.g., Qin & Karabacak, 2010). It should be mentioned that the amount of information drawn from the source texts had not been used as criteria of the rubric, because of the nature of this kind of integrated writing task, which allows using writers' own personal experiences and knowledge. In fact, this study was intended to answer the question how much is appropriate to draw textual information from source texts to write a good argumentative paper.

Results and Discussions

RQ 1. Information Drawn from Source Texts

The first research question investigated how Arabic EFL university students used source texts in their argumentative writing regarding the functions. It was found that in general, the students employed the source text information to various degrees. Three students out of 20 relied heavily on source texts by just summarizing the texts without presenting his or her own point of view. Interestingly, one student out of 20 copied heavily from only one of the source texts which is consistent with her point of view. Table 1 presents the descriptive statistics for each functional category. For example, the source text information has been used in the introduction part nine times in total in all the 20 essays, which means on average 0.45 times per essay.

Table 1. Descriptives of the functions of use of source texts

Functional Category	Sum	Mean
Introduction	9	0.45
Thesis	3	0.13
Claim	23	1.13
Data	24	1.2
Warrant	5	0.25
Counterargument claim	11	0.55
Counterargument data	5	0.25
Rebuttal claim	2	0.1
Rebuttal data	1	0.05

It can be seen from the above Table, mostly students used source text information for the functions of *claims* and *data*. Half of the students were able to take the opposing view from the other source text and present it as *counterargument*

claim, and fewer students even went further to use information from source texts to use it as *counterargument data*. Seven students used information from the source texts for the *introduction* part of their essays. Very rarely did students present the *rebuttal claim* and *rebuttal data*.

On a continuum of source text use, some of the participants relied on the source texts too heavily and simply summarized the main ideas of the source texts without realizing that they were expected to use source texts to help support their own arguments. At the other end of the continuum, some participants simply ignored the source texts without using them at all. These findings may have occurred because these participants had not received any training in how to use source texts for their writing purposes.

Regarding the use of source texts, it was found that most of the students did not know how to use source texts appropriately for the writing purpose. There was a continuum from using almost no source text information, despite the requirement to do so, to relying too much on the source texts. This was due to the students' lack of experience in writing from sources. It is suggested that students be taught how to use source texts both in terms of content and also style. Specifically, students should be informed that they can use information from sources for their writing purposes in a variety of ways, such as using source information as evidence to support their point of view, or as possible opposing views contrary to their own point of view, or as background information on the related topic and so on.

RQ2. The relationship between the use of source texts and the quality of argumentative papers

Table 2. Descriptives of quality of argumentative papers

Student	N	Mean	SD
Arabic-speaking EFL students	20	3.12	.993

Table 2 presents the quality of the 20 Arabic EFL university students' essays with a Mean of 3.12 out of 5. To answer the second research question, Pearson product-moment correlation was computed for each functional category, and no statistically significant relationship was found for any of the functions. It suggests that

the more use of the source texts regarding any of the functions does not mean the better quality of the essays. However, a quick examination of a good argumentative paper (see the excerpt below) with a score of 4.5 reveals that it used information from source texts to support her point of view. Some other good-quality papers used source text information to introduce opposing views and refute these opposing views. In addition, the writer used her own words to paraphrase the ideas from source texts or attribute the source of information if the source texts were used. The writer also expanded the ideas gained from the source texts and add her own interpretation.

> The English as a lingua franca is continuing to grow because of many reasons. First, English is one of the easiest languages and anyone could be able to learn it. English unlike other competitive languages such as Chinese is easier to read, write and speak. Second, because of globalization, most countries had to assign English as one of the official languages nationally. For example, The governments in the GCCs assigned English writings in sign roads in addition to Arabic because those countries embraces not only Arabs but many other different nationalities. Moreover, English became so popular and is the most practical way to deal with foreign people in businesses. Third, universities who want to be competing to place a high world rank should apply teaching all courses with English. <u>According to Dyer (2012), one of the best top universities in Italy "Politecnico Di Milano" will apply English in all courses and they have to do this to be one of the best universities globally</u>. (Data) This step will give the graduates high chances to work in many different countries and have more job opportunities in the global economy.

In contrast, it was found that the ineffective argumentative papers were ones which relied too much on the source texts where heavy verbatim use of source texts were found or which failed to use any information from source texts. In some cases, every piece of information is borrowed from source texts without much of the writer's own ideas. It should be noted all the results are just suggestive, as the amount of data is too small to draw any generalization.

In general, it seems that greater use of source texts created the impression of more reliance on source texts, and consequently failed to show a writer's own efforts in arguing for his or her point of view. On the other hand, because the writing task required the students to use information from source texts in their writing, it would be appropriate to extract some source text information for their writing purposes, such as using research studies in source texts as evidence to support their points of view or using the opposing views and data presented in the other source text as

counterargument claim and counterargument data. Therefore, it can be argued that the more effective papers demonstrated an optimal level of the use of source texts.

Conclusion

This study has addressed the gap in the line of research on writing from sources in L2 contexts, that is how L2 university students used source text information in argumentative writing. As indicated in previous research, a great number of L2 university students in English-medium universities have felt the lack of preparation for tasks based on writing from sources (Carson, 2001; Leki & Carson, 1994, 1997), which constitute a crucial part of academic studies in English-medium universities. The study, in general, has confirmed the previous finding that Arabic EFL university students are not proficient in using source texts in that they failed to use the source texts information properly. It is suggested that explicit instruction should be provided to L2 writers on how to use outside sources in their writing, as reading-to-write tasks are inherently challenging.

References

Bridgeman, B., & Carlson, S. (1984). Survey of academic writing tasks. *Written Communication, 1,* 247–280.

Campbell, C. (1990). Working with others' words: Using background reading text in academic compositions. In B. Kroll (Ed.), *Second language writing: Research insights for the classroom* (pp. 211-230). New York: Cambridge University Press.

Carson, J. (2001). A task analysis of reading and writing in academic contexts. In D. Belcher & A. Hirvela (Eds.), *Linking literacies: Perspectives on L2 reading-writing connection* (pp. 48–83). Ann Arbor, MI: The University of Michigan Press.

Cumming, A., Kantor, R., Baba, K., Eouanzoui, K., Erdosy, U., & James, M. (2006). *Analysis of discourse features and verification of scoring levels for independents and integrated prototype tasks for the new TOEFL.* TOEFL Monograph Series, Report No. 30. Princeton, NJ: Educational Testing Service.

Gebril, A., & Plakans, L. (2009). Investigating Source Use, Discourse Features, and Process in Integrated Writing Tests. *Spaan Fellow Working Papers in Second or Foreign Language Assessment, 7,* 47-84.

Grabe, W. (2003). Reading and writing relations: Second language perspectives on research and practice. In B. Kroll (Ed.), *Exploring the dynamics of second language writing* (pp. 242-262). New York: Cambridge University Press.

Guo, L. (2011). *Product and process in TOEFL iBT independent and integrated writing tasks: An investigation of construct validity.* Unpublished doctoral dissertation, Georgia State University.

Hale, G., Taylor, C., Bridgeman, B., Carson, J., Kroll, B., & Kantor, R. (1996). A *study of writing tasks assigned in academic degree programs.* Princeton, NJ: Educational Testing Service.

Horowitz, D. (1986). What professors actually want: Academic tasks for the ESL classroom. *TESOL Quarterly, 20,* 445-462.

Johns, A. (1993). Reading and writing tasks in English for academic purposes classes: Products, processes and resources. In J. Carson & I. Leki (Eds.), *Reading in the composition classroom* (pp. 274-289). New York: Newbury House.

Keck, C. (2006). The use of paraphrase in summary writing: A comparison of L1 and L2 writers. *Journal of Second Language Writing, 15,* 261-278.

Kim, S. (2001). Characteristics of EFL readers' summary writing: A study with Korean university students. *Foreign Language Annals, 34,* 569-581.

Leki, I., & Carson, J. (1994). Students' perceptions of EAP writing instruction and writing needs across the disciplines. *TESOL Quarterly, 28*, 81-101.

Leki, I., & Carson, J. (1997). "Completely different worlds": EAP and the writing experiences of ESL students in university courses. *TESOL Quarterly, 31,* 39–69.

McCann, T. (1989). Student argumentative writing knowledge and ability at three grade levels. *Research in the Teaching of English*, *23*, 62-76.

Nussbaum, E. M., & Kardash, C. A. M. (2005). The effects of goal instructions and text on the generation of counterarguments during writing. *Journal of Educational Psychology, 97,* 157-169.

Plakans, L. & Gebril, A. (2012). A close investigation of source use in integrated writing tasks. *Assessing Writing Journal, 17* (1), 18-34.

Plakans, L. & Gebril, A. (2013). Using multiple texts in an integrated writing assessment: Source text use as a predictor of score. *Journal of Second Language Writing, 22,* 217-230.

Qin, J., & Karabacak, E. (2010). The analysis of Toulmin elements in Chinese EFL university argumentative writing. *System, 38,* 444-456.

Shi, L. (2004). Textual borrowing in second-language writing. *Written Communication, 21,* 171-200.

Spivey, N. N. (1983). *Discourse synthesis: Constructing texts in reading and writing.* Unpublished doctoral dissertation, University of Texas, Austin.

Spivey, N. N., & King, J. (1989). Readers as writers composing from sources. *Reading Research Quarterly, 24,* 7-26.

Toulmin, S. (1958). *The uses of argument.* Cambridge: Cambridge University Press.

Toulmin, S. (2003). *The uses of argument* (updated ed.). Cambridge: Cambridge University Press.

Weigle, S. C., & Parker, K. (2012). Source text borrowing in an integrated reading/writing assessment. *Journal of Second Language Writing, 21*, 118-133.

Zhu, W. (2004). Writing in business courses: An analysis of assignments types, their characteristics, and required skills. *English for Specific Purposes, 23,* 111–135.

Comparison of Code-Switching Behaviors of Male and Female Classroom Teachers across Age Groups

Christina D. Vicencio
Bulacan State University, Phillippines

Christopher S. Vicencio
Bulacan State University, Phillippines

Bioprofiles:

Christina D. Vicencio, M.A.E is a college instructor in Bulacan State University. She finished her Master of Arts in Education major in Language Education and is currently pursuing her Doctor of Philosophy in English Language Studies at Bulacan State University, City of Malolos, Province of Bulacan. Email: athaliah1017@yahoo.com

Christopher S. Vicencio, Ed.D finished his Master of Arts in Educational Management at Northern Quezon Colleges Incorporation in Quezon Province and his Doctor of Education in Educational Management at Eulogio "Amang" Rodriguez Institute of Science and Technology in Sampaloc, Manila. He is an Asst. Professor IV in Bulacan State University. Email: allstar_252002@yahoo.com

Abstract

The purpose of this study is to investigate the comparison of code-switching behaviors of male and female classroom teachers across different age groups, since code-switching is a phenomenon in bilingual societies. The participants are 18 college teachers in a satellite campus of a State University in Region III teaching different subjects using English as a medium of language. The participants are grouped according to sex and three age groups. As a descriptive study, it focuses on the occurrence of code-switching of the classroom teachers. Specifically, this study traces the curve of code-switching functions of the respondents using a 2x3 matrix involving sex and three age groups.

Keywords: code-switching, age, sex, behavior

Bulacan State University, City of Malolos,
Philippines

Introduction

People who have learned two languages demonstrate an interesting phenomenon known as "code-switching" by mixing words or phrases from the two tongues together during the course of speech or writing. Code-switching defines as an alternation of words and phrases between two languages or dialects. This usually occurs between people who share those particular languages. It is a phenomenon that exists in bilingual societies where people have the opportunity to use two or more languages to communicate. It is regarded as a communicative phenomenon of constantly switching between two languages in a bilingual's speech repertoire. Chloros (1998), describe Code-switching as a means of communication which involves a speaker alternating between one language and another in communicative events. Crystal (2003) views the concept as a code-switch when an individual who is bilingual alternates between two languages during his or her speech engagement with another bilingual. All these definitions infer that the speaker in a code-switching situation must have communicative competency in two languages for them to be able to switch from one language to another; it may be the mother tongue and a second language. Code-switching performs various functions in its naturally occurring context.

The researchers' aim in this work is to understand the differences of male and female classroom teachers across age groups in code-switching inside their classrooms since classroom life is socially constructed and it also acknowledges the dominance of social interactions in all that occurs in classrooms. Teachers interact with their students as members of the class and construct different functions of code-switching as means of classroom communication. Learning in classroom is not merely a process of learning a language, but it is also a reflection of language through which teachers determine their code to cater for their students' need by applying different codes of interaction in the domain of EFL classrooms (Gulzar, 2013).

This paper shows that it is the content knowledge or language focus of the lesson that influences the discourse functions of choice of the official language of the classroom or other languages. The findings point to selected code-switching functions being useful for teachers to achieve teaching goals in classrooms where lack of proficiency in the instructional language might compromise learning.

Objectives of the Study

The general objective of the study is the comparison of code-switching behaviors of male and female classroom teachers across age groups. The specific objectives of the study are the following: (1) To describe the personal and professional characteristics of the classroom teachers; (2) To describe the sex difference in code-switching between Filipino and English in teaching inside the classroom; (3) To determine the frequency and percentage on teachers consciousness of Filipino switching in the class; (4) To define the teachers' perception on code-switching to Filipino as an efficient strategy of learning and teaching; and (5) To determine the functions that influence teachers' switching to the first language.

Review of Related Literature

Code-switching can be used as part of an actual teaching methodology. When a teacher is aware of the language of the students, the classroom is a setting that potentially uses code-switching. Code-switching is necessary in the classroom if the teachers and students share the same language and should be regarded as a natural part of a bilingual's behavior. Systematic investigation of learners' code-switching is undertaken by Arnfast and Jorgensen (2003), which show how code-switching may develop into a bilingual competence among learners within the first year of intensive training. Bilingual teachers use two languages to teach the academic content. Within the context of lessons, they switch between the languages in at least three ways: (a) spontaneously, (b) directly, (c) intentionally. Teachers may decide immediately when L1 should be used and when a switch to L2 is appropriate in order to enable comprehension and meaningful involvement of the students (Cook, 2001).

In the study conducted by Greggio and Gil (2007), it is evident that the teachers code switch in the beginner group in four different occasions such as [1] explaining grammar, [2] giving instructions, [3] monitoring/assisting the students [4] when correcting activities and interestingly to attract learners' attention. In most cases, the teachers claimed that they need to code switch in order to "clarify words, expressions, structures and rules of utterances" (Greggio & Gil, 2007:376). Thus, from the above functions, code-switching does play an important role in ESL classrooms as it helps learners to better understand the target language they are learning.

Schweers (1999) conducted a research into the field of code-switching, and found that a high percentage (about 88.7%) of the participants felt that the use of mother tongue in their English classes is effective.

Cipriani (2001) worked on real participation strategies in a beginner found that use of code-switching was a good way to foster oral participation among the learners and teacher. Her study also showed that the teachers' use of code–switching strategy to clarify words in communicative tasks make a good atmosphere among learners to speak English.

Cook (2001) believed that to permit students use their mother tongue can be a factor to help students improve the way they learn a second language and it provides students with opportunities to say what they tend to say in classrooms. Levine, (2003) also argued that if the use of L1 was in the right way it serves an effective functions in L2 classrooms, such as, vocabulary, grammar, and writing assignments.

Storch and Wiggleworth (2003) conducted a study to examine the use of L1 in writing assignments and problem-solving among adult L2 language learners. Furthermore, the results identifies the following uses of mother tongue in the classroom: classroom management, language analysis and presenting rules that govern grammar, giving instructions or prompts, explaining errors, and checking comprehension. As this brief literature review shows, code-switching carries out important functions in the communicative activities in L2 classrooms.

The present study would illustrate the use of code-switching in interactive exchanges among learners and male and female teachers across age groups in bilingual classrooms.

Methods of the Study

To have a general and genuine reflection of teachers' code-switching in classroom used across age groups, the researchers used qualitative and quantitative research methods, consisting of questionnaire and classroom recordings. The questionnaire was adopted from the studies of Duff and Polio (2003). It has three parts: personal background, guidance and questions. In the first part of personal background, the subjects were asked to give their personal information. The second part pertains to sex differences in code-switching to Filipino, and the third part was questions about how code-switching influenced the teachers and its functions.

Classroom recordings were also used to collect the data on functions of code-switching to Filipino and enrich the data from the questionnaires. Some of the materials recorded have been transcribed into written forms which are used as examples to support the authors' analysis.

Participants of the Study

The study was conducted in one of the satellite campuses in a state university in Region III. The participants were 18 college teachers teaching different subjects using English as a medium of instructions.

Results and Discussions

The following discussions presents, analyzes and interprets data gathered from 18 college teachers in a satellite campus of a state university in Region III teaching different subjects using English as a medium of instructions. The participants are grouped according to sex and three age groups. The researchers gathered the teachers' responses through survey questionnaires and tape recording.

Section 1- Personal and Professional Characteristics of the Teachers

This section describes the personal and professional characteristics of the teachers in terms of their sex, age and professional rank.

Table 1. Distribution of Teachers in terms of Personal Demographics

A. Sex	Frequency	Percentage
1. Male	9	50%
2. Female	9	50%
Total	**18**	**100**
B. Age		
1. 30 and below	6	33.33%
2. 31-40	6	33.33%
3. 41 and above	6	33.33%
Total	**18**	**100**
C. Professional Rank		
1. Part- time instructor	14	77.77%
2. Instructor II	1	5.56
3. Asst. Professor	2	11.11%

4. Assoc. Professor	1	5.56%
Total	**18**	**100**

Table 1 shows frequency and percentage of the characteristics of the sample in terms of their personal details like: sex, age, and professional rank. It could be gleaned that the respondents in terms of sex were equally distributed with nine (9) or 50%. The three age groups were also equally distributed with six (6) or 33.33%

In terms of their academic rank, most of the teachers were part-time instructors (14, or 77.77%), followed by assistant professors (2, or 11.11%). There were one instructor and one associate professor (5.56% each).

Section 2- Sex Differences in Code-switching to Filipino

This section refers to the questions regarding the alternate use of the first language and the target language.

Table 2. Frequency and Percentage on how often do teachers use Filipino in the class

How often do you use Filipino in the class?	30 and below		31-40		41 and above	
	Male	**Female**	**Male**	**Female**	**Male**	**Female**
Always	0 (0%)	3 (100%)	1 (33.33%)	1 (33.33%)	1 (33.33%)	1 (33.33%)
Sometimes	2 (66.67%)	0 (0%)	1 (33.33%)	2 (66.67%)	2 (66.67%)	1 (33.33%)
Occasionally	1 (33.33%)	0 (0%)	1 (33.33%)	0 (0%)	0 (0%)	1 (33.33%)
Never	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)
Percentage	**100%**	**100%**	**100%**	**100%**	**100%**	**100%**

Table 2 represents how often teachers use Filipino in the class. It shows that in the "always" category, none of the male aged 30 and below answered always or 0%, while female teachers under the same age bracket have 100% or they always code-switch due to many reasons. Male and female aged 31-40 and male and female aged 41 and above have 1 or 33.33% answered that they always code-switch to Filipino. For the "sometimes" category, there are 2 or 66.67% male aged 30 and below answered they occasionally code-switched to Filipino, while male aged 31-40 has 1 or 33.33% as same to female aged 41 and above. And female aged 31-40 and male aged 41 and above have 2 or 66.67%. For "occasionally" category male aged 30 and below

and 31-40 has 1 or 33.33% same with female aged 41 and above. None of them answered they had never code- switched in the class.

Table 3. Frequency and Percentage on Teachers Consciousness of Filipino Switching in the Class

Are you conscious of Filipino Switching?	30 and below		31-40		41 and above	
	Male	**Female**	**Male**	**Female**	**Male**	**Female**
Yes, Always	2 (66.67%)	3 (100%)	2 (6.67%)	3 (100%)	2 (66.67%)	1 (33.33%)
Yes, Sometimes	1 (33.33%)	0 (0%)	1 (33.33%)	0 (0%)	1 (33.33%)	2 (66.67%)
Yes, Occasionally	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)
No, Never	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)
Percentage	**100%**	**100%**	**100%**	**100%**	**100%**	**100%**

Table 3 indicates that female teachers aged 30 and below and 31-40 have 3 or 100% are always conscious of Filipino switching, unlike female teachers aged 41 and above, 2 or 66.67% answered always and 1 or 33.33% answered he/she sometimes conscious of code-switching to Filipino. While all male teachers answered 2 or 66.67% for always conscious in code-switching to Filipino, and 1 or 33.33% answered they are sometimes conscious of Filipino switching which indicates that teachers are always or sometimes conscious of Filipino switching.

Table 4. Frequency and Percentage of Teachers' Perception on Code-switching to Filipino as an efficient strategy of learning and teaching.

Do you think code-switch to Filipino is an efficient strategy of learning and teaching?	30 and below		31-40		41 and above	
	Male	**Female**	**Male**	**Female**	**Male**	**Female**
Yes	2 (66.67%)	3 (100%)	3 (100%)	3 (100%)	1 (33.33%)	1 (33.33%)
No	1 (33.33%)	0 (0%)	0 (0%)	0 (0%)	2 (66.67%)	2 (66.67%)
Percentage	**100%**	**100%**	**100%**	**100%**	**100%**	**100%**

Table 4 shows the frequency and percentage of teachers' perception on code-switching to Filipino as an efficient strategy of learning and teaching. Most of them, male and female aged 31-40 with 3 (or 100%) and female aged 30 and below,

answered "yes" showing that code-switching to Filipino is an efficient strategy of learning and teaching. While 4 male and female aged 41 and above answered "no" which means code-switching to Filipino is not an efficient strategy of learning and teaching.

Section 3- Functions that Influence Teachers' Switching to Filipino

This section refers to the questions regarding the functions that influence teacher's switching to Filipino.

Table 5. Frequency of the Functions of Code-Switching to Filipino

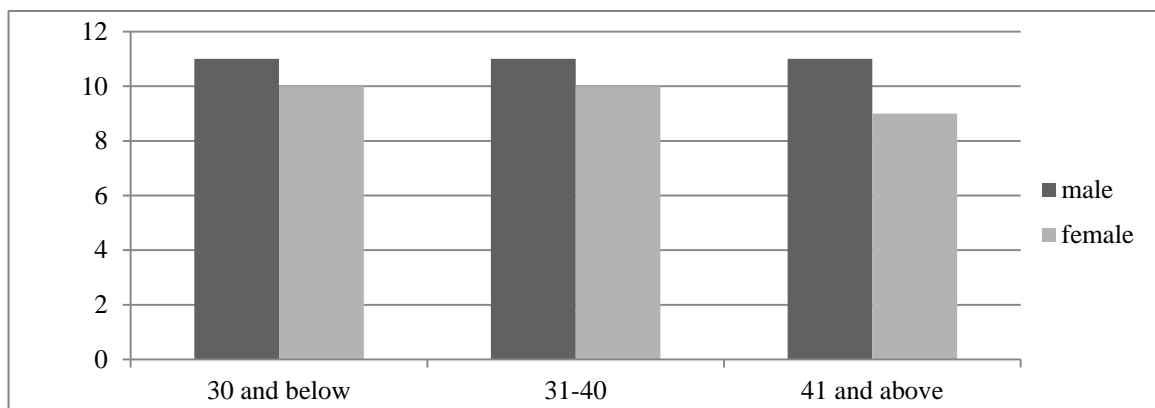| What are the functions of code-switching in your class?	30 and below		31-40		41 and above	
	Male	**Female**	**Male**	**Female**	**Male**	**Female**
1.To explain grammar	3	3	3	3	3	3
2.To manage class	2	2	2	2	2	2
3.To translate unknown vocabulary items	3	3	3	3	3	3
4.For humorous effect	3	2	3	2	3	1
Total:	**11**	**10**	**11**	**10**	**11**	**9**

*multiple response(s)

Figure 1. Graphical Representation of Code-Switching to Filipino

Table 5 and Figure 1 present the frequency distribution of the functions of code- switching to Filipino. Since it is a multiple response/s, all of them believed that code-switching were important to explain grammar and to translate unknown vocabulary, and some of them answered code-switching to Filipino was important to

manage class. Meanwhile, male teachers across age groups used code-switching for humorous effect during class discussions.

Table 6. Frequency of the Functions of Code-Switching to Filipino based on classroom recordings

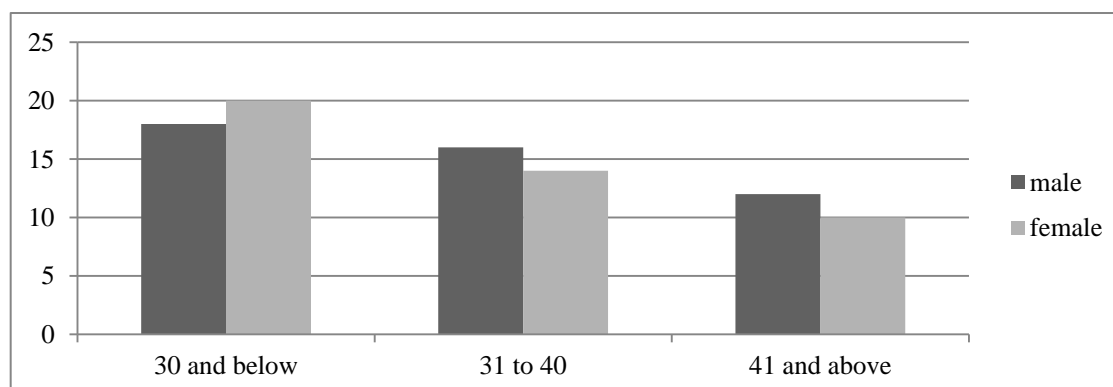| Functions	Times of Occurrence					
	30 and below		31-40		41 and above	
	Male	Female	Male	Female	Male	Female
1.Explaining grammar	18	20	16	14	12	10
2.Managing class	6	4	5	5	6	6
3.Translating unknown vocabulary items	21	23	20	21	18	19
4.For humorous effects	9	5	9	5	7	4
Total:	54	52	50	45	43	39

Figure 2. Graphical representation of times of occurrences of code-switching to
    Filipino in Explaining Grammar

1.   *Explaining Grammar*

It shows that female teachers aged 30 and below code-switched 20 times and male teachers of the same age group code-switched 18 times, while male teachers aged 31-40 code-switched 16 times and female teachers of the same age group code-switched 14 times. Male and female teachers aged 41 and above code-switched 12 times and 10 times respectively. This shows that teachers tended to alternate to Filipino in explaining grammar.

Examples:

(1) T: Okay class, Let's talk about the tense of the verb. What is a verb tense? This indicates the time or duration of the action. ***Diba kapag sinabing tense ito yung kung kailan naganap ang*** action. (tense indicates the time or duration of the action) There are three tenses of a verb, the past, present and future tense. ***Ito yung pangnagdaan, pangkasalukuyan at panghinaharap.***

(2) T: In subject verb agreement, the verb agrees with the subject, not with a noun or pronoun in the phrase. Example, the book, which talked about different Filipino traditions, is interesting. In this sentence ***ang*** subject ***natin ay libro at hindi yung traditions kaya*** is ***ang ginamit nating*** verb. (Our subject is 'book' and not 'traditions', that's why we used 'is' as a verb.)

(3) T: … declarative sentence expresses or states an idea, a fact or an opinion and ends with a period. Interrogative sentence seeks information for an answer or asks a question and ends with a question mark. ***Yung*** declarative class ***simpleng nagkukwento ka lang, o nagsasabi, samantalang yung*** interrogative ***ay nagtatanong…***

In the examples above, the teachers' code-switching behavior to Filipino is observed in grammar instruction, that the teacher shifts his/her language to Filipino for his/her students in dealing with particular grammar points, which are taught at that moment. In these cases, the students' attention is directed to the new knowledge by making use of code switching and accordingly making use of Filipino (Bautista 2004).
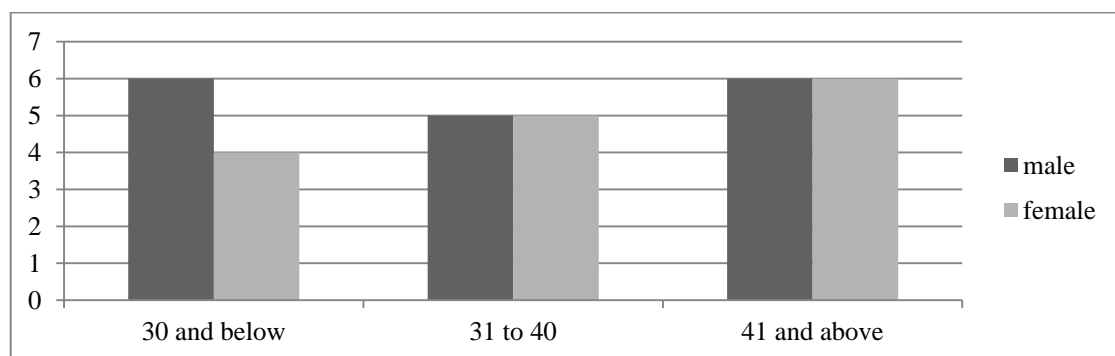

Figure 3. Graphical representation of times of occurrences of code-switching to Filipino in Managing Class

2. Managing Class

The behaviors in which the teachers organize the class also involve a choice of language, so they choose to code-switch to Filipino. The data of the present study

shows that male teachers aged 30 and below code-switched 6 times and female of the same age group code-switched 4 times, while male and female teachers aged 31-40 code-switched 5 times, and male and female teachers aged 41 and above code-switched 6 times.

(1)T: so guys, you will recite your poem, you will be called alphabetically… *walang mag iingay ha? Makinig kayong mabuti sa classmate nyo.* (Please be quiet and listen carefully to your classmate.)

(2)T: … the interviewer sheet is another observation tool which is also called the conference recording form, so *ganito ang gagawin nyo,*(this is what you are going to do) … *naiintindihan nyo ba* class? (Do you understand?)

(3)… before leaving the room please arrange your chairs, *pakidampot yung mga papel at kalat sa tabi nyo.* (Please pick up the pieces of paper and trash near you).
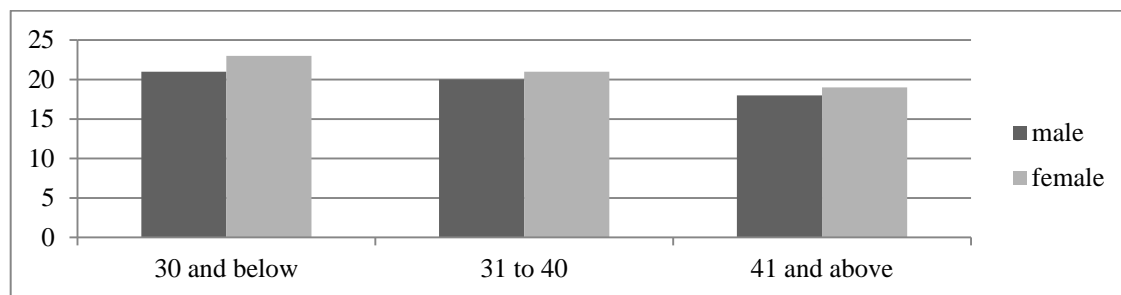

Figure 4. Graphical representation of times of occurrences of code-switching to Filipino in Translating Unknown Vocabulary

3.   *Translating Unknown Vocabulary*

Translating unknown vocabulary had the most number of code-switching occurrences to Filipino, simply because to explain new vocabulary would be difficult for the students to understand. 21 times of code-switching to Filipino occurred in the group of male teachers aged 30 and below while female teachers of the same age group had 23 times of occurrences. While male teachers aged 31-40 had 20 times and female of the same age group had 21 times of code-switching to Filipino, male teachers aged 41 above had 18 times and the female of the same age group had 19 times of code-switching to Filipino. Teachers across age groups frequently provided Filipino equivalents in translating unknown vocabulary to ensure their students' comprehension (Bautista, 2004).

Examples:

(1) T: the affective domain is the least studied and most often overlooked domain in educational literature… ***yung*** affective ***ito yung may kinalaman sa ating pakiramdam o emosyon…***(it has something to do or related with our feelings, emotions).

(2) T: A scoundrel Deborah is going to marry! Now who is going to marry? What is the characteristic of Deborah? … In this sentence, a scoundrel means ***malupit, gumagawa ng masama***… (cruel, evil doer).

(3) T: …ethnographic research has something to do with the values and institutions of any given society. ***Itong*** ethnographic research ***na ito ginagawa o pinag-aaralan sa pamamagitan ng pakikihalubilo mismo ng researcher/s sa mga tao*** (This… it could be done or studied through immersion of the researcher/s) to know their culture, traditions and even their language.
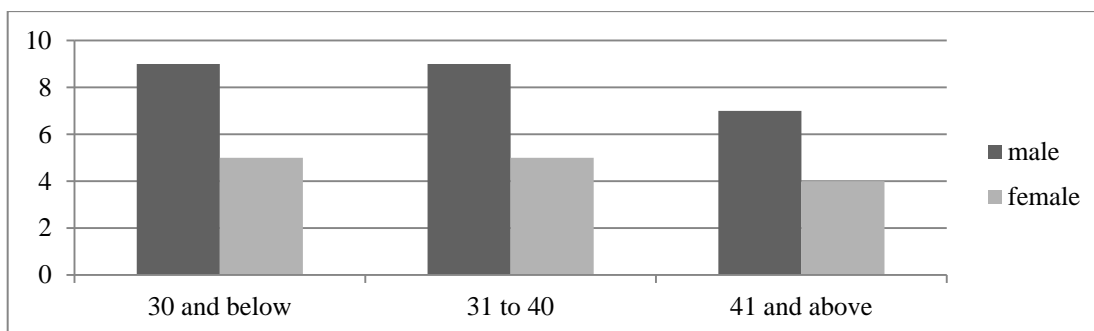

Figure 5. Graphical representation of times of occurrences of code-switching to Filipino for Humorous Effect

4.   *For Humorous Effect*

The figure above shows that the most number of occurrences when it comes to code-switching for humorous effects are the male teachers across age groups. Male teachers aged 30 and below and 31-40 had the same number of code-switching to Filipino with 9 times, followed by male teachers aged 41 and above. While female teachers aged 30 and below and 31-40 also had the same number of code-switching to Filipino and fewer times of occurrences are the female teachers aged 41 and above.

Examples:

(1) T: Personification pertains to giving the inanimate objects the characteristics of a person… example of personification, "***nadurog ang puso ni JP***

(name of the student) nang bastedin siya ni Richelle (name of the student). (The heart of JP was broken into pieces when Richelle rejected him). (Classmates laughed and teased them.)

(2) T: Shakespeare is the literary giant… some of his works are Hamlet, Macbeth… when did he die? And what was written on his tomb?.. ***o sige Jay, tutal classmate mo si Shakespeare dati anu sagot?*** (okay, Jay, since Shakespeare and you were classmates before.) (Students laughed.)

(3) T: Are you listening to the class?... oh yeah, Daniel is listening, ***ayun at naghihikab na, inaantok na yata?*** (Look, he is yawning and feeling sleepy.) (Students laughed.)

According to the study of Chloros (1998), telling jokes and being humorous is traditionally considered unfeminine and so is generally less common among females that might be the reason why in this study male teachers had the most number of code-switching to Filipino for humorous effects. It seemed that teachers tried to say everything first in English and then in Filipino to make sure that everyone has understood.

Generally, it could be gleaned from the table that male and female aged 30 and below had the most number of times of code-switching depending on its functions, followed by male and female aged 31 to 40 and lastly the male and female teachers aged 41 and above. With regards to gender, male occurred many times of code-switching to Filipino than female teachers. Similar to the study of Chloros (1998), gender and age are major variables in language use and consequently in code-switching. It shows that male and female teachers of younger age had the most number of code-switching for them to be understood by their students because of their lack of experience in teaching, while older teachers, aged 41 and above, reported less code-switching because they were more formal inside the classroom and more confident in teaching.

Conclusions

Code-switching in classroom setting was found to be an unavoidable and inevitable phenomenon because teachers' code-switching was not only a part of communicative resources of bilingual classes but also an active part in the learning experience. This study has shown that teachers' code-switching behavior has something to do with their age and gender. The results also show teachers' code-

switching behavior to Filipino has been found to serve various functions like translating vocabulary items, explaining grammar, managing class, and for humorous effects which are in accord with the findings of Macaro (2009) and Levine (2003). In addition, quoting others' words, emphasizing some points are also the functional uses in ESL class. Lastly, code-switching to Filipino is a good strategy of efficiency and benefits classroom teaching and learning.

Recommendations

Based on the foregoing findings and conclusions, it is strongly recommended that the teachers' perception on code-switching to Filipino be analyzed and further studied since code-switching is an efficient strategy in teaching and learning process. The administrators may take into considerations to further analyze the different functions of code-switching in classroom discussions.

References:

Arnfast, J. S., & Jorgensen J. N. (2003). Code-switching as a communication, learning, and social negotiation strategy in first-year learners of Danish. *International Journal of Applied Linguistics*, 13(1), 23-53.

Bautista, L. (2004). Tagalog-English Code Switching as a Mode of Discourse. *Asia Pacific Education Review,*5(2), 226-233.

Chloros, P. (1998). Code-Switching, Gender and Politeness: A study in London Greek-Cypriot Community. *ACTAS/Proceedings Simposio Internacional Bilinguismo. London.*

Cipriani, F. (2001). *Oral participation strategies in the foreign language classroom: An ethnographic account*. Dissertaçao (Mestrado em Inglês) - Universidad Federal de Santa Catarina, Florianopolis.

Cook, V. (2001). *Second language learning and language teaching. Third Edition*. London: Arnold.

Crystal, D. (2003). *The Cambridge Encyclopaedia of Language*. Cambridge: Cambridge University Press.

Duff, P., & Polio, C. (2003). How much foreign language is there in foreign language classroom? *Modern Language Journal,* 74, 154-166.

Greggio, S., & Gil, G. (2007). Teacher's and learner's use of code-switching in the

English as a foreign language classroom: a qualitative study. *Linguagem and Ensino*, 10(2), 371-393.

Gulzar, M. et al. (2013). Inter-Sentential Patterns of Code-Switching: A Gender-Based Investigation of Male and Female EFL Teachers. *International Education Studies*, 6(11), 144-159.

Levine, G. (2003) Student and instructor beliefs and attitudes about target language use, first language use, and anxiety: report of a questionnaire study. *Modern Language Journal,* 87, 343-364.

Macaro, E. (2009). Teacher use of codeswitching in the second language classroom: Exploring 'optimal' use. In M. Turnbull & J. Dailey-O'Cain (Eds.), *First Language Use in Second and Foreign Language Learning* (pp. 35-49). Bristol: Multilingual Matters.

Schweers, C. W. (1999). Using L1 in the L2 Classroom. *English Teaching Forum*, 37(2), 6-9.

Storch, N., & Wiggleworth, G. (2003). Is there a role for the use of the L1 in an L2 setting? *TESOL Quarterly*, 37, 760-770

Appendix

Dear teacher,

This questionnaire is administered as part of the research undertaken by the undersigned on **"Comparison of Code- Switching Behaviors of Male and Female Classroom Teachers across Age Groups."** This is not a test. It has no right or wrong answers. The information obtained through the questionnaire will be kept confidential and will be used solely for research purposes. Please read the statements and the questions carefully and answer them honestly. The success of the investigation depends on your sincerity in filling the questionnaire. Your cooperation will be greatly appreciated and acknowledged.

Thank you.

Yours sincerely,
Christina D. Vicencio, M.A.E
Christopher S. Vicencio, Ed.D.
Researchers

Teachers' Questionnaire:

I.Personal and Professional Characteristics of the teachers.

Name: (optional)

Personal Background

Gender: ___Male   ___Female

Age: ___ 30 and below   ___31-40   ___41 and above

Professional Rank   ___Part-time Instructor

    ___Instructor

    ___Assistant Professor

    ___Associate Professor

    ___Professor

II. Gender Differences in code-switching to Filipino.

1. How often do you use Filipino in the class?

A. always   B. sometimes   C. occassionally   D. never

2. Are you conscious of Filipino switching in the class?

A. Yes, always.   B. Yes, sometimes.

C. Yes, occassionally.   D. No, never.

3. Do you think code-switch to Filipino is an efficient strategy of learning and teaching?

A.yes   B. No

III. Functions that influence teachers' switching to Filipino.

4. What are the functions of code-switching in your class? (You may give more than one choice.)

A. to explain grammar.

B. to manage class.

C. to translate unknown vocabulary items.

D. for humorous effect.

How Teachers Can Make the Keyword Method More Challenging for Students

Satima Rotjanawongchai
Prince of Songkla University, Trang Campus, Thailand

Bioprofile

Satima Rotjanawongchai is an English lecturer at the Prince of Songkla University, Trang Campus in Thailand. She holds a Master's degree in Translation and a Bachelor's degree in Spanish with first-class honors from Chulalongkorn University, Thailand. Her research interests concern theoretical applications to classroom practices. Email: katibbchan@gmail.com

Abstract

In many ESL and EFL classrooms, the responsibility for increasing vocabulary knowledge lies with the teacher. Although there are various vocabulary learning strategies that can be used and have been studied, the ones used are almost always chosen by the teachers. This could hinder students from becoming autonomous learners. The keyword method, a mnemonic technique involving making an acoustic link between a native-language word or phrase that sounds or looks like the foreign word, is one of the most extensively researched vocabulary learning strategies. In many studies, the method has been proven to enhance vocabulary learning however, many of the studies provided the keywords for the subjects rather than having students generate the keywords themselves. Therefore, the question of whether teachers can make the method more challenging and more beneficial to learner autonomy arises. Analyzing and evaluating the data from vocabulary tests, questionnaires, and in-depth interviews shows that a new application of the keyword method can be effective in improving both vocabulary learning and learning autonomy.

Keywords: vocabulary learning, the keyword method, learning autonomy

Prince of Songkla University, Trang Campus,
Kuanpring, Muang, Trang, Thailand

Introduction

"If language structures make up the skeleton of language, then it is vocabulary that provides the vital organs and the flesh." (Harmer, 1993, p.153). Wilkins (1972) explains that without grammar knowledge it is still possible to use a language for communication, but it is impossible to do so without vocabulary knowledge. Vocabulary competency is a crucial element of second language (L2) acquisition (Avila & Sadoski, 1996). However, ESL/EFL learners around the world face problems with learning English vocabulary. There is substantial evidence that many learners, particularly in an English as a foreign language (EFL) context, do not have sufficient vocabulary knowledge and vocabulary size to apply in order to communicate effectively in many English language contexts (Alonso, 2013; Barrow, Nakanishi, & Ishino, 1999; Chui, 2006; Nurweni & Read, 1999).

In Thailand, English is taught to students starting in primary school and is a compulsory course at the tertiary level. However, Thai learners still face all sorts of problems in learning English, especially the English lexicon. Jenpattarakul (2012) listed 9 problems Thai students encounter which affect their reading skills. For example, the inadequacy of independent reading creates fewer chances for them to encounter new vocabulary. Another problem is that they cannot guess a word's meaning from the context, so they cannot get the gist of the text. This might be due to their lack of practice in using the learned words, as Jenpattarakul also pointed out, Thai students do not review the words they had learned and do not complete the vocabulary exercises by themselves. The most crucial problem might be the fact that they like being "spoon-fed" vocabulary knowledge and prefer the teacher to translate vocabulary for them. Reports show that the average English proficiency of Thai students is lower than those of students from other ASEAN countries (Prapphal, 2003).

In order to boost the vocabulary knowledge of learners, studies of vocabulary learning are considered "a promising area of inquiry" (Ellis, 1990, cited in Tavakoli & Gerami, 2013). Vocabulary learning strategies have been studied by many researchers, such as Oxford (1990), Gu and Johnson (1996), and Schmitt (1997). Most of these

studies compare one vocabulary strategy to another strategy, or to multiple other strategies, in relation to one of the several aspects of vocabulary knowledge, e.g. retention and recall. On the contrary, in Thailand, most research on English vocabulary learning has focused on the vocabulary learning strategies used by learners (e.g. Boonkongsaen & Intaraprasert, 2014; Intaraprasert, 2004; Komol & Sripetpun, 2011; Nirattisai & Chiramanee, 2014; Saengpakdeejit, 2014; Siriwan, 2007).

In spite of the increasing number of studies in this field of research, vocabulary specialists are in general agreement on the point that there is a need for further experimental studies to determine the effectiveness of different vocabulary teaching approaches on learning. For this reason, there have been a number of studies on effectiveness of vocabulary learning strategies undertaken (e.g. Brown & Perry, 1991; Fan, 2003; Gu, 2010; Little & Kobayashi, 2015; Lu, 2008; Singleton, 2008). One of the most extensively researched vocabulary strategies is the keyword method (Jenpattarakul, 2012), which is a mnemonic technique involving making an acoustic link between a native-language word or phrase that sounds or looks like the foreign word or phrase. There have been a number of studies done on the keyword method, most of which explain its advantages in vocabulary learning, recall, retention, and learners' attitudes (e.g. Atkinson & Raugh, 1975; Godley, Fournet, & Estes, 1987; Hauptmann, 2004; Jenpattarakul, 2012; McDaniel & Pressley, 1989; Moore & Surber, 1992). For example, McDaniel and Pressley (1989) found that the keyword method was more facilitative to learning than the context method. The study of Atkinson and Raugh (1975) indicated that the keyword method was more effective on Russian vocabulary learning than an unconstrained control procedure. Similarly, Wing-wa (2012) studied the effectiveness of the keyword method and the context method in L2 vocabulary acquisition and the result showed that the keyword method was superior to the context method in vocabulary retention. However, learners could encounter problems when applying the method due to unfamiliarity with the method and the time required to create links between the keywords and their corresponding foreign words (Hall, Wilson, & Patterson, 1981). As Djigunovic (2000) suggested, using a new strategy may make vocabulary learning more difficult because learners who have been learning foreign languages for some time develop their own strategies to help

them master foreign language vocabulary. It is necessary to explore the effectiveness of the keyword method compared to student's old method in order to prove if this limitation affects the method's usefulness.

Although many studies investigated the effectiveness of the keyword method compared to other vocabulary strategies. Most of those studies used ready-made keyword technique examples with the students. They studied the application of the keyword method using supplied keywords and images (e.g. Atkinson & Raugh, 1975; Avila & Sadoski, 1996; Jenpattarakul, 2012; Siriganjanavong, 2013). Some studies have shown that the keyword-provided approach is more effective than the keyword-generated approach for college students (e.g. Atkinson & Raugh, 1975; Campos, Amor, & Gonzalez, 2004; Hall, Wilson, & Patterson, 1981). It was suggested that the reason may be that the students were not able to generate the effective keywords themselves. Therefore, it was inferred that using provided keywords is more suitable. It could also be implied that if learners are provided with the keywords, the method will benefit their vocabulary learning. However, this application of the method might not be appropriate for increasing learner autonomy, which is the purpose of strategy training (Nation, 2008, cited in Mizumoto, 2013). Moreover, those studies were conducted in laboratory-like settings, which were markedly different from what occurs in a classroom (Fuentes, 1976). For these reasons, it is worth trying to shed the light on whether the usefulness of the keyword method remains true with total learner autonomy in a real-life classroom environment.

To investigate whether this adaptation of the keyword method is effective in vocabulary learning, the method was compared to the control group. The control group studied new words in a traditional way, repetition, translation, and example sentence explanation.

The study was targeted at answering these questions:

1.   Can the keyword method promote vocabulary learning in a real classroom if students generate the keywords themselves?

2.   What are students' attitudes toward the keyword method if they generate the keywords themselves?

The results of this study shed light on the use of vocabulary learning strategies among Thai students, as well as other EFL students, who wish to improve their vocabulary knowledge independently.

Methodology

The population in this study was 300 Thai speaking freshmen at a university in Thailand. Participants were selected from two intact classes of first-year students who were taking a compulsory three-hour-per-week Fundamental English Reading and Writing course. Eighty-one participants were purposively selected and grouped into an experimental group (keyword method group) and a control group (a traditional method group). However, the number of participants finally fell at 74 because seven students dropped out of the course, making 41 students and 33 students left in the experimental group and the control group respectively. The participants were 17-24 years old. All participants had been learning English for at least 6 years. Both classes were taught by the same instructor using the same teaching material from "Unlock Level 1 Reading and Writing Skills".

Each group was assigned to study the 212 target training words from the vocabulary sections of units 1-10 in the textbook. The experimental group was introduced to using keyword method as their vocabulary learning strategy, while the control group involved in traditional vocabulary instruction, including repetition, translation, and example sentence explanation.

To begin the study, the participants were tested on their existing vocabulary knowledge level by using a vocabulary level test. The test was adapted from the monolingual vocabulary level test by Schmitt et al. (2001). It is claimed to be an effective tool for supplying a profile of a learner's vocabulary, particularly for placement and diagnostic purposes (Schmitt, 2010) and has been widely used and accepted by researchers since 1980s (Laufer, 1998; Laufer & Nation, 1995; 1999). The test consists of five levels of word frequency in English: the 2,000-word level, the 3,000-word level, the 5,000-word level, the 10,000-word level, and Academic Vocabulary. The test format is multiple choice word-definition matching. There are

30 items for each frequency level yielding 150 items in total. In this study, the 10,000 word level and Academic Word List were not employed because they are beyond the participants' knowledge level, hence the total number of items was decreased to 90.

Instrument

A questionnaire was used to collect data on the students' attitudes towards the keyword method and the vocabulary notebooks. The questionnaire consisted of three sections. Section A and B were Likert scale items and section C contained open-ended questions. Section A focused on the students' personal information. Sections B and C investigated the students' opinions about the methods. The total number of all items in the questionnaire was 14. In order to establish content validity, the questionnaire was reviewed by three Teaching English to Speakers of Other Languages (TESOL) professors and a pilot test was administered to 50 students at a university to ensure the validity and reliability of the questionnaire. The overall reliability coefficient (Cronbach's alpha) in the pilot study was .87.

Semi-structured interviews were conducted with five students randomly chosen from the keyword method group to explore their opinions of the method. Although a written list of questions was used as a guide, the interviewees were informed that there may be additional questions asked to seek more in-depth information. The interview was audio recorded and field notes were taken. To reduce the interviewees' anxiety and get the clearest answers, the interviews were carried out in the Thai language.

Two intact classes were assigned to the keyword method group and the traditional method group with the same instructor and the same course book. Before the treatment, the vocabulary level test adapted from the monolingual vocabulary level test by Schmitt et al. (2001) was administered to determine the homogeneity of their English vocabulary levels. The participants were given thirty minutes to complete the test. A correct answer to an item received one point and there was no penalty for incorrect responses. During the treatment, the keyword method group

received keyword method training, while the control group was instructed by the traditional instruction.

The research period lasted one semester (17 weeks) and the students were taught ten lessons from the book "Unlock level 1 Reading and Writing Skills", with a total of 212 vocabulary items. On the first day of the course, one group was introduced to using the keyword method to learn new words, while the other was taught by a traditional method.

During the semester, the words from each lesson were presented to both groups. The keyword method group was assigned to develop their own link between an English word and its meaning in Thai, as well as a mental image relating the keyword with the Thai translation. Meanwhile, the control group was instructed using the traditional vocabulary learning method. At the end of the semester, vocabulary knowledge was compared between the two groups using a vocabulary quiz. The quiz was a 40 item multiple-choice format in which students were required to choose the best words to complete sentences. To prove the validity of the test, it was reviewed by 10 Thai TESOL instructors and 3 native English speakers. Then troublesome wording and other difficulties were corrected.

After scoring the quiz, the students in the experimental group were asked to complete the attitude questionnaire and five of them were asked to participate in an in-depth interview.

Results

To answer the research questions, the data from the vocabulary quiz of 40 items, the attitude questionnaires, and the in-depth interviews were analyzed. To answer research question 1 relating to the effects of the keyword method on the students' vocabulary learning, the vocabulary quiz scores and the t-test score will be shown in Table 1 and Table 2 respectively. As for research question 2 about the students' attitudes toward the keyword method on the condition that they created the keyword themselves, Table 3 will show the mean score of their attitudes on each area, and Table 4-7 will reveal the summary of their answers to the four open-ended

questions in the questionnaire. To get more in-depth information about the students' attitudes, the data from the in-depth interviews will be presented.

Table 1. Vocabulary quiz scores (out of 40)

group	N	S.D.	Mean	Max	Min
Keyword group	48	6.06	26.17	37	14
Control group	33	5.99	22.3	35	7

Table 2. T-test score of the vocabulary quiz

	Levene's Test for Equality of Variances	t-test for Equality of Means
	Sig.	Sig. (2-tailed)
Equal variances assumed	0.956	0.004

Question 1.   Can the keyword method promote vocabulary learning in a real classroom if students generate the keywords themselves?

According to Table 1, the statistical results from the vocabulary quiz shows that the keyword method group outperformed the control group, with the mean score of 26.17 compared to 21.30 for the traditional instruction group. To assure that the mean score difference between the two groups was due to the different learning method, a t-test was applied. According to table 2, the Sig (2-Tailed) value is 0.04 (<0.05). Therefore, it can be concluded that there is a statistically significant difference between the mean score on the quiz of the two groups. The experiment reveals that the keyword method could promote vocabulary learning in a real classroom even when the students generate the keywords themselves.

Question 2.   What are students' attitudes toward the keyword method if they generate the keywords themselves?

Another question which this study sought to answer was what the students thought about the keyword method. To achieve this aim, a questionnaire on the attitudes toward the keyword method was distributed to the participants and semi-structrued interviews were performed.

Table 3. Students' attitudes towards the keyword method from the questionnaire

item	attitude	mean
1	I can remember word meaning better	4.06
2	I can recognize word meaning better	4.25
3	Apart from its meaning, I can learn other aspects of word	3.87
4	The method makes me more interested in learning new words	4.00
5	I have a chance to manage my own vocabulary learning	4.25
6	I feel more positive about vocabulary learning because of this method	4.15
7	I can use this method in the future when I want to learn new words.	4.23
8	This method is suitable for me.	3.73
9	Learning words by using this method is worth my time.	3.96
10	I like this method.	3.75

According to Table 3, the participants rated their opinions about each aspect on a 5-point Likert Scale – (1) strongly disagree, (2) disagree, (3) no opinion, (4) agree, or (5) strongly agree. Overall, most students had a postive attitude toward the keyword method, with mean score of 4.06. Questionnaire items 4, 5, and 6 were related to learning autonomy and had high mean scores (4.00, 4.25, 4.15 respectively). When we looked at which item had the greatest number of participants answered "strongly agree" to, it was item 5, which is about learning autonomy. It can be concluded from this that the participants agree that the keyword method can promote vocabulary learning in terms of enhancing learning autonomy. Considering the participants' attitude toward the usefulness of the method for remembering and recoginizing vocabulary (mean scores of 4.06 and 4.25 respectively), most of the participants agreed that the keyword method helped them remember and recognize words better. Another interesting point is that the least positive response was to item 10 (mean score of 3.75), which asked about fondness for the keyword method. It

might be possilbe that even if the students thought the keyword method was useful for vocabulary learning, they did not like the method.

Open-ended question results

In order to get more detailed information on the participants' atttitudes toward the keyword method, four open-ended questions were included in the questionnaire. The questions were: (1) "What are the problems with using the method?" (2) "Why did you use or not use the method outside of the classroom?" (3) "What did you like most about the method?" and (4) "What did you dislike most about the method?". Table 4-7, shows the students' answers to those four open-ended questions.

Table 4. Problems with using the keyword method

Problems	No. of participants
-too lazy to use this method	5
-couldn't think of a keyword	5
-couldn't think of a keyword and its relation to word meaning	9
-didn't like the way I had to remember what I thought	2
-didn't know how to describe my way of remembering	1
-spent a lot of time thinking about a keyword and its relation to word meaning	3
-didn't know how the word is pronounced, so couldn't find a keyword	2
-the method didn't work	5
-the method was confusing	1
-N/A	8
Total	**41**

Table 5. Reasons for disliking the keyword method

What I disliked about this method	No. of participants
- could not think of a relation between the keyword and the meaning	9
-could not use the method with all words	4
-was hard and confusing	6
-hard to find a keyword	3
-N/A	18
Total	**41**

Table 6. Reasons for liking the keyword method

What I liked most about the method	No. of participants
-could use creativeness and felt sense of control	14
-could remember and recognize words more easily and quickly	11
-aroused interest in learning more words	2
-helped learning pronunciation	2
-was an interesting and easy method	2
-liked making connections between the keyword and word meaning	2
-enjoyed sharing my way of remembering words with friends	1
-could use the method outside of the classroom	1
-N/A	6
Total	**41**

After analyzing the responses, it was found that the inability to find a keyword that related to a word's meaning was considered the biggest problem by most participants (Table 4). This result is in accordance with the response to question 4, which asked about what they disliked about this method. Most students said that what they disliked most was the fact that they could not think of a relation between the keyword and the meaning (Table 5). Although there were some problems and obstacles with this method, the participants reported that they liked the sense of control they felt when using the method (Table 6). Moreover, they realized its usefulness for vocabulary learning (word meanings and pronunciation). This could explain why most of them also used the keyword method outside of the classroom and in other classes. The major reason why they used the method out of the classroom expressed by participants was that they thought it helped remember and recognize words better and more easily (Table 7).

Table 7. Use of the keyword method outside of the classroom

Use of the method out of the classroom	No. of Participants
-Yes	32
-No	3
-N/A	6
Total	**41**
Reason for "Yes"	
- helped remember and recognize words better and more easily	12
-saved a lot of time learning new words	
-made me enjoy learning new words	1
-liked self-created way of remembering	7
-N/A	1
	11
Reason for "No"	
-words from the lessons were good enough	1
-liked another method better	2

Interview Results

Five participants from the keyword method group were interviewed to complement the findings from the quantitative data. There were some interesting aspects from the questionnaire results that was worth investigating in depth. First, the responses from the questionnaire revealed that the students found the keyword method useful for remembering and recognizing the meaning of new words while they also thought that this method was not easy to apply. The interview results showed how the students coped with this problem and why they still believed this strategy was helpful.

Four of the interviewees insisted that the keyword method was useful for remembering and recogninzing the meaning of new words, although in the early stages some of them felt that it was difficult to find a relationship between the keyword and the word's meaning. They said that while they tried hard to find a keyword (a Thai word that sounded like the new English word) and related the keyword to the meaning of the English word, they would uncousciously remember the translation of that word.

"At first, I thought this way of learning vocabulary was quite complicated. But when I figured it out, it seemed like I could remember the meaning of the word automatically." (Interviwee 1)

"I like the keyword method because I can use my imagination. When I kept thinking about which keyword I should use and how I could relate it to the meaning, the meaning stuck in my head for a while and then I could remember it." (Interviewee 2)

"It took me a long time to find a keyword and its relation to the word's meaning in Thai, that's why I was a bit against this method. But when I got past that stage, I felt like I could remember and recall the words effortlessly. It really helped me when I saw the word in the test." (Interviewee 4)

"At the beginning I thought this method was difficult and was too lazy to connect a Thai word with the meaning. Later, I thought it was a lot of fun and sometimes I laughed at my imagination. That helped me learn words more happily and effectively." (Interviewee 5)

However, one of the five students thought that the keyword method could not help her remember words better because she was accustomed to learning words by rote memorization.

"I thought the keyword method did not suit my learning style. I prefered rote learning because I found it more effective. What's more, I didn't like the method because I didn't enjoy thinking." (Interviewee 3)

The answer of interviewee 3 could explain a reason why some students did not like the keyword method. It showed that he was not accustommed to the new method and preferred his own learning style and that a vocabulary learning strategy might not be suitable for every learner. In addition, it supported that learners favored easy stretegies which did not require thinking more than complex strategies (Schmitt, 1997).

The second point emphasized in the interview concerned the interviewees' opinions on the usefulness of the method and how they applied the strategy with other classes. As the quesionnaire results showed, some students thought the keyword method was also helpful for learning pronunciation and spelling. The four interviewees who had postive attitudes towards the method agreed that it could help with learning pronunciation and spelling. They explained that the reason for this was

that they had to know how the word was pronounced in order to find a Thai word that sounded like the target word. Moreover, they said that they had also applied this method to technical vocabulary in other classes, such as Marketing and Business Administration. Also, four of them believed that the keyword method promoted learner autonomy. They felt they had full responsibility of their learning and thought they could be more autonomous learners in the future.

"The keyword method led me to be an autonomous learner. This method helped me practise my thinking process, which caused self-directed learning. I am quite sure that I could learn by myself better in the future." (Interviewee 1)

"This method aroused my interest in learning vocaubulary by myself outside of the classroom. I have alrealy used it with new vocabulary in my Business course." (Interviewee 2)

"I always enjoy learning English outside of the classroom from movies and other media. I feel the keyword method is really useful for enhancing my vocabulary learning. When I see new words, I can try adapting this method my own way." (Interviewee 4)

"I used to try many ways to remember the meanings of English words, but they did not work. When the teacher introduced the keyword method and I decided to try it and I found that it was an effective way to learn vocabulary. I liked this method because I could feel a sense of freedom to design my way of learning. This helped me be an independent learner." (Interviewee 5)

The third focus in the interview was about advantages and disadvantages of letting the students create the keyword and its relationship to the word's meaning, two of the four students who successfully applied the method pointed out that it was challenging to do all the process of the method themselves. They thought it was too hard to find a keyword themselves for some words. Therefore, they suggested that teachers should give them some examples so that they could more easily think of their own way of remembering the words. However, they still agreed with the idea that teachers should make them think deliberately and be totally responsible for all the decisions concerned with their learning. In contrast, the other two students believed

that teachers should not provide the keyword to students in any circumstances because, if students try hard enough, they will find the keyword themselves and that this would be better for long-term retention.

"I think it would be better if the teacher could help us when we can not think of a keyword or how to relate it to the word's meaning. In the cases where the teacher's way of remembering does not match ours, we could use it as a guideline to create our own ways." (Interviewee 1)

"It would be nice if the would teachers help us with finding the keyword when we cannot figure it out, but this should be only for difficult words like abstract words or long words." (Interviewee 2)

"If the teacher had told us which word was the keyword and shown us how to connect it with the English word's meaning, I would not have been able to remember the meaning of the word because it was not from my own mind and then I would need to learn it by rote as before." (Interviewee 4)

"I disagree with the idea of the teacher providing the keyword for us. I believe I could remember words better if I take complete charge of my learning." (Interviewee 5)

Another interesting point is that the four students admitted that they might have been less motivated if the teacher had not told them to submit a record of how they used the method. In addition, recording in a notebook how they created a keyword and its relationship to the word's meaning helped them revise for the exams.

"If the teacher had not assigned us to report our application of this method, I might not have been as enthusiastic as this." (Interviewee 1)

"Thinking of it as an assignment, I was more actively involved in this method." (Interviewee 2)

"Jotting down the keywords and their reltionship to the word's meaning was very useful for the exams. This was different from rote learning because I could go over it more quickly and didn't have to stuff many words into my head. What I needed to do is to go over what I thought before." (Interviewee 4)

"I was happier with my vocabulary learning. I didn't have to torture myself by repeating words and trying to remember them. I just learned my way of remembering from what I noted in my notebook." (Interviewee 5)

Discussion

The main goal of this study was to investigate the effectiveness of the keyword method on vocabulary learning when students generate the keywords themselves. The strategy has been proven to enhance the vocabulary learning of college students (Atkinson & Raugh, 1975; Jenpattarakul, 2012; Kasper & Glass, 1982; Moore & Surber, 1992; Pressley, Levin, Hall, Miller, & Berry, 1980; Raugh & Atkinson, 1975; Siriganjanavong, 2013). Similarly, the data from this research suggested that the keyword method could enhance vocabulary learning. However, the students admitted that there were some problems with this method. The most important problem was an inability to create a keyword and its relationship to the English words' meanings. Crutcher (1990) showed that one of the obstacles to retrieving the vocabulary definitions was related to unsuccessful retrieval of the keyword and/or unsuccessful retrieval of a native-language equivalent from the keyword. The present study and Crutcher's study shared interesting findings about the problems with the keyword method, even though the keywords were supplied in Crutcher's research. It could be claimed that these problems need to be discussed more if teachers want to introduce the method into their classrooms. Nonetheless, it does not mean that the keyword method should be avoided. The questionnaire data showed that most of the students realized the advantages of the strategy for remembering and recognizing word meanings and also chose to apply it outside the classroom and with other subjects of study.

The results of this study proved that the keyword method is still effective in vocabulary learning even when the students create the keywords themselves. The assumption that the students are not able to generate effective keywords by themselves is not completely true. Although some students admitted that creating the keywords was hard, they found that they could get through it if they put more effort into it. Not only is it useful for vocabulary learning, letting students create the

keywords reinforces learning autonomy and learning satisfaction, as they feel a sense of control and creativity. It is possible that, due to more than 6 years of learning English, the subjects found generating the keywords was more effective than being provided the keywords. According to Atkinson and Raugh (1975), the keyword-provided condition is superior to the keyword-generated condition. They argued that the results of their experiment were based on beginners who had no previous training in the target language (Russian) and that provided keywords became less useful when subjects were more familiar with the language. Therefore, it could be suggested that students who possess some knowledge of the target language should be encouraged to create the keywords themselves.

Hall, Wilson, and Patterson (1981) studied the effectiveness of the keyword method on Spanish word learning. One group of subjects viewed each provided keyword for 10 sec and at the beginning of each 10-sec period heard or saw the Spanish word pronounced three times within the first 4 sec of the keyword's presentation. Then the group was compared to another group which also used the keyword method but was free to study how they wished. It was found that if students studied the words freely, the keyword method was not superior to the control group. Interestingly, the results of the present study showed that the keyword method was still effective even when the students used the method freely by allowing them to create the keywords and manage word learning on their own. Moreover, they could manage the word presentation by reviewing the keywords in their notebook. These results also negate the claim that learners who have been learning foreign languages for a long time may be too unfamiliar with a new strategy and it could interfere with its usefulness. Actually, according to the findings, even though students did face some troubles in taking up the new method in the early stages, they were able to work through it on their own.

Classroom implications

It is clear that the keyword method is useful for vocabulary learning even in a keyword-generated condition. Also, it yields positive attitudes in terms of learning

freedom and autonomy. However, to implement the keyword method in a real classroom, teachers should keep in mind that this method does not suit every student and that some students might not be able to generate the keywords on their own. Therefore, it is recommended that teachers should provide them some suggestions and let them decide whether they will continue to use the method. Teachers could also suggest to students that the method be integrated with other vocabulary strategies to best suit individual's preferences. Ultimately, teachers need to look back to the major objective of teaching strategies, which is to enhance learner autonomy (Nation, 2008, cited in Mizumoto, 2013).

To avoid the resistance to the new learning strategy, teachers might provide the keywords for some new words at the introduction of the keyword method. In the Thai context as the findings showed, some students found the teachers' control of using the method helped them keep on the track and put more efforts to the method. However, teachers need to realize their students' characteristics and learning styles when making decisions about this. For example, if the students are quite autonomous, teachers could be facilitators more than directors.

Limitations

Although the size of the study was limited, only including 74 Thai undergraduate students, which is clearly not enough to make generalizations about all Thai students or all undergraduate students, a clear pattern emerged from the findings. The overview of applying keyword-generated condition with undergraduate students it provides is worth noticing and future research could use the findings to explore using the keyword method in an integrated way with a larger sample and adapting the study to account for advantages and disadvantages of the method in order to uncover a new way of applying the method to achieve the best results. This could lead to innovative ideas for inspiring students who have lost hope for learning L2 vocabulary.

References

Alonso, A. C. (2013). Receptive vocabulary size of secondary Spanish EFL learners. *Revista de Lingüística y Lenguas Aplicadas*, 8, 66-75.

Atkinson, R. C. & Raugh, M. R. (1975). An application of the mnemonic keyword method to the acquisition of a Russian vocabulary. *Journal of Experimental Psychology: Human Learning and Memory*, 104, 126-133.

Avila, E. & Sadoski, M. (1996). Exploring new applications of the keyword method to acquire English vocabulary. *Language Learning*, *46,* 379-395.

Barrow, J., Nakanishi, Y., & Ishino, H. (1999). Assessing Japanese college students' vocabulary knowledge with a self-checking familiarity survey. *System*, 27, 223-247.

Boonkongsaen, N. & Intaraprasert, C. (2014). English Vocabulary Learning Strategies Employed by Thai Tertiary-Levels Students with Different Genders and Levels of Vocabulary Proficiency. *International Journal of Scientific and Research Publications*, 4(3), 1-9.

Brown, T. S. & Perry, F. L. (1991). A Comparison of Three Learning Strategies for ESL Vocabulary Acquisition. *TESOL Quarterly*, 25, 655–670.

Campos, A., Amor, A., & Gonzalez, M. A. (2004). The importance of the keyword-generation method in keyword mnemonics. *Experimental Psychology*, *51*(2), 1-7.

Chui, A. S. Y. (2006). A study of the English vocabulary knowledge of university students in Hong Kong. *Asian Journal of English Language Teaching*, 16, 1-23.

Crutcher, R. J. (1990). *The Role of Mediation in Knowledge Acquisition and Retention: Learning Foreign Vocabulary Using the Keyword Method*. Retrieved from http://www.colorado.edu/ics/sites/default/files/attached-files/90-10.pdf

Djigunovic, M. J. (2000). Language learning strategies and affect. *CLCS Occasional Paper No. 59*, Trinity College, Dublin.

Fan, M. Y. (2003). Frequency of Use, Perceived Usefulness, and Actual Usefulness of Second Language Vocabulary Strategies: A Study of Hong Kong Learners. *The Modern Language Journal*, 87(2), 222–241.

Fuentes, E. J. (1976). An investigation into the use of imagery and generativity in learning a foreign language vocabulary. *Dissertation Abstracts International*, 37, 2694.

Godley, R. M., Fournet, G. P., & Estes, R. E. (1987). The effects of administration, treatment and pacing factors on the keyword method of foreign vocabulary learning. *Educational and Psychological Research*, 7(3), 173-181.

Gu, Y., & Johnson, R. K. (1996). Vocabulary learning strategies and language learning outcomes. *Language Learning*, 46(4), 643-697.

Gu, Y. (2010). Learning Strategies for Vocabulary Development. *Reflections on English Language Teaching*, 9(2), 105–118.

Hall, J. W., Wilson, K. P., & Patterson, R. J. (1981). Mnemotechnics: Some limitations of the mnemonic keyword method for the study of foreign language vocabulary. *Journal of Educational Psychology*, 73, 43-49.

Harmer, J. (1993). *The Practice of English Language Teaching*. Harlow, UK: Longman.

Hauptmann, J. (2004). *The effect of the integrated keyword method on vocabulary retention and motivation*. Unpublished doctoral thesis. University of Leicester. Leicester, England. Retrieved from
https://lra.le.ac.uk/bitstream/2381/30861/1/U187838.pdf

Intaraprasert, C. (2004). EST students and vocabulary learning strategies: A preliminary investigation. *Suranaree Journal of Science and Technology*, *12*(2), 163-171. Retrieved from http://www.ird.sut.ac.th.

Jenpattarakul, W. (2012). The impact of Keyword Technique on the Students' Vocabulary Retention Ability in an EFL Class. *Mediterranean Journal of Social Sciences*, 3(3), 265-573.

Kasper, L. F., & Glass, A. L. (1982). The role of the keyword method in the acquisition of Spanish nouns. *Human Learning*, 1, 235-250.

Komol, T., & Sripetpun, W. (2011). *Vocabulary Learning Strategies Employed by Undergraduate Students and its Relationship to their Vocabulary Knowledge*. Paper Presented at the 3rd International Conference on Humanities and Social Science, 1-18.

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

Laufer, B., & Nation, P. (1995). Vocabulary Size and Use: Lexical Richness in L2 Written Production. *Applied Linguistics*, 16(3), 307-322.

Laufer, B., & Nation, P. (1999). A vocabulary-size test of controlled productive ability. *Language Testing*, 16, 33-51.

Little, A. & Kobayashi, K. (2015). Vocabulary Learning Strategies of Japanese Life Science Students. Vocabulary Learning Strategies of Japanese Life Science Students. *TESOL Journal*, 6(1), 81-111. doi: 10.1002/tesj.141.

Lu, M. (2008). Effectiveness of vocabulary learning via mobile phone. *Journal of Computer Assisted Learning*, 24, 515–525.

Mizumoto, A. (2013). Enhancing Self-efficacy in Vocabulary Learning: A Self-regulated Learning Approach. *Vocabulary Learning and Instruction*, 2(1), 15-24.

McDaniel, M. A., & Pressley, M. (1989). Keyword and context instruction of new vocabulary meanings: Effect on text comprehension and memory. *Journal of Educational Psychology*, 81(2), 204-213.

Moore, J. C., & Surber, J. R. (1992). Effects of context and keyword methods on second language vocabulary acquisition. *Contemporary Educational Psychology*, 17(3), 286-292.

Nation, I. S. P. (2008). *Teaching vocabulary: Strategies and techniques*. Boston, MA: Heinle, Cengage Learning.

Nirattisai, S., & Chiramanee, T. (2014). The Contribution of Vocabulary Learning Strategies to University Students' Vocabulary Size. *Silpakorn University Journal of Social Sciences, Humanities, and Arts*, 14(2), 45-66.

Nurweni, A., & Read, J. (1999). The English vocabulary knowledge of Indonesian university students. *English for Specific Purposes*, 18, 161-175.

Oxford, R. L. (1990). *Language Learning Strategies: What every teacher should know*. New York, NY: Newbury House/Harper Collins.

Prapphal, K. (2003). English Proficiency of Thai learners and Directions of English Teaching and Learning in Thailand. *Journal of English Study*, 1(1), 6-12.

Pressley, M., Levin, J. R., Hall, J. W., Miller, G. E., & Berry, J. K. (1980). The keyword method and foreign word acquisition. *Journal of Experimental Psychology*, 6, 163-173.

Raugh, M. R., & Atkinson, R. C. (1975). A mnemonic method for learning second-language vocabulary. *Journal of Educational Psychology*, 67, 1-16.

Saengpakdeejit, R. (2014). Strategies for Dealing with Vocabulary Learning Problems by Thai University Students. *Silpakorn University Journal of Social Sciences, Humanities, and Arts*, 14(1), 147-167.

Schmitt, N. (1997). Vocabulary learning strategies. In N, Schmitt. & M, McCarthy (Eds.), *Vocabulary: Description, Acquisition, and Pedagogy* (pp. 199-227). Cambridge, UK: Cambridge University Press.

Schmitt, N. (2010). *Researching Vocabulary: A Vocabulary Research Manual*. Basinstoke: Palgrave Macmillan.

Schmitt, N., & McCarthy, M. (Eds.). (1997). *Vocabulary: Description, acquisition, and pedagogy*. Cambridge: Cambridge University Press.

Schmitt, N., Schmitt, D., & Clapham, C. (2001). Developing and Exploring the Behavior of Two New Versions of the Vocabulary Level Test. *Language Testing*, *17*, 55-58.

Singleton, D. (2008). *Vocabulary learning strategies and foreign language acquisition*. England: Multilingual Matters.

Siriganjanavong, V. (2013). The Mnemonic Keyword Method: Effects on the Vocabulary Acquisition and Retention. *English Language Teaching*, 6(10), 1-10.

Siriwan, M. (2007). *English Vocabulary Learning Strategies Employed by Rajabhat University Students*. Unpublished doctoral thesis. Suranaree University of Technology. Nakorn Ratchasrima, Thailand). Retrieved from http://sutir.sut.ac.th:8080/sutir/bitstream/123456789/284/1/mayuree_fulltext.pdf

Tavakoli, M., & Gerami, E. (2013). The Effect of Keyword and Pictorial Methods on EFL Learners' Vocabulary Learning and Retention. *Porta Linguarum*, *19,* 299-316.

Wilkin, D. A. (1972). *Linguistics in Language Teaching*. London, UK: Edward Amold.

Wing-wa, Y. (2012). *An investigation into the effectiveness of and attitudes towards the keyword method and the context method in L2 vocabulary acquisition by Cantonese-speaking Chinese learners in an English as the medium of instruction (EMI) secondary school in Hong Kong*. Unpublished Master thesis. University of Hong Kong. Hong Kong. Retrieved from http://hub.hku.hk/handle/10722/180103.

Gender Roles and Representations in the English E-Textbooks for Junior High School in Indonesia

Ouda Teda Ena
Sanata Dharma University, Yogyakarta

Made Frida Yulia
Universitas Negeri Malang
Home base: Sanata Dharma University, Yogyakarta

Carla Sih Prabandari
Sanata Dharma University, Yogyakarta

Bioprofiles:
Ouda Teda ENA is a lecturer at the English Education Department, Sanata Dharma University, Yogyakarta, Indonesia. He obtained his Master's degree from Universitas Negeri Malang in 2005 and his Doctor of Education from Loyola University Chicago in 2013. His research interests are ELT, Sociolinguistics, and Critical Pedagogy. Email: ouda.art@gmail.com

Made Frida YULIA is a lecturer at the English Education Department, Sanata Dharma University, Yogyakarta, Indonesia. She obtained her Master's degree in English Education from Universitas Negeri Malang in 2005. She is currently taking postgraduate study at the same university. Her research interests are ELT, Sociolinguistics, teacher education, and professional development. Email: madefrida@gmail.com

Carla Sih PRABANDARI has been teaching at the English Education Department, Sanata Dharma University, Yogyakarta, Indonesia, since 1999. She earned her Master's degree in English Language Studies of Sanata Dharma University in 2008. Her main interests are Linguistics, English Language Teaching, and teacher education. Email: ibucarlasp@gmail.com

Abstract

Indonesia is a large country with diverse population. The diversity lies among others in gender aspect. Ideally, the diversity should be accommodated in the educational system, for example in teaching and learning materials for schools. However, the diversity seems to be ignored due to the enactment of a centralized policy in

Sanata Dharma University, Jln. Affandi Mrican
CT Depok, Sleman – Yogyakarta
Universitas Negeri Malang, Jln. Surabaya no. 5,
Malang – Jawa Timur

Indonesian educational system about the publication of the materials in the form of e-textbooks. This research investigated whether the English e-textbooks for Junior High School in Indonesia which were published by the Ministry of Education addressed gender roles and representations impartially. There were respectively three e-textbooks for 7th, 8th, and 9th grades; thus, in total there were nine e-textbooks. Document analysis was used to analyze the visual images that were used as illustrations of the e-textbooks. Human images were the focus of the study. To observe whether gender issues had been addressed proportionately in the English e-textbook, an analysis on the roles of male and female images was conducted. The study showed that women were underrepresented in the English e-textbooks for Junior High School in Indonesia. The illustrations mainly portrayed men as a more favorable gender. The study also depicted biased representations and some stereotyping of women's roles. Women were associated with certain types of jobs and roles, such as being a mother, sensitive, caring, weak. By contrast, men were associated with more powerful jobs and strong roles.

Keywords: gender roles, gender representation, English e-textbooks

Introduction

Indonesia is a large country with diverse population. Ideally, the diversity should be accommodated in the educational system. However, the diversity seems to be ignored due to the enactment of a centralized policy in Indonesian educational system. Despite the fact that the Decentralization Law was ratified in 1999 (Tyson, 2010), the government still continues to adopt a centralized policy system. One example of the centralized policy is the publication of teaching and learning materials. The government controls and publishes the materials for schools. Centralization has the potential for creating imbalanced power relations between the central and local governments. There is a possibility that the publication of the materials does not include certain local interests and therefore marginalize these groups and prioritize other groups.

E-textbooks publication is one of the centralized publication policies. This policy has also been criticized as costly and marginalizing areas where the Internet facilities do not exist (Maryulis, 2008). Publishing school e-textbooks for a very diverse population of students like in Indonesia should be done with great cautions.

The e-textbooks should include all students' cultural traits so that they are inclusive and would increase students' learning motivation. Although the government stated that these e-textbooks had been reviewed and certified as qualified by the National Bureau of Education Standardization (Nuh, 2010), there have not been any research regarding whether these books are culturally representative of the diverse population of students.

The Ministry of Education had published 927 e-textbooks for all school levels up to January 2012. They are 291 e-textbooks for elementary school, 154 for junior high school, 276 for senior high school, and 204 for vocational school. There are three different titles for all levels of junior high school; therefore, there are nine e-textbooks in total.

Issuing centralistic policy for the whole nation should be done with great cautions especially in Indonesia. Indonesia with its 220 million populations, more than 700 languages, and around 1000 ethnic groups is probably the most diverse nation in the world. Indonesia is a large country covering an area of 5,193,250 square kilometers, out of which 2,027,087 square kilometers are land and the remainder is water. The Indonesian population is estimated to be 220,953,634 people. It is the fourth in the world for its population density. According to the data from the Central Bureau of Statistics, the number of school children is 25,389,000 (*Biro Pusat Statistik* or The Central Bureau of Statistics, 2010). These school children represent the diversity of the Indonesian population. The most notable diversity in Indonesia includes ethnicity, religion, gender, and social economic status. Centralistic curriculum with its centralistic textbooks could easily hamper students' motivation by not including their cultural traits in the textbooks (Ena, 2013). Publication of textbooks in Asia (Yen, 2000 Yi, 1997) and in Indonesia (Ena, 2013) often failed in including the cultural aspects that are important for the students such as gender, ethnicity, socioeconomic status, and religion.

This study was a content analysis observing gender roles and representations in the English textbooks for Junior High School in Indonesia. The purpose of the study was to evaluate the visual images of the English textbooks. Qualitative visual content analysis was employed as the method of the research (Krippendorff & Bock, 2009; Merriam, 2009; Rosengren, 1981; Willis, 2007). The analysis focused on whether the textbooks accommodate the gender roles of the students and represent them impartially. It sought to answer the question of whether gender role diversity

was represented in the visual images of the English textbooks for Junior High School in Indonesia. This study was crucial to be conducted because the textbooks are intended for a very diverse students' population. The research was limited to studying one cultural aspect of the students, namely gender.

Literature Review

The Curriculum and the English Textbooks for Junior High School

The first standardized curriculum applied nationally was developed and implemented in 1975. The national curriculum of English was revised in 1984, 1994, 2000, and 2003 (Mistar, 2005; Sadtono, 1997; Yuwono, 2005). According to the 1975 curriculum, the purpose of English teaching in high school was to facilitate the development of advance science, technology, culture, and arts, and to enhance international relations. The four language skills, namely reading, writing, listening, and speaking, were taught discretely. In 1984 the Ministry of Education revised the national curriculum and therefore it was known as the 1984 Curriculum. The teaching method adopted in this curriculum was known as the Communicative Approach. Another curriculum was introduced in 1994. There were not many changes in English language teaching. However, there was a change of priority of language skills in the 1994 curriculum. The most important language skill was reading, followed by listening, speaking, and writing. In 2004 a competency-based curriculum was adopted. There were not any significant changes in terms of learning materials. In 2006, in line with the implementation of the Regional Autonomy Law, the Ministry of Education promoted school-based management and school-based curriculum. Schools were expected to develop their own curriculum and be independent financially. The adoption of a school based management system had only benefited schools in the cities and in the well-developed areas.

In 2013 a centralistic curriculum was implemented. It is centralistic in nature because it was developed by the central government and was implemented throughout the country. The curriculum documents, syllabuses, students' textbooks, and teachers' books were produced by the central government in this case by the Ministry of Education. All the learning objectives named as core competence and basic competence were set by the government. Teachers were only allowed to develop the lesson plans based on the syllabuses given (Nuh, 2013).

English E-textbooks in Indonesia

The publication of e-textbooks in Indonesia was based on the assumption that technology enhances learning. The e-textbook policy instituted by the Indonesian Ministry of Education was meant to provide inexpensive textbooks for schools therefore increasing access to resources. It was introduced for the first time by the Decree of the Ministry of Education No. 46/2007. The Ministry of Education bought the copyrights from textbook writers and made the textbooks accessible for free on the Internet. The e-textbook publications continue under the new minister of education with the Decree of Minister of Education No. 41/2008. The Ministry of Education has published 927 E-textbooks for all school levels up to January 2012. The e-textbooks are accessible and downloadable from http://bse.kemdiknas.go.id/. According to the government, these e-textbooks have been reviewed and certified as qualified by the National Bureau of Education Standardization (Nuh, 2010).

The main advantages of e-textbooks are ease of use, low cost, and fast content delivery. However, it also has limitations due to such issues as slow Internet connections and access to computers. The implementation of the e-textbook policy needs to be evaluated to make sure that it is beneficial for all intended users. Critical education theory is suitable to evaluate whether the policy provides better access to learning resources for all students or if it offers benefits only to certain groups and therefore marginalizes other groups.

Critical Education

Critical education refers to a philosophy of education as well as a process of critique towards the practices of education with the emphasis on the power relations of different parties involved in it. McLaren (1998, cited in Wink, 2000) defines critical pedagogy as a way of thinking, while Giroux (2001) characterizes critical education as theoretical work as well as a conscious practice of education. McLaren claims that

> "Critical pedagogy is a way of thinking about, negotiating, and transforming the relationship among classroom teaching, the production of knowledge, the intuitial structures of the school, and the social and material relations of the wider community, society, and nation state (McLaren, 1998, cited in Wink, 2000, p. 31)."

Critical education is an application of critical theory in education. According to Giroux (2001), critical theory refers to the theoretical work developed by certain members of the Frankfurt School as well as to the nature of self-conscious critique and to the need to develop a discourse of social transformation and emancipation that does not cling dogmatically to its own doctrinal assumptions. It refers to both a school of thought and also a process of critique. Critical theory focuses on analyzing the contemporary power interests between groups and individuals within society – identifying who gains and who loses in specific situations. The central focus of critical research is the dynamics of how privileged groups support the status quo to maintain their privileges (Kincheloe & McLaren, 2000, p. 281).

According to Kincheloe and McLaren (2000), "critical social theory is concerned in particular with issues of power and justice and the ways that the economy, matters of race, class, gender, ideologies, discourses, education, religion and other social institutions, and cultural dynamics interact to construct a social system" (p. 281). Therefore, critical education is concerned with issues of power and justice within the realm of education. These issues of power relations and justice exist because of different economic status, race, class, gender, ideologies, discourses, education, religion, and other social factors embedded in the education system. These issues are often very subtle that people do not notice and are considered to be normal. Critical education is a way of analyzing imbalances in power relations within education as well as a conscious practice to attempt to bring about changes toward a more just education system.

Critical education is suitable as the theoretical framework for the study because it is concerned with issues of power and justice within the realm of education. The objective of the study is to observe whether the learning materials provided by the government privileges certain groups of students and thereby marginalizing other groups. Critical theory in education will be employed to study the unfair dynamics between marginalized and privileged groups. Critical education focuses on analyzing the contemporary power interests between groups and individuals within schools, identifying who gains and who loses in specific situations. Critical theory will be the framework of this study in an examination of the English e-textbooks for grades 7th, 8th, and 9th published by the government.

Gender

Gender is a combination of nature and culture, biological traits as well as learned behaviors (Ryan, 2010). The Indonesian government applies a heteronormative standard where there are only two genders, male and female; however, homosexuality and transgender exist in public life and are not criminalized. Transgendered people are accepted as part of the society although they often become victims of discrimination. They generally can only work in certain fields, such as fashion, beauty salons, entertainment, and some work as street singers or prostitutes. Apart from a male-dominated society in general, some communities adopt matrilineal systems in which women are in charge of the household and men take their wife's surnames, such as in Minangkabau and West Timor (World Trade Press, 2010).

According to the 2010 Indonesian government census, the populations of women and men in Indonesia were 118,010,413 and 119,630,913 respectively. Indonesia is basically a patriarchal society, where women are expected to play traditional subordinate roles as daughters, wives, and mothers; however, Indonesian women have come a long way in their majority Islamic and male-dominated society. They have become more economically independent over the last few years. Women gained the right to vote at the same time as Indonesia's independence in 1945. Women can legally engage in any social and economic activities. They can inherit equally with men in spite of customary Islamic inheritance laws, which are unfavorable to women (World Trade Press, 2010).

Gender roles are among other cultural identities that play an important role in learning. Sheets (2005) identifies how they develop:

> "Gender roles develop through a socialization process in the family and community. These roles are screened through specific cultural norm. Other family characteristics such as race, culture, socioeconomic level, class, and religion can also significantly shape children's gender learning (p. 43)."

Although gender inequality still exists, the situation is changing. Indonesian law mandates that political parties should include at least 20% women as their candidates for the House of Representatives and the local senate (World Trade Press, 2010).

In education, women are often marginalized. In rural areas and under-developed areas where resources are scarce when a choice has to be made, parents often discourage girls from continuing their education. Therefore, boys outnumber girls in enrollment in elementary schools and only 12.8% of girls finish high school. The literacy rate of women from all age groups is as low as 42% (Biro Pusat Statistik, 2010).

Learning Materials

Learning materials play an important role in teaching and learning activities in the classrooms. Textbooks are often the only if not the main sources of learning in Indonesia. Indonesia with a population of more than 200 million is a diverse country in terms of ethnicity, language, religion, and socio-economic status (SES). Publishing any learning materials for schools should take diversity into account in order that the materials cater to the learning needs of students with different ethnic, language, religion, and socio-economic backgrounds.

In order to promote learning, materials should be culturally relevant and inclusive, accommodating the many cultural differences of the students. Culturally relevant teaching materials should include students' culture in order to maintain it and to avoid misinterpretations of other cultures (Ladson-Billing, 2009). Culturally relevant teaching is a pedagogy that empowers students intellectually, socially, emotionally, and politically by using cultural referents to impart knowledge, skills, and attitudes. These cultural referents are not merely vehicles for bridging or explaining the dominant culture; they are aspects of the curriculum on their own right (Ladson-Billing, 2009).

Tomlinson and Lynch-Brown (2010) recommend some criteria for evaluating and selecting culturally relevant learning materials for school and classroom libraries. These criteria include authentic depiction of the cultural experience from the perspective of that group, accuracy of cultural details in text and illustrations, positive images of minority characters, balance between historic and contemporary views of groups, and adequate representation of any group. Similarly, Campbell (2010) puts forward fifteen minimum qualitative criteria to align or choose textbooks to make sure that they are culturally relevant to the students. The first criterion is that the learning materials are unbiased and non-stereotypical. Secondly, they should have a comprehensive, complete, and inclusive view of society and its history. They should

also have diverse viewpoints. They should not only represent the culture of the majority group. Good learning materials should build on and extend students' experiences. Finally, they should help students analyze and comprehend how real-life situations are. The literature on critical education and instructional material development help the researchers to develop the criteria to evaluate whether the learning materials which are culturally appropriate reside in both the written texts as well as the visual images. Both of them could be used to study whether certain learning materials or textbooks have balanced representation (Sheets, 2005).

Based on what has been presented above, the research aimed at evaluating the visual images of the English e-textbooks designated for the $7^{th,}8^{th}$, and 9th graders in Indonesia. The analysis focused on whether and to what extent these e-textbooks accommodated gender diversity of the students.

Research Question

The research question was formulated as follows.

To what extent is gender role diversity represented in the visual images of the English e-textbooks for Junior High School (7th, 8th, and 9th grades) in Indonesia?

Research Methods

Research Subjects

There were nine (9) English e-textbooks for grades 7th, 8th, and 9th of Junior High Schools which became the subjects of investigation. Three e-books were published for each grade. The purpose of the study was to provide descriptions of visual images incorporated in those nine English textbooks and critically observed how they represented the cultural diversity of the students. The study also provided interpretations as to how these visual images marginalized or prioritized certain group of students.

The following was the description of the e-textbooks which were studied. The three e-textbooks provided for the 7th grade were:

1. *Bahasa Inggris Sekolah Menengah Pertama Kelas VII: Contextual Teaching and Learning* by Th. Kumalarini, Achmad Munir, and Slamet Setiawan (2008), retrievable from

   http://bse.kemdikbud.go.id/index.php/buku/details/20080724125512

2. *English in Focus* for Grade VII Junior High School by Artono Wardiman, Masduki B. Jahur, and M. Sukiman (2008), retrievable from http://bse.kemdikbud.go.id/index.php/buku/details/20080723130147

3. *Scaffolding:* English for Junior High School Students Grade VII by Joko Priyana, Riandi, and Anita P. Mumpuni (2008), retrievable from http://bse.kemdikbud.go.id/index.php/buku/details/20080723130258

The three e-textbooks for the 8th grade were:

1. *Bahasa Inggris Sekolah Menengah Pertama Kelas VIII: Contextual Teaching and Learning* by Utami Widiati, Gunadi H Sulistyo, and Nunung Suryati (2008), retrievable from http://bse.kemdikbud.go.id/index.php/buku/details/20080724125512

2. *English in Focus* for Grade VIII Junior High School by Artono Wardiman, Masduki B. Jahur, and M. Sukiman (2008), retrievable from http://bse.kemdikbud.go.id/index.php/buku/details/20080723130408

3. *Scaffolding*: English for Junior High School Students Grade VIII by Joko Priyana, Arnis Rahayu Irjayani, and Virga R. (2008), retrievable from http://bse.kemdikbud.go.id/index.php/buku/details/20080723130534

The three e-textbooks for the 9th grade were:

1. *Bahasa Inggris Sekolah Menengah Pertama Kelas IX: Contextual Teaching and Learning* by Gunarso Susilohadi, Suharso and Dwi Anggani (2008), retrievable from http://bse.kemdikbud.go.id/index.php/buku/details/20080724125129

2. *English in Focus* for Grade VIII Junior High School by Artono Wardiman, Masduki B. Jahur, and M. Sukiman (2008), retrievable from http://bse.kemdikbud.go.id/index.php/buku/details/20080723130625

3. *Scaffolding*: English for Junior High School Students Grade VIII by Joko Priyana, Riandi, and Anita P. Mumpuni (2008), retrievable from http://bse.kemdikbud.go.id/index.php/buku/details/20080726124625

Data Analysis Technique

The study employed document analysis. The analysis focused on whether the visual images of the e-textbooks accommodated the gender roles of the students and represented them impartially. It sought to answer the question of whether gender role diversity was represented in the visual images of the English textbooks for Junior

High School in Indonesia. This study was crucial to be conducted considering that the textbooks are intended for a very diverse students' population.

Visual images were used as the sources of data to observe the representation of gender roles in the English e-textbooks because of their richness in providing cultural information. They did not merely accompany the text, because they often provided much more important cultural information than the texts. They often provided the subconscious beliefs of the writers. What images included was as important as what images which were absent from the textbooks because they revealed the writers' reasons behind the choice (Taylor, 2002). Banks (2007) states that there are two main reasons for using visual images in research. First, visual images are easy to find and to access. The second reason for incorporating them is that "they might be able to reveal some insight that was not accessible by any other means" (p.4).

Human images were used in the study. It was divided into male and female. The binary category of male and female was used to correspond to the official category adopted by the Indonesian government although transgender individuals were found in popular media such as televisions and newspapers. Frequency of occurrence was assigned to the subcategory to enable the researchers to interpret the frequency data of the visual images. Descriptions were given when the visual images signified certain roles. An undetermined category was assigned when the visual image did not belong to any category or there was not enough information to classify the visual image. A data collection form was used to record each visual image. A summary of the frequency of occurrences was given. A summary of the description of each variable for the books was also made when all images were described.

To maintain consistency and to obtain validity, multiple coding was employed. To observe whether gender issues had been addressed proportionately in the English textbooks, an analysis on the roles of male and female images was conducted. A careful observation on the roles they were depicting and whether they were trivial or important was conducted.

Results and Discussion

I. Visual Images Representing Gender

The study was concerned with whether the English e-textbooks designated for Junior High School in Indonesia represented gender equally in their illustrations. The representativeness of gender was examined by comparing the roles and number of

male and female characters in the visual images used as illustrations in the e-textbooks. The categories of female, male, and undetermined were used to classify the visual images examined. Three types of visual images, namely images of people, animals, and objects were examined to determine gender representations. The data showed that the visual images of both animals and objects did not contain information about gender. Therefore, the representation of gender was observed mainly through the images of people found in the nine English e-textbooks for grades 7, 8, and 9.

The total number of images of human beings in the nine e-textbooks studied was 1619 as shown in Table 1. The number of images depicting male characters was 884 or 54.60 %. The female group was represented by 687 images or 42.43 % in the e-textbooks. Therefore, there was an imbalanced representation in terms of number between male and female groups in the textbooks. In terms of number, the images of men were more dominant than the images of women as shown in Table 1.

Table 1. The number of human images representing gender

Grade	7				8				9				Total
Textbook	**a**	**b**	**c**	**Total**	**D**	**e**	**f**	**Total**	**g**	**h**	**i**	**Total**	
Female	66	138	97	301	96	43	103	242	47	46	51	144	687 (42.43%)
Male	138	148	64	350	131	116	105	352	50	41	91	182	884 (54.60%)
Undetermined	0	0	0	0	1	0	4	5	36	0	7	43	48 (2.96%)
Total	204	286	161	651	228	159	212	599	133	87	149	369	1619

Gender role comparison between females and males was examined to determine the fairness of the representations. Gender role was defined as a set of social and behavioral norms that were generally considered appropriate for either a man or a woman in the society (Ryan, 2010; Sheets, 2005). The English e-textbooks for grades 7, 8, and 9 represented both females and males as having different roles and social status. However, they also depicted the existing presumptions about males and females. These presumptions were often in the form of gender role stereotypes.

The images that depicted female stereotypes in traditional gender roles were the pictures of women cooking or preparing meals and doing household chores. There was no male depicted as cooking in the e-textbooks, while the female characters depicted as cooking or preparing food were abundant. Similarly, only very few males were depicted as doing household chores and there were many females doing them.

Other images that showed male and female role stereotypes were the pictures that showed emotions. Sadness and compassion were mainly shown by females in the English textbooks examined. There were girls shown crying, girls holding dolls and flowers which might be interpreted as showing care or tenderness. The e-textbooks also showed pictures of different women planting flowers and taking care of children or babies. No such activities were depicted done by males. The male role stereotypes were depicted by men as being assertive and having power or control. These were shown in e-textbooks where boys were playing soccer, painting, and working outdoors.

Most occupations in the e-textbooks were represented by males and females. In terms of number, the profession such as teacher was almost equally represented. However, some professions were only represented by men. All doctors, politicians, and soldiers in the e-textbooks were men.

II. Gender Role Stereotypes and Biases

Gender is not a mere biological trait but it is also learned behavior (Ryan, 2010). Gender roles are learned in the family and through education. Gender stereotypes exist in the society where males and females are often overly expected to behave in certain manner based on the societal beliefs and norms. These expectations often limit the opportunities of women and girls to perform certain tasks or assume certain jobs. In education, stereotyping could limit students' future decisions regarding various aspects of their lives, including choice of profession and career development. Stereotypes keep both sexes in traditional professions and are incoherent with the various kinds of existing opportunities (Mills & Mills, 1996).

The study showed that the English e-textbooks for Junior High School grades 7, 8, and 9 in Indonesia depicted some stereotypes of women roles. It may have negative impacts on female students because it might limit them from performing certain tasks or assuming certain jobs or taking certain opportunities.

The findings of the study were found to be similar to those revealed in an earlier study conducted by Ena (2013). His focus of analysis was on e-textbooks for Senior High School in Indonesia.

Conclusion

The research question was concerned with how the English e-textbooks for Junior High School in Indonesia represented gender through the use of visual images in their illustrations. First, in terms of numbers, women were underrepresented in the overall English textbooks examined. There were 687 images of women and girls compared to 884 images of men and boys. Secondly, men were depicted as having wider range of roles as shown by their occupations and tasks they performed. Certain professions such as politicians, doctors and soldiers were only depicted by men. Other occupations such as teachers and police officers were illustrated by both men and women. The visual images used as illustrations in the English e-textbooks also contained gender stereotypes related to emotions and feelings. Women were depicted as caring, sympathetic, and passionate, while men were depicted as having power and kept their emotions in control.

The visual images of males and females in the textbook did not represent the reality that exists in the society. In reality, the number of women is more than the number of men according to 2000 population census (Biro Pusat Statistik, 2010). Women play important roles in the society. Women own about sixty percent of micro, small, and medium scale business enterprises. Women also hold more than 11 percent of the seats in the parliament and represent about 11 percent of government ministers. More than sixteen percent of judges and about fifteen percent of Supreme Court Judges are women (World Trade Press, 2010). Women are increasingly playing more important roles in the society. This fact should be better represented in all textbooks used in schools.

Publishing school e-textbooks for a very diverse population of students like in Indonesia should be done with great cautions. The e-textbooks should include all students' cultural traits so that they are inclusive and would increase students' learning motivation. Textbook writers and publishers should create balanced representations of males and females so that the English textbooks would be motivating and engaging for both male and female students. Balanced and accurate depictions of both genders in instructional materials would make students aware of the inequalities that have been experienced by women in Indonesian culture and encourage students to take corrective actions (Mills & Mills, 1996). The balanced representations should not only be in equal numbers of both genders but also they should be depicted as having the same roles. Gender role stereotypes that limit men

and women to perform certain tasks should be avoided. Showing egalitarian roles might lead students to have broader views on the choices of professions.

References

Banks, M. (2007). *Using visual data in qualitative research.* Sage: London.

Biro Pusat Statistik. (2010). *BPS Strategic Data: Statistics Indonesia*. Jakarta: BPS-Statistic Indonesia.

Campbell, D. E. (2010). *Choosing democracy: a practical guide to multicultural education.* Boston: Allyn & Bacon.

Ena, O. T. (2013). *Content analysis: Visual analysis of e-textbooks for senior high school in Indonesia*. South Carolina: Createspace.

Giroux, H. A. (2001). *Theory and resistance in education: towards pedagogy for the opposition.* Westport: Bergin & Garvey.

Kincheloe, J. L., & McLaren, P. (2000). Rethinking critical theory and qualitative research. In N. K. Denzin & Y. S. Lincoln (Eds.), *Handbook of qualitative research* (pp. 279-314). Thousand Oaks: Sage Publication, Inc.

Krippendorff, K., & Bock, M. A. (2009). *The content analysis reader*. Los Angeles: Sage.

Ladson-Billings, G. (2009). *The dreamkeepers: successful teachers of African American children.* San Francisco, CA: Jossey Bass.

Maryulis, M. (2008). *Buku digital bikin ribet anak sekolah (Digital books are such hassle to school children)*. Retrieved on April 19, 2010, from http://maryulismax.wordpress.com/2008/07/24/buku-digital-bikin-ribet-anak-sekolah.

Merriam, S. B. (2009). *Qualitative research: a guide to design and implementation.* San Francisco, CA: Jossey Bass.

Mills, R. & Mills, R. (1996). Adolescents' attitude toward female gender roles: Implications for education. *Adolescence*. 31(123), 741-745.

Mistar, J. (2005). Teaching English as a foreign language (TEFL) in Indonesia. In G. Braine (Ed.), *Teaching English to the world*: *History, curriculum, and practice* (pp. 75-85). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Nuh, M. (2010). *Buku teks pelajaran murah Departemen Pendidikan Nasional (Inexpensive textbooks from Department of National Education) (Depdiknas).* Retrieved on April 19, 2010, from http://bse.depdiknas.go.id.

Nuh, M. (2013). *Permendikbud dan kepmendikbud tentang implementasi Kurikulum 2013*. Retrieved on 1 May 2014 from http://www.kopertis12.or.id/2013/09/13/permendikbud-no-81a-tahun-2013-tentang-implementasi-kurikulum.html.

Rosengren, K. E. (1981). *Advances in content analysis*. Beverly Hills: Sage Publication.

Ryan, M. (2010). *Cultural studies: a practical introduction*. Chichester: Wiley-Blackwell.

Sadtono, E. (1997). ELT development in Indonesia a smorgasbord. In E. Sadtono (Ed.), *National seminar: the development of TEFL in Indonesia* (pp. 1-20). Malang: Penerbit IKIP Malang.

Sheets, R. H. (2005). *Diversity pedagogy: Examining the role of culture in the teaching-learning process.* Boston: Pearson Education Inc.

Taylor, E. W. (2002). Using still photography in making meaning of adult educators teaching beliefs. *Studies in the education of adults*, 34(2), 123-140.

Tomlinson, C. M., & Lynch-Brown, C. (2010). *Essential of young adult literature*. Boston: Pearson.

Tyson, A. D. (2010). *Decentralization and adat revivalism in Indonesia: the politics of becoming indigenous.* Hoboken: Routledge.

Willis, J. W. (2007). *Foundations of qualitative research: interpretive and critical approaches.* Thousand Oaks, CA: Sage.

Wink, J. (2000). *Critical pedagogy: notes from the real world.* New York: Addison-Wesley Longman Inc.

World Trade Press (2010). *Indonesia society and culture complete report.* Petaluma: World Trade Press.

Yen, Y. (2000). *Identity issues in EFL And ESL textbooks: A sociocultural perspective*. (Doctoral dissertation). Available from ProQuest Dissertations & Theses database. (UMI No. 9971668)

Yi, H. (1997). *A Content analysis of Korean textbooks for adult learners of Korean as a foreign language.* Unpublished doctoral dissertation. Available from ProQuest Dissertations & Theses database. (UMI No. 9929015)

Yuwono, G. I. (2005). *English language teaching in decentralised Indonesia: Voices from the less privileged schools.* Paper presented at the Australian Association for Research in Education, International Education Research Conference, The University of Western Sydney, Parramatta.

Employing Inspiring and Appealing Materials as Supplemental Reading to Improve EFL Learners' English Language Competence

Dr. Hsiu Chiao (Sally) Fan
Department of Applied English Studies, China University of Technology, Taiwan

Bioprofile:
Hsiu Chiao (Sally) Fan, Ph.D., is a full-time assistant professor in the department of Applied English Studies at China University of Technology in Taiwan. She graduated from the University of Kansas in TESL. Her research and teaching interests focus on reading comprehension, metacognitive strategies of reading, self-regulated learning, and by extension, linking reading to writing. E-mail: sallyfan28@gmail.com

Abstract:

Much research has been done on supporting improving the quality of pleasure reading in developing ESL/EFL learners' English language competence in the ESL/EFL realm in recent years. Students in Taiwan, however, pay little attention to English pleasure reading. Accordingly, students spend less time fostering their reading habit and are inferior to other EFL/ESL students in English reading comprehension. The purpose of the study was to determine if including inspiring and interesting pleasure reading can improve EFL learners' reading abilities. Additionally, in order to see if "gender difference" has a meaningful impact on pleasure reading, gender as a variable is also important in the current study. Three instruments were utilized to collect data: (1) metacognitive awareness of reading strategies inventory, (2) pre-test, and (3) post-test. The results indicated that there was a significant difference between the experimental group and the control group on the posttest reading scores. When students employed inspiring and appealing materials like *Guideposts* as part of their supplemental reading, they experienced greater improvement in both their reading comprehension and English language competence.

Keywords: appealing materials, English language competence, inspiring materials, pleasure reading

China University of Technology: No. 56, Sec. 3, Xinglong Rd., Wunshan District, Taipei City 11695, Taiwan (R.O.C.)

Introduction

Much research on supporting improving the quality of pleasure reading in developing ESL/EFL learners' English language abilities has been done in the ESL/EFL realm over the past years. Examples included Janopoulo's (1986) study on the correlation between pleasure reading and writing proficiency, Ross's (1999) study for proving the academic success not only in the success of academic performance but also in lifelong experience, and Thorndike's (1973) study on examining the lifelong reading habits of 17- and 18-year-old high school students. These studies results indicated that subjects who read more, experienced superior comprehension and better overall English competency compared to those who read less.

According to Krashen (1981), pleasure reading and intrinsic interest had the potential for bringing in the needed input for language acquisition. Parlette and Howard (2010) also stated that positive experiences with pleasure reading may assist in the transition to adulthood during the university experience and may even enhance students' retention. Anderson, Hiebert, Scott, Wilkinson, and Becker (1988) also believed that reading is a cornerstone for success, no matter if it's in school or throughout life. It could refer to a way of connecting one's relaxedness of mind with the text materials, which would help the learners, specifically the English language learners, advance their competence.

Reading for pleasure has been notably beneficial, i.e., not just for increasing literacy, but for the profound effect on individuals' lives (Gilbert & Fister, 2010). As such, the author would like to know whether using pleasure reading as an approach can motivate EFL learners to immerse themselves in a thrilling world of interesting and impressive stories, and help them get involved in developing their English language competence as well. Inspiring and appealing materials such as popular magazines, daily newspapers, and books of easy short stories (Sapayan, 2009) will be the focus on the treatment to look for evidence.

Another aspect is that students are likely to pay less attention to reading for pleasure while in high school due to the emphasis the college entrance exam places on the expository text rather than the narrative text that is typically found in pleasure reading, specifically in English reading. Even though there was some research conducted to support the benefits of pleasure reading such as Chen's (2007) study of extracurricular reading habits of college students in Taiwan and Lin's (2009) study of the current situation of reading in high school and the corresponding improvement in

pleasure reading with their school learning, the volume of research in the field is still limited, with most of these studies focusing on Chinese language materials. Moreover, there was little evidence to support the assumption of improvement in English language competence via pleasure reading. Consequently, it is assumed that students who spend less time fostering their "reading habits" (Howard, 2013, October, p. 3) are inferior to other ESL/EFL students in English reading proficiency who take the time to include pleasure reading.

For the above reasons, the aim of the study was to determine if including inspiring and interesting pleasure reading can improve ESL/EFL learners' English reading as well as English language competence, specifically for adult EFL learners at university level. Additionally, in order to determine whether "gender difference" (Green & Oxford, 1995) has a meaningful impact on pleasure reading and whether male or female students could benefit more from pleasure reading has also drawn much interest and attention; many studies conducted in America have found that heavy readers are more likely to be female than male, and to have achieved a higher educational level than the population at large (Ross, 1999); the author would nonetheless like to see if this situation is same as in Asia.

The following research questions are addressed: (1) Can adult EFL learners at university level improve their English language competence when using appealing and inspiring materials as supplemental reading for pleasure? (2) Can male or female students benefit more from pleasure reading?

Literature Review

Theoretical Framework

The important aspects of the theoretical framework presented here are the main components of inspiring and appealing materials for pleasure reading and language competence.

Pleasure reading materials are those light materials, such as a popular magazine, a daily newspaper, or a book of easy short stories (Sapayan, 2009) that could make reading enjoyable, interesting and applicable to the adult English as a foreign language (EFL) learners at university level. Ujiie and Krashen (1996) suggest that comic book reading and other kinds of light reading may serve as an important bridge from everyday "conversational" language to what Cummins (1991) terms "academic language." Reading for pleasure is believed to be the best way to improve students'

command of foreign languages even in the midst of so much time that is devoted to their learning (Lasagabaster, 2008).

On the basis of the claims delineated above, the present investigation with pleasure reading was focused on inspiring and appealing materials such as *Guideposts* as part of the subjects' supplemental reading. *Guideposts* is a popular magazine which provides hope, encouragement, and inspiration to its readers. Through the forms of magazines, books, websites, etc., *Guideposts* helps people deepen their faith and inspires them to reach their true potential (*Guideposts*, n.d.). Though religious in nature, its wording and content is very easy to read and is filled with uplifting, touching and inspiring stories and quotes that could easily motivate readers and EFL learners to keep moving forward in their reading. Grabe (1991), a prominent scholar in the field of theory and practice of ESL/EFL reading and learning, identified reading as an essential skill and probably the most important skill for ESL[/EFL] language learners to master in academic contexts. Schiefele (1999) also illustrated that personal interest [in a long-term intrinsic interest] is a significant predictor of comprehension and learning from reading texts (as cited in Grabe, 2004). If the EFL learners can activate their intrinsic motivation, read over and over again, and put emphasis on essential reading skills, their English language competence will advance to a higher level.

Language competence basically includes linguistic or grammatical competence, discourse competence, and textual competence. It is best developed in the context of learning activities where the language is used for real purposes, e.g., reading, pleasure reading, and/or writing (North, 2000). The practical applications are the vocabulary, grammatical structures, and text form learning (North, 2000). In this study, the focus of the language competence was on the expansion of vocabulary, grammatical competence and fluent comprehension.

The Research Literature

There is compelling evidence supporting the premise that pleasure reading is important in the development of ESL/EFL learners' English proficiency. Some examples have been described above. Other examples include Constantino's (1995) and Gallik's (1999) research. Constantino references Krashen's theory of the benefits of free voluntary reading and the positive impact pleasure reading has on the results of the Teaching of English as a Foreign Language (TOEFL) test. Krashen (1984)

compared the influence of pleasure reading habits on college freshmen essay writing, i.e., whether they were rated either highly competent or of low competence. Those whose essay writing was highly competent reported more pleasure reading at all ages, especially in high school. Gallik (1999) conducted a study to find how much time college students spent reading for pleasure and to determine whether there was a relationship between the amount of recreational reading and academic achievement. The findings revealed a strong and positive relation between the two. First, students' recreational reading habits are helpful in making predictions about their future academic success; second, students who spend more time reading have superior academic skills; third, spending time in recreational reading benefits individuals in their ability to sustain attention and concentration for academic success.

Consistent with Gallik's study (1999) that found a positive relationship between pleasure reading and academic success, Sarangi, Nation, and Meister (as cited in Krashen, 1993) correspondingly demonstrated the effect of pleasure reading on vocabulary development. As Gilbert and Fister (2010) stated, "Reading for pleasure has been associated with creativity and with improved academic achievement (p. 2)," the author was thinking of applying the valuable leisure learning concepts to the adult EFL learner at the undergraduate level and seeing whether students in Taiwan could benefit from this readily-accessible and enjoyable language development approach.

In addition, whether male or female students could benefit more from pleasure reading also presented an important variable in the current study. Some research related to gender difference has been briefly delineated. Marinak and Gambrell (2010) conducted an investigation of gender differences in all readers, including 288 third-grade, average readers. The results suggested that boys and girls who are average readers are equally self-confident about their reading ability; however, boys value reading less than girls. Logan and Johnston (2009) explored gender differences in the relationship between reading ability, frequency of reading and attitudes and beliefs relating to reading and school. The subjects included 232 ten-year-old children (117 male and 115 female). The results indicated that girls were significantly better at reading than boys. Girls read more frequently and had a more positive attitude toward reading and school. Chiu and Mcbride-Chang (2006) carried out a macro study including 199,097 fifteen-year-old learners among 43 countries to complete a reading comprehension test and a questionnaire. Study results indicated that in every country, girls outscored boys; reading enjoyment mediated 42% of the gender effect. In terms

of reading enjoyment, a student who enjoys reading will tend to be a better reader relative to someone who does not enjoy reading. Girls are more likely to enjoy reading, and a culture of peers who enjoy reading is relatively strongly associated with reading performance itself.

Methods

Subjects

The sample consisted of the first and second year undergraduate students enrolled in the department of Applied English at a technology and science university in the northern part of Taiwan. A total number of 90 subjects were randomly recruited and administered the pre-test and, at the completion of the treatment, post-test instrument. Twenty-six were not included in the data analysis due to the fact that some subjects only took either the pretest or posttest, some subjects marked the same answer for every question on the answer sheet, and some subjects skipped numerous items on the answer sheet. These data were then excluded from the study. The effective subjects became 64 students. The mean age of these students was 19.31 years ($M = 19.31$, $SD = 1.22$), and their English proficiency was mixed, but largely from pre-intermediate to intermediate level based on the scores of their English entrance exams. There were 21 male (33%) and 43 female (67%) students in the sample. The uneven sample size in the current study reflects the overall ratio of male to female students in the humanities department. The subjects were randomly assigned into two groups - an experimental group and a control group. The experimental group of 31 ($N = 31$) students was instructed to read the appealing and inspiring materials, such as *Guideposts*, as their after-school assignment to improve their vocabulary, grammar, sentence structure, and fluent reading. The control group containing 33 students ($N = 33$) was instructed via the traditional curriculum. Both groups met once a week for 100 minutes of instruction, and the treatment lasted for eight weeks.

Instruments

The data in the study were collected through three instruments: (1) metacognitive awareness of reading strategies inventory, (2) pre-test, and (3) post-test.

The Metacognitive Awareness of Reading Strategies Inventory ("MARSI" from hereon: see Appendix A for partial text in English & Appendix B for partial text in

Chinese) was used to assess students' metacognitive awareness in reading comprehension (Mokhtari & Reichard, 2002; Sheorey & Mokhtari, 2001). The MARSI questionnaire was related to a good reader's skills which consist of reading with a purpose in mind, previewing, re-reading, organizing text, context clues, summarizing, visualizing images, guessing meaning and prior knowledge. This questionnaire can help readers to understand what they need to do to improve their comprehension while they are reading. The administration of the MARSI in this study was purposed to help EFL learners at the undergraduate level to have a common resource to use in order to facilitate their comprehension while they are reading for pleasure in the treatment. Therefore, the MARSI was administered at the very beginning of the study for both groups. All subjects in the study were given the Chinese version of the MARSI (see Appendix B).

The pretest and posttest were adapted from the reading exercise book (Lougheed, 2006), which included many reading passages, for example, interesting advertisements, daily communication emails, and many other examples found in authentic environments to fit the format purpose. They were used to measure differences between the experimental and control groups at the beginning and the end of the treatment period and they were not the same versions. The equivalence between the pretest and posttest was judged through a separate pilot study (Fan, 2009). The value for coefficient alpha (Cronbach's Alpha) was .96, which indicated that both tests are of satisfactory reliability (Fan, 2009).

The Levence's test of equality of variance and the independent-samples t test were conducted to evaluate whether the means for the two independent groups were significantly different from each other (see Table 1). According to Green and Salkind (2005), if the Levene's Test for Equality of Variances was significant ($p < .05$), the two variances (pretest and posttest) were significantly different. If the Levene's Test for Equality of Variances was not significant ($p > .05$), the two variances were not different, which means the two variances were approximately equal (p. 171). The significance of .11 that we found was greater than .05, which leads to the conclusion that the two variances were approximately equal. The t (25) = .43, p = .67 also indicated that the two forms of pretest and posttest were not significantly different from each other but they were equivalent.

Table 1. The Levene's Test of Equality of Variance and *t* Test for the Equivalence
Between the Pretest and Posttest

Levene's Test for Equality of Variances	*t test* for Equality of Means		
Sig.	*t*	df	Sig.
Equal Variances assumed .11	.43	25	.67

Procedures

The purpose of the study was to investigate whether using appealing and inspiring materials for supplemental reading for pleasure had a positive impact on the English language competence of adult EFL learners at university level. The emphasis of the study was put on the pleasure reading of *Guideposts* and some other easy-read short stories. At the beginning of the fall semester of the year, the instructor (also the author) provided both groups with traditional curriculum, which included reading selections, comprehension questions, grammar exercises and vocabulary. Yet, while the control group continued with the traditional instruction, the experimental group was required to read *Guideposts* as supplemental reading after school. Students were asked to read several passages selected by the instructor, who thought these passages might be inspiring, interesting, and readable for the students. Students were also instructed to underline or mark the key words that they did not understand, and take notes on the sentence structure and the quotes that they preferred (see Appendix C for an example of an assignment) when they read these stories after school. Each passage ranged from around 500 to 1900 words (for example, *The Wedding Planner* - 576 words, *Rediscovering Faith* - 1137 words, *Thank You* - 1253 words, *An Angel in Her Father's Eyes* - 1488, *Gram Faith* - 1577, *My Truest Hope* - 1625 words, *Operation Haiti* - 1773 words, *Hope for Love* - 1906 words, etc.) During class sessions, the subjects were divided into groups to discuss and express thoughts about their reading.

The detailed administration procedure for data collection is presented below: 1) The MARSI was tested in both groups at the beginning of the treatment. 2) Following the MARSI questionnaire, the pre-test was administered, which took about one hour for the subjects to complete. 3) The treatment period was extended to 8 weeks of class sessions (100 minutes' instruction per week). 4) The post-test was administered after the treatment taking about one hour for completion.

Data Analysis

A one-way analysis of variance (ANOVA) was employed to assess whether the adult EFL learners at university level can improve their English language competence when using appealing and inspiring materials as supplemental reading for pleasure. Thus, in this study, there is one independent variable – the amount of emphasis placed on pleasure reading between the experimental and control groups - and one dependent variable - the posttest scores.

Moreover, gender was examined through descriptive statistics to see whether male students or female students can benefit more from the appealing reading and whether "gender difference" (Green & Oxford, 1995) has a meaningful impact on pleasure reading results.

Results

Research Question One

The first research question was addressed as follows: (1) Can adult EFL learners at university level improve their English language competence when using appealing and inspiring materials as supplemental reading for pleasure? Is there a difference between the experimental group that received appealing and inspiring materials as supplemental reading for pleasure and the control group that received no further reading for pleasure? The resultant analysis through the descriptive statistics and one-way analysis of variance was presented in Table 2. The mean of the experimental group $M = 40.58$, $SD = 11.76$ was greater than the mean of control group $M = 34.42$, $SD = 8.57$, which indicated that the experimental group subjects performed better than the control groups in the posttest. The experimental group of university level adult EFL learners improved their English language competence when employing appealing and inspiring materials as supplemental reading for pleasure. Standard deviation (SD) presents variability of scores about the mean of a distribution so the greater the standard deviation, the greater the spread of scores (Shavelson, 1995). The results for the ANOVA indicated a significant main effect for the groups. The $F (1, 62) = 5.78$, $p = .019 < .05$, partial η^2 (Partial Eta Squared) = .085 indicated that there was a significant difference between the experimental group and the control group. The partial $\eta^2 = .085$ indicated the group was likely medium in effect sizes. Partial η^2 ranges in value from 0 to 1. The conventional rule is .01, .06 and .14 for small,

medium, and large effect sizes, respectively (Green & Salkind, 2005, p. 187).

Table 2. The Descriptive Statistics of Mean Scores and the One-Way Analysis of
Variance

Groups	Mean	Std. Deviation	N
Experiment	40.58	11.76	31
Control	34.42	8.57	33

Source of Variation	SS	*df*	MS	*F*	*p*-value (Sig.)	Partial η^2
Group	605.83	1	605.83	5.78	.019	.085
Error	6497.61	62	104.80			
Total	7103.44	63				

Research Question Two

The second research question was addressed as follows: (2) Can male or female students benefit more from pleasure reading? The results of mean difference and standard deviation were presented in Table 3, which indicated that the mean of female students $M = 38.84$, $SD = 10.80$ was greater than the mean of male students for $M = 34.48$, $SD = 9.83$, which indicated that the female subjects performed better than male subjects in pleasure reading learning.

Table 3. The Descriptive Statistics of Mean Scores for Genders

Groups	Mean	Std. Deviation	N
Female students	38.84	10.80	43
Male student	34.48	9.83	21

Additionally, the MARSI administered at the beginning of the study was to help the subjects to obtain a reflective direction to facilitate their comprehension while they are reading. The mean of the MARSI for male students was $M = 3.56$, $SD = 0.47$, which was higher than the mean of the female students for $M = 3.11$, $SD = 0.52$ (see Table 4). This means that the male students who rated their MARSI reading ability as excellent had a lower use of the guidance of the metacognitive awareness strategies. Conversely, the female students who rated their MARSI reading ability as

bad have a higher use of the guidance of the metacognitive awareness strategies. Overall, the purpose of the MARSI instrument in the current study was to function as the management of the perception for the metacognitive reading strategies while reading for pleasure to assist in readers' English language competence elevation.

Table 4. The Descriptive Statistics of Mean Scores for MARSI Test

Gender	Mean	Std. Deviation	N
Female students	3.11	.52	43
Male student	3.56	.47	21

Discussion and Implication

The results obtained from the study revealed a significant difference between the experimental group and the control group on the posttest reading scores. The adult EFL learners at the undergraduate level who were assigned to read inspiring and appealing articles, (e.g., *Guideposts*) as supplemental reading for pleasure positively outperformed the control group on the reading comprehension posttest scores. This study suggests that reading for pleasure is a way to motivate EFL adults to learn English. When students employed inspiring and appealing materials and/or other short easy-read stories as part of their supplemental reading, they not only experienced greater improvement in their reading comprehension and their English language competence, but also increased their confidence and interest in learning English. The more that students incorporate opportunities to read for pleasure into their daily routines, the more they may improve their academic competence and academic achievement as well.

In terms of the second research question: Can male or female students benefit more from the pleasure reading? The mean and standard deviation obtained from the male and female students' performance on the posttest reading scores revealed that female students benefited more than the male students from pleasure reading. This finding is consistent with previously stated studies on literature. Female students are more likely to get involved and interested in the inspiring and touching stories and they closely followed the reading of the appealing, inspiring and touching stories they read outside of class (Chiu & Mcbride-Chang, 2006; Logan & Johnston, 2009; Marinak & Gambrell, 2010).

Implications

The positive results in the current study revealed several implications. First, reading for pleasure can be a natural and easy way to "turn the texts as a favored source of information" (Ross, 1999, p. 787) to foster adult EFL learners' language competence. Second, reading for pleasure can play a positive role in improving EFL learners' academic performance, enhance their range of background knowledge as well as develop their active vocabulary. In the same way as Krashen (2004) claimed that everything that is fun is good for literacy development and language competence, and the path of pleasure is the only one that works, rather than the current view of focusing on form and explicit direct instruction to gain language competence. Third, reading for pleasure can be interwoven into the texture of their lives (Ross, 1999). As researchers stated that teachers should help students build their positive attitudes towards reading and a love of books (Clark & Rumbold, 2006; Ross, 2000). Finally, reading for pleasure can be developed to increase the EFL learners' job and career development, especially in responding to changes of social context (Kirsch & Guthrie, 1984).

Regardless of the fact that the current study yielded significant results, there were some limitations. First, the subjects in the study were recruited via uneven distribution by gender (Male=21, Female=43), which could limit external validity, i.e., the generalizability might not be testable. Second, even with the fact that the pretest and posttest equivalence had been established, both tests were not standardized reading tests, which made it difficult to completely test students' reading comprehension abilities. Third, subjects still encountered many difficulties on the vocabulary recognition in both the context and in elements of the main idea comprehension. Thus, the experiment might need to be reexamined to further improve students' comprehension performance.

Expansion for future study is recommended in order to explore the correlation between pleasure reading and academic achievement in an attempt to advance understanding of the value of reading for pleasure and to provide information that will enable parents and teachers to get involved in promoting reading motivation and wider reading. Another trial access will be recommended to monitor students' further achievement at a higher educational level.

References

Anderson, R. C., Hiebert, E. H., Scott, J. A., Wilkinson, I. A., & Becker, W. C. (1988). Becoming a nation of readers: The report of the commission on reading. *Education and Treatment of Children,* 389-396.

Chen, S. Y. (2007). Extracurricular reading habits of college students in Taiwan: Findings from two national surveys. *Journal of Adolescent & Adult Literacy*, *50*(8), 642-653.

Chiu, M. M., & McBride-Chang, C. (2006). Gender, context, and reading: A comparison of students in 43 countries. *Scientific Studies of Reading*, *10*(4), 331-362.

Clark, C., & Rumbold, K. (2006). Reading for pleasure: A research overview. London, United Kingdom: National Literacy Trust.

Constantino, R. (1995). The Effects of pleasure reading. *Mosaic: A Journal for Language Teachers*, *3*(1), 15-17.

Cummins, J. (1991). Language development and academic learning. *Language, culture and cognition*, *161*, 75.

Fan, H. C. S. (2009). *The effectiveness of metacognitive strategies in facilitating Taiwanese university learners in EFL reading comprehension*. Unpublished doctoral dissertation. University of Kansas, Lawrence, USA.

Gallik, Jude. D. (1999). Do they read for pleasure? Recreational reading habits of college students. *Journal of Adolescent & Adult Literac*y, *42*(6), 480-488.

Gilbert, J., & Fister, B. (2010). Reading, risk, and reality: College students and reading for pleasure. *College & Research Libraries*, l-148.

Grabe, W. (1991). Current developments in second language reading research. *TESOL Quarterly*, *25*(3), 375-406.

Grabe, W. (2004). 3. Research on teaching reading. *Annual Review of Applied Linguistics*, *24*, 44-69.

Green, J. M., & Oxford, R. (1995). A closer look at learning strategies, L2 proficiency, and gender. *TESOL Quarterly*, *29*(2), 261-297.

Green, S. B., & Salkind, Neil J. (2005). *Using SPSS for Windows and Macintosh analyzing and understanding data* (Fourth ed.). Upper Saddle River, New Jersey 07458: Pearson Education, Upper Saddle River, New Jersey.

Guideposts. (n.d.). Retrieved July 28, 2015 from https://www.guideposts.org/

Howard, Vivian. (2013, October). Teens and pleasure reading: A critical assessment

from Nova Scotia. *Proceedings of the Annual Conference of CAIS/Actes du Congrès Annuel de l'ACSI*. (pp. 1-11). Canada.

Janopoulos, Michael. (1986). The relationship of pleasure reading and second language writing proficiency. *TESOL Quarterly*, *20*(4), 763-768.

Kirsch, I. S., & Guthrie, J. T. (1984). Adult reading practices for work and leisure. *Adult Education Quarterly*, *34*(4), 213-232.

Krashen, S. D. (1981). Bilingual education and second language acquisition theory. *Schooling and language minority students: A theoretical framework*, 51-79.

Krashen, S. D. (1984). *Writing, research, theory, and applications*. Oxford: Pergamon.

Krashen, S. D. (1993). The Case for Free Voluntary Reading. *Canadian Modern Language Review*, *50*(1), 72-82.

Krashen, S. D. (2004). *The power of reading: Insights from the research (Second Edition).* Santa Barbara, CA: ABC-CLIO Corporate.

Lasagabaster, D. (2008). Foreign language competence in content and language integrated courses. *The Open Applied Linguistics Journal*, *1*(1), 30-41.

林家慈. (2009). *綜合高中學生課外閱讀現況之研究—以台東縣為例. 臺東大學進修部暑期學校行政碩士班學位論文.*

[Lin, Jia Ci. (2009). Study on the current situation of leisure reading among comprehensive high school students — case study on Taitung county. (Master's thesis, In-service master program, National Taitung University)].

Logan, S., & Johnston, R. (2009). Gender differences in reading ability and attitudes: Examining where these differences lie. *Journal of Research in Reading*, *32*(2), 199-214.

Lougheed, L. (2006). *Barron's TOEIC Test 4/e.* Barron's Educational Series.

Marinak, B. A., & Gambrell, L. B. (2010). Reading motivation: Exploring the elementary gender gap. *Literacy Research and Instruction*, *49*(2), 129-141.

Mokhtari, K., & Reichard, C. A. (2002). Assessing students' metacognitive awareness of reading strategies. *Journal of Educational Psychology*, *94*(2), 249.

North, B. (2000). *The development of a common framework scale of language proficiency*. New York, NY: Peterlang.

Parlette, M., & Howard, V. (2010). Personal growth, habits and understanding: Pleasure reading among first-year university students. *Evidence Based Library*

and Information Practice, *5*(4), 53-69.

Ross, C. S. (1999). Finding without seeking: The information encounter in the context of reading for pleasure. *Information Processing & Management*, *35*(6), 783-799.

Ross, C. S. (2000). Making choices: What readers say about choosing books to read for pleasure. *The Acquisitions Librarian*, *13*(25), 5-21.

Sapayan, M. S. (2009). The Pleasure of Reading. Retrieved from http://ezinearticles.com/?The-Pleasure-of-Reading&id=2474444

Schiefele, U. (1999). Interest and learning from text. *Scientific Studies of Reading*, *3*(3), 257-279.

Shavelson, R. J. (1995). *Statistical reasoning for the behavioral sciences.* Boston, MA: Allyn and Bacon.

Sheorey, R., & Mokhtari, K. (2001). Differences in the metacognitive awareness of reading strategies among native and non-native readers. *System*, *29*(4), 431-449.

Thorndike, R. L. (1973). *Reading comprehension education in 15 countries* (International Studies in Evaluation III). New York: Halsted Press.

Ujiie, Joanne, & Krashen, Stephen D. (1996). Comic book reading, reading enjoyment, and pleasure reading among middle class and Chapter I middle school students. *Reading Improvement*, *33*, 51-54.

Appendix A.

Metacognitive Awareness of Reading Strategies Inventory (Version 1.0)

Directions: Listed below are statements about what people do when they read *academic or school-related materials* such as textbooks or library books.

Five numbers follow each statement (1, 2, 3, 4, 5), and each number means the following:

After reading each statement, **circle the number** (1, 2, 3, 4, or 5) that applies to you using the scale provided. Please note that there are **no right or wrong answers** to the statements in this inventory.

Type	Strategy	Scale
GLOB	1. I have a purpose in mind when I read.	1 2 3 4 5
SUP	2. I take notes while reading to help me understand what I read.	1 2 3 4 5
GLOB	3. I think about what I know to help me understand what I read.	1 2 3 4 5
GLOB	4. I preview the text to see what it's about before reading it.	1 2 3 4 5
SUP	5. When text becomes difficult, I read aloud to help me understand what I read.	1 2 3 4 5
SUP	6. I summarize what I read to reflect on important information in the text.	1 2 3 4 5
GLOB	7. I think about whether the content of the text fits my reading purpose.	1 2 3 4 5
PROB	8. I read slowly but carefully to be sure I understand what I'm reading.	1 2 3 4 5
SUP	9. I discuss what I read with others to check my understanding.	1 2 3 4 5
GLOB	10. I skim the text first by noting characteristics like length and organization.	1 2 3 4 5

Scoring Rubric

Student name: _____ Age:_____ Date:_____
Grade in school: ▫ 6th ▫ 7th ▫ 8th ▫ 9th ▫ 10th ▫ 11th ▫ 12th ▫ College ▫ Other
_____

Global Reading Strategies (GLOB subscale)	Problem-Solving Strategies (PROB subscale)	Support Reading Strategies (SUP subscale)	Overall Reading Strategies
_____GLOB score	_____PROB score	_____SUP score	_____Overall score
_____GLOB mean	_____PROB mean	_____SUP mean	_____Overall mean

Key to averages: 3.5 or higher _ high 2.5–3.4 _ medium 2.4 or lower _ low
_____

Appendix B.

後設認知策略理解問卷調查

學生姓名:_____ 年齡:_____ 日期:_____

指示：下列的說明是關於一般人如何閱讀讀學校相關書籍, 例如教科書或圖書館的書. 數字代表依照順序的排列:

- 1 "我從來或幾乎沒有讀過."
- 2 "我偶爾會讀."
- 3 "我有時會讀." (大概佔 50%的時間)
- 4 "我通常都會讀."
- 5 "我總是或幾乎都讀."

請將以下的說法讀過之後, 將你的範圍程度(1, 2, 3, 4, or 5)圈起來.

請注意這些說法在以下的清單裡是沒有對與錯的答案.

種類		策略						等
級程度								
GLOB	1.	我閱讀是有一定的目標的.		1	2	3	4	5
SUP	2.	我閱讀時會做些筆記幫助我了解意思.		1	2	3	4	5
GLOB	3.	我會思考我所知道的去幫助自己了解閱讀內容.	1 2 3 4 5					
GLOB	4.	在閱讀之前我會先預習以了解本文內容.		1	2	3	4	5
SUP	5.	當本文更難時, 我會唸出聲幫助我了解.		1	2	3	4	5
SUP	6.	我會總結我讀的本文來點出重點.		1	2	3	4	5
GLOB	7.	我會思考我讀的本文是否符合我的目的.		1	2	3	4	5
PROB	8.	我讀的很慢但我很小心在讀, 確認了解.		1	2	3	4	5
SUP	9.	我會與同學討論我讀的內容來確認了解.	1 2 3 4 5					
GLOB	10.	我會先略讀本文並注意內容特性,像是長度及組織架構.		1	2	3	4	5

.

.

.

* GLOB: 全球閱讀策略(如設定閱讀目的, 啟動先前知識, 瀏覽略讀理解閱讀內容)

* PROB: 問題處理策略(調整閱讀速度, 再閱讀, 想像故事內容)

* SUP: 支持閱讀策略(做筆記, 重點劃線, 自己問問題)

全球閱讀方案	問題處理方案	支持閱讀方案	全盤的閱讀方案
_____GLOB 分數	_____PROB 分數	_____SUP 分數	_____總分
_____GLOB 平均	_____PROB 平均	_____SUP 平均	_____總平均

關鍵總平均: 3.5 或更高 = 高分 2.5-3.4 = 中等 2.4 或更低 = 低分

Appendix C.

Assignment for the Pleasure Reading

Name: _____ Date: _____

Directions: Answer each question below:

1) Identify the title, author, and main points. For example: In the passage "The Kindness of Angels," the author, Roma Downey, and her main points in the passage …	2) Highlight the important information, quotes or phrases in the passage.
3) Write down the words and phrases that you don't understand and give an explanation with it in English and Chinese.	4) Make an outline to check whether you have covered the important points.

Investigating Grade-Nine EFL Teachers' and Learners' Beliefs towards CLT and Perceived Difficulties in Implementing Communicative Language Teaching (CLT) in EFL: The Case of Three Debre Markos Secondary Schools, in Ethiopia

Mihretu Yihunie Yalew
Department of English Language and Literature, Bahir Dar University, Ethiopia

Bioprofile:
Mihretu is an instructor of English Language and Literature at Bahir Dar University, Ethiopia. He has his MEd. from Bahir Dar University in TEFL. He has also taught English as EFL for many years at different levels. His research interests include language teaching, learning and assessment. Email: mihretuyihunie@yahoo.com

Abstract

The study is aimed at investigating beliefs and conceptions, and major CLT implementing difficulties in three secondary schools of Debre Markos, in Ethiopia, which is an EFL setting. Eighteen grade-nine English language teachers and eighty-four students of the same schools were participants of the study. The researcher used comprehensive sampling for teachers and simple random sampling for students. Questionnaires and observation checklists were major data gathering tools. As the data lend itself for both qualitative and quantitative data analysis, mixed-method approach was employed for the data analysis and interpretation. The collected data from both groups was tallied and tabulated in categories. It was also analyzed and interpreted using T-test, descriptive and inferential statistics. In doing so, 2.5 has used as an expected mean score to determine a statistically significant mean difference of each item in the category as well as the grand mean score of that category. Finally, the major findings of the research can be generalized, though most teachers of English language in the study area do not have serious misconceptions and wrong beliefs about CLT, their classroom practices are entangled with CLT implementing difficulties in their endeavor of developing students' communicative competence in

the target language (English). Hence, there is also a need to reconsider the Ethiopian existing educational setting for implementing CLT since it has been a method /an approach designed primarily for ESL settings.

Introduction

Background of the study

Communicative Language Teaching (CLT) as a methodology was first proposed in England in the 1970s. This methodology was regarded as revolutionary since it placed emphasis on communication in language learning classrooms. Being primarily an ESL (English as a second language) methodology, it rapidly gained wide acceptance in the western countries. Following the emergence of CLT in the English speaking countries, it began to spread all over the world signifying the new and being endorsed as a reaction against the traditional language teaching methodologies, CLT has served as a major source of influence on English language teaching practice in both ESL and EFL environments.

Even though national policies and school curricula shifted toward CLT in EFL contexts, researchers have pointed out a gap between policy and practice (Littlewood, 1981; Nunan, 2003). The implementation of CLT has encountered problems and resistance in EFL classrooms (Ellis, 1996; Yu, 2001). Various research studies reported that instruction in EFL classrooms was still based on the traditional approach (Littlewood, 1981; Nunan, 2003; Savignon & Wang, 2003). The ideas of CLT are different from the educational values and traditions of many EFL settings (Burnaby & Sun, 1989; Li, 1998). In addition, situational factors such as large-sized class, test-oriented instruction, and students' low proficiency also influenced CLT practices. Even if the policies and curricula support the adoption of CLT, only the teachers can decide what really happens inside their classrooms. Thus, it is essential to learn the teachers' views regarding the implementation of CLT.

The English language teachers at each level are required to apply the new approach in the teaching learning process. Since the former ways of language teaching were not able to bring the desired results in students learning, the need to apply communicative approach became a necessity in language teaching methodologies.

Following this paradigm shift, the communicative approach of language teaching (CLT) was accepted to be used in the Ethiopian context too.

Background Information on Education in Ethiopia

English is taught in Ethiopia as a subject from grade one up wards. It is also a medium of instruction for secondary schools and tertiary levels. However, the vast majority of students in secondary schools lack the fundamental communicative competence, which is the goal of the present language syllabus (curriculum) (ELIP, 2004). The communicative approach in the English curriculum of secondary schools is likely to be entangled with implementation and assessment problems. The significant role that English language plays in various sectors of human life (social, economic and academic) has initiated a great need for good communication skills of this language around the globe (Harmer, 2001). The need for good communication skills in English, in turn, has created an immense demand for quality in English language teaching. Learners are interested in mastering English and many employers want their employees to have good English language skills (Richards, 2006). In Ethiopia, there is also a large demand for good English command as it gives a wide access to employment, opens the doors for both local and international scholarship opportunities and provides access to most scientific materials. Moreover, as an international lingua franca, it serves as a medium of communication with the international community. Although English serve as a medium of instruction at the secondary and higher education levels in Ethiopia, the quality of its teaching has been challenged by various factors (MoE, 2005).

In contrast to traditional grammar teaching, CLT has been viewed as an effective and widely accepted approach to language teaching (Gray & Klapper, 2009; Harmer 2001; Richards, 2006). CLT encourages interactive language learning through the process of real life communication in a meaningful way (Richards, 2006). Therefore, CLT has widely been accepted and applied in teaching languages in various parts of the world (Richards, 2006). Similarly, in Ethiopia, it has been found to be a promising approach to improve English language teaching (Lakachew, 2003). Hence, to understand the implementation of CLT in Ethiopian secondary schools, the teachers and students beliefs about it and their perceived difficulties in implementing it in EFL are investigated in this research article.

Statement of the problem

Though the curriculum, the English syllabus, and teaching materials for secondary schools in Ethiopia are prepared in a communicative way, the learners' communicative competence in the target language is still under question. Therefore, investigating how communicative language teaching is being implemented, finding perceived difficulties in implementing CLT and forwarding sound and reliable recommendations based on the finding is the central issue of this research.

Objectives of the Study

The objectives of the study are:

● Assessing teachers' and learners' beliefs towards communicative language teaching (CLT);

● Investigating CLT implementing difficulties faced by EFL teachers and learners in the study area;

● Forwarding sound recommendations based on findings.

Review of Related Literature

An Overview of Teaching English in Ethiopia

With the introduction of modern education in the early 20th century, the foreign languages that were operating in the country were mainly Italian, French, and English. With the British assistance to Ethiopian to fight against the occupying Italian army, the emperor of Ethiopia and the people in general were grateful for the assistance they got from Britain. Aleka Lemma (Mengistu's father) said that we had to learn their language because they were the ones who helped us expel the Italians (Mengistu, 1996).

Almost all of the curriculum and teaching materials, especially for high schools, were imported from Britain. Ethiopian students were encouraged to sit for the same examinations, as would students in Britain and allowed to sit for the college entrance examination in the USA.

A reintroduction decision was made to teach English as a subject and make it the language of instruction in most subjects commencing with the fifth year (Markakis, 1974:174). Thus, English was used as a medium of instruction in primary schools, until Amharic replaced it in the early 1960s. When the Amharic policy was put into place, English became the medium of instruction for most of the subjects taught in

grades seven and eight. English was also the language of instruction at secondary level and in teacher training institutions.

Thus, the students have been expected mainly to learn English and use it as a medium of communication. Some scholars have questioned the extent to which students can be successful in using English as a medium of instruction. Based on his field survey in the mid- 1980s, Stoddart had the following to say about the ability to use English language by the vast majority of students in Ethiopia, "Unless students participate actively through their own speaking and writing, they do not possess sufficient knowledge of English to understand what they hear from their teachers and read in their text-books. Moreover, they spend most of their class hours copying down notes that the teacher has written on the blackboard." (Stoddart, 1986: 6-7) as cited in Bogale (2007). As a result, the adoption of communicative language teaching as an approach came in to being.

Communicative language teaching (CLT) is a recognized theoretical model in English language teaching in Ethiopia today. Many applied linguists regard it as one of the most effective approaches to ELT. Since its inception in Europe in the early 1970s, CLT has served as a major source of influence on language teaching practices around the world. As Li (1998) comments, CLT has extended in scope and has been used by different educators in different ways. Even though national policies and school curricula shifted toward CLT in EFL contexts, researchers have pointed out a gap between policies and practices (Littlewood, 1981).

The following are also believed to be the major factors that hinder CLT from being fully implemented in Ethiopia, which is an EFL context.

Teachers' and Learners' Beliefs and Conceptions towards CLT. Beliefs are important factors in determining the success of any teaching and learning. Beliefs are a central construct in every discipline that deals with human behavior and learning (Ajzen, 1988; Fishbein & Ajzen, 1975). Learner views of learning cannot be ignored, in particular, when there is a mismatch between teacher beliefs and learner beliefs (Schulz, 1996). Similarly, teachers' and learners' beliefs towards CLT matter to its successful implementation. Understanding the attitudes, beliefs, assumptions, preferences, and needs of the next generation is inevitable if EFL instruction is really the aim of empowering students linguistically and culturally (Savignon, 2007). Clear

understanding of learners' attitudes and perceptions of CLT as a widespread framework in shaping current definitions of the goals set for EFL teaching is quite important to help learners attain these goals (Savignon & Wang, 2003). Therefore, investigating the existing situation of teachers and learners beliefs and conceptions towards CLT was found to be an important part of this research.

Contextual Differences between ESL and EFL Environment. Since ESL learners have the chance to continue learning English outside the classroom through interactions in their everyday lives, learning the target language for these students is more than a curriculum discipline, it is rather part of survival. Therefore, ESL learners need to learn the language to survive and grow (Ellis, 1996). The English classroom in ESL settings typically functions the principle of immersing the learners in the target language society. Hiep (2007) argues that it is, therefore, essential in the ESL classroom to establish what Halliday(1994:54) calls "the optimum interactional parameters", within which, learners, by interacting with each other on meaningful things, can best develop the communicative skills they immediately use in their real life. Yet, learning English for EFL learners is generally a part of the school curriculum rather than a necessity to survive. Hence, it is usually during class time that EFL students have exposure to English, so they are unable to test and practice strategies as easily (Ellis, 1996).

In addition, learners in an ESL setting generally have different native languages from their peers. This being so, for ESL learners, using the target language becomes salient in interacting and making friends with classmates in and outside the language classroom. As pointed out by Ellis (1996), it is most likely that heterogeneous language classroom produces higher motivation and faster adaption of learning strategies on the part of the learners. On the other hand, EFL learners usually share the native language with their classmates. As a result, they generally feel tempted to use their native language when they need to initiate a conversation in the language classroom.

Barriers in Adopting CLT to EFL Contexts. It has been argued by researchers and writers that taking a set of teaching methods developed in one part of the world and using it in another part brings out problems and challenges (Holliday, 1994; Pennycook, 1989). According to these authors, education is bound to a particular

cultural environment, and good teaching practices are socially constructed in this environment. Accordingly, as cited in Hiep (2007), assuming that what is suitable in one particular educational setting will naturally be suitable in another is to disregard the fact that ELT methodology is rooted in an Anglo-Saxon view of education. Likewise, Phillipson (1992) maintains that since Anglo-American ELT trends lack appreciation of various distinct linguistic, cultural and educational contexts around the world, they cannot thus produce appropriate teaching and learning materials that will address the local and culture specific needs of the learners. The particular context in which an innovation is introduced determines its success or failure.

CLT was initially developed as a western ELT methodology in the 1970s. However, since then, it has been extensively adopted in both ESL and EFL contexts all around the world. Although implementing CLT in EFL context results in a number of problems and challenges, it would be dubious to claim that these problems cancel out its potential as a language teaching methodology in EFL environments. Larsen-Freeman (2000:67) warns that in the battle against imported methods, "We may fail to understand the cause of the problem and run the risk of overacting and losing something valuable in the process".

In addition, in a study that addressed the issue of CLT use in Taiwan, Liu (2005) found out that despite the prevalent popularity of CLT in Taiwan, it was rather difficult to apply CLT to the actual language classroom. Since the education system is mainly exam-oriented in Taiwan, EFL teachers laid a heavy emphasis on preparing their students for the national college entrance examination. They essentially teach grammatical structures of English because the examination largely consists of questions that assess the structural forms of language.

Teachers' and Students' Role in CLT Classroom. The learner-centered characteristic of CLT and the new type of classroom activities imply different roles in the language classroom for teachers and learners than from those found in more traditional second language classrooms. Learners in CLT classrooms are supposed to participate in the classroom activities that are based on a collaborative rather than individualistic approach to learning. They are portrayed as active participants in the language learning process. Therefore, CLT alters the role of the teacher. In addition, CLT as a methodology has much to do with interaction. The teacher uses

communication as a means to reach the goal, which is improving learners' communicative competence. Accordingly, it would be wise to claim that the teachers' and students' roles in CLT classroom are dynamic in nature, and thus they tend to vary all the time.

Breen and Candlin (1980), in defining the role of the teacher in CLT classroom, notes the following central roles:

The first role is to facilitate communication process between all participants and the various activities and texts. The second role is to act as an independent participant within the learning-teaching group. A third role of the teacher is that of a researcher and learner, with much to contribute in terms of appropriate knowledge and abilities, actual and observed experience of the nature of learning and organizational capacities.

Deckert (2004:13), referring to the student-centered characteristics of CLT, emphasizes that "CLT's approach features low profile teacher roles, frequent pair work or small group problem solving, students responding to authentic samples of English, extended exchanges on high interest topics, and the integration of the four basic skills namely, Speaking, Listening, Reading and Writing". He further states that CLT discourages pervasive teachers-controlled drills, quizzing of memorized material, and extensive explanation of forms of English.

Assessment in CLT. Not only the proper implementation of CLT that brings the desired change in our teaching and learning but also the use of proper assessment procedures that can go hand in hand with the objective of the curriculum (course). Hence, its power to have positive backwash effect to the teaching plays a pivotal role for the success of our teaching (Weigle & Jensen, 1997). The opposite will be true if our assessment does not have relevance with the objective of the curriculum. It cannot strengthen our teaching. Assessment can inspire, motivate, and provide corrective feedback, but it can also lead us to ignore what cannot be easily measured (Geoff Petty, 2004). Teaching is, after all, the primary activity, if testing comes in conflict with it, and then it is testing which should go, especially when it has been admitted that so much testing provides inaccurate information (Hughes, 1989). Contrary to these principles and features of testing, implementation and assessment of CLT seems deviating from its objectives. Especially, this is true when teacher-made and national

exams of English are restricted to grammar, vocabulary and reading, marginalizing the other language skills.

Data Presentation, Analysis and Discussion

Questionnaire and observation checklist were the major data gathering tools employed by the researcher. The data gathered by means of questionnaire from both grade-9 English language teachers and students of three Debre Markos Secondary Schools, which is the study area, is presented, analyzed, discussed and interpreted both quantitatively and qualitatively hereunder. It was also triangulated with other similar research findings as well.

For the purpose of data analysis and interpretation in each category, the researcher used t-test (descriptive statistics) and inferential statistics. In doing so, 2.5 served as an expected mean score (a determinant mean score) in order to see a statistically significant difference on teachers' and learners' responses regarding their beliefs and conceptions towards CLT as well as the difficulties faced in implementing it in EFL setting.

Data from Grade-9 Teachers and Learners on Their Beliefs & Conceptions towards CLT

As figure1 below signifies the grand mean score of grade-9 teachers' responses on their beliefs and conceptions towards CLT is above the expected mean score, which is 2.69. From this, we can infer that many teachers do not lack knowledge about CLT as well as their beliefs towards CLT do not deviate from its principles. Conversely, the grand mean score of grade-9 student respondents on their beliefs and conceptions towards CLT is 2.26;which is less than the expected mean score 2.5.This implies that many students' beliefs and conceptions towards CLT are not in line with its principles and hence they have misconceptions about its principles. This in turn affects the success of CLT implementation. In line with this, in the classroom context, the perceptions, beliefs, attitudes, and meta-cognitive knowledge that students bring with them to the learning situation have been recognized as a significant contributory factor in the learning process and ultimate success (Breen & Candlin, 2001). This difference on teachers and students on their beliefs and conceptions towards CLT might be due to the familiarity and exposure of teachers to

those theoretical issues unlike with the students though teachers also found having troubles on the practical and implementation aspects of CLT.

Figure1.Data from both grade-9 teachers & students on their beliefs & conceptions towards CLT

Data from Grade-9 Teachers & Students on Teacher - Related Difficulties

As can be viewed from Figure2 below, the grand mean score obtained from grade-9 teachers on "teacher-related difficulties in adopting CLT" is 2.74, which is above the expected mean score 2.5. Hence, this implies that most teachers have faced implementing difficulties of CLT. Again, the vast majority of students face teacher-related implementing difficulties while CLT is being implemented in EFL classes. This is true as the average mean score of each item in the category is above 2.7 and the grand mean score of the category is 2.85. This result is also found to be in line with the findings of the study carried out by Li (1998), in order to investigate South Korean teachers' perception of CLT in EFL environments.
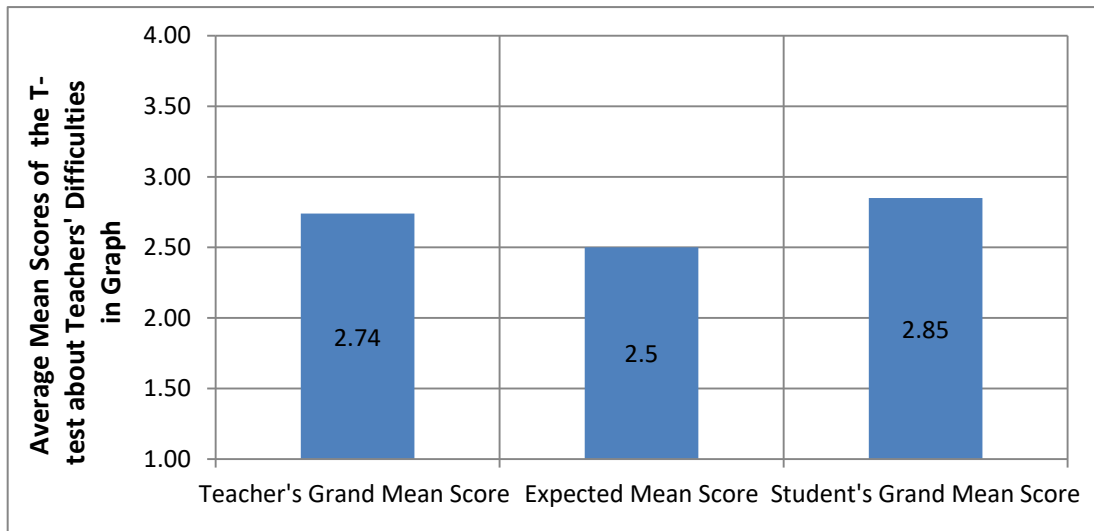
Figure2.Data from both grade-9 teachers & students on teacher - related difficulties

Data from Teachers and Students on Student-Related Difficulties

As can be revealed in Figure3 below, grade-9 teachers' responses on student-related difficulties while implementing CLT has the grand mean score of 3.71, which is statistically significant as it exceeds the expected mean score (2.5) by 1.21. Therefore, one can conclude that student-related difficulties play undisputedly significant role in hindering the success of implementing CLT in EFL setting. Similarly, the grand mean score of the items for student respondents in the category is 2.8, which is again greater than the expected mean score 2.5. This informs us that students themselves do not deny that they are among the sources of difficulty in the implementation of CLT in EFL classes.
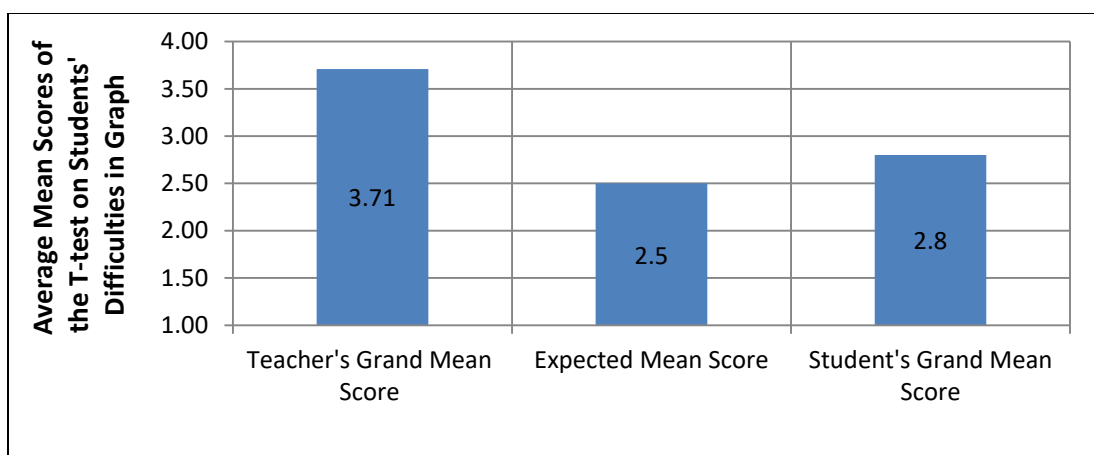

Figure3. Data from both grade-9 teachers & students on student- related difficulties

Data of Grade-9 Teachers and Students on Difficulties Related to the Educational System

As shown in figure 4 below, the grand mean score of teacher and student respondents in the category is 3.4 and 2.91 respectively; both are more significant than the expected mean score. This implies that difficulties related to educational system are prevalent to both grade-9 teachers and students in CLT. In other words, the educational system in practice is not favorable for the successful implementation of CLT or it does not lend itself for CLT. For example, lack of support from the administration and grammar-based assessments are evidences that attribute to difficulties in educational system.
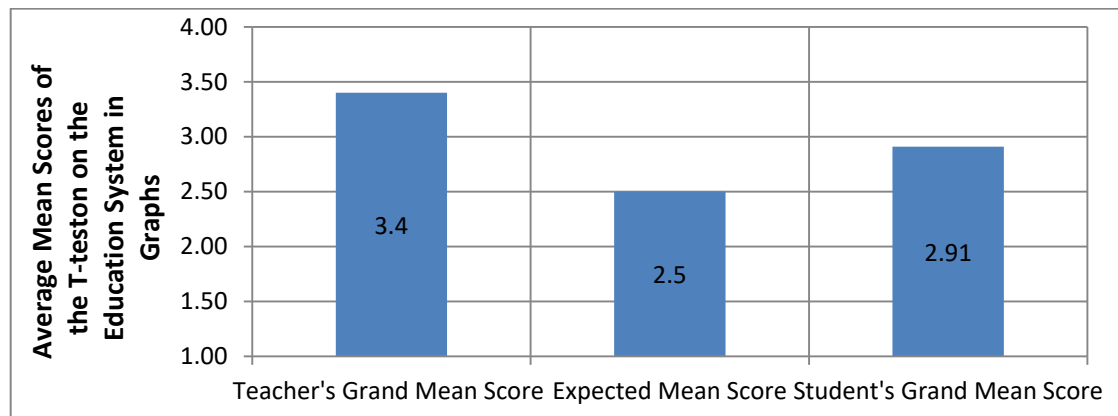

Figure4.Data from grade-9 teachers and students on difficulties related to the educational system

Data from Grade-9 Teachers and Students on Difficulties Related to Assessment

As displayed in figure 5, the grand mean score of teacher respondents on difficulties-related to assessment is 3.4, highly significant from the expected mean score 2.5. In addition, the grand mean score of student respondents of the category is 3.04.
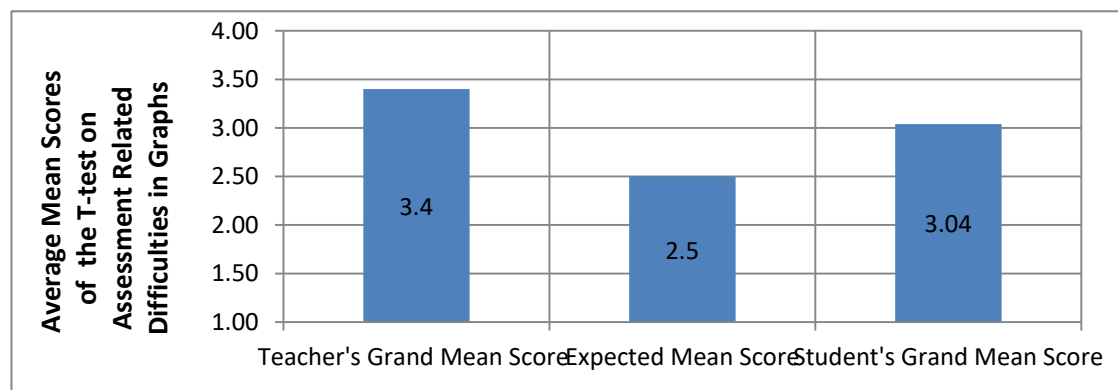

Figure 5. Data of grade-9 teachers and students responses on difficulties related to assessment

Hence, one can deduce that difficulties related to assessment are present challenges to most teachers and students in implementing CLT in EFL setting. This in turn warns us that the assessment procedures and types of assessment we are employing need to be in line with the intentions of the current CLT curriculum so as to narrow the existing gap. To support this idea with evidence obtained from the observation-checklist and experience, many teachers often tend to assess grammar instead of assessing the four language skills. This is again true in national exams, which emphasize on grammar, vocabulary and reading, marginalizing the other language skills.

Findings, Conclusion, Pedagogical Implication, and Recommendations

Findings and conclusion

After analyzing the questionnaire data from both teachers and students as well as from observation checklist, the researcher reached the following findings and conclusions respectively.

The finding of the study has shown that though the beliefs and conceptions of many grade-9 teachers are in line with the principles of CLT, they also have significant problems in their classroom implementation. This has also been claimed by Bataine (2011) and Yu (2001). However, the beliefs and conceptions of many students towards CLT are not in line with the principles. This difference between teachers and students on their beliefs and conceptions towards CLT may have resulted from the fact that the students might not have taken CLT courses that deal with the principles, practical applications as well as its implementation unlike the teachers.

In general, the result from both grade-9 teachers' and students' questionnaire on implementing difficulties in the four categories demonstrated that all are highly significant and existing problems. However, the result from teachers' questionnaire informs us that their beliefs and conceptions towards CLT are on the right track, though they face practical implementation problems in the classroom. This also has been found to be consistent with other research findings (Chang, 2011; Li, 1998; Tsai, 2007) in other EFL countries. Specifically, CLT implementing difficulties and the findings in each category are presented here as follows.

Concerning the first category, which is related to grade-9 teachers implementing difficulties, it is found that they face significant challenges and difficulties in their practical classroom implementation of CLT. There is a

discrepancy between teachers' beliefs and conceptions and their classroom practical implementation. This is because they lack the skills and expertise of implementing CLT. Therefore, teacher related problems are real problems that hinder the successful implementation of CLT in the EFL class. This result is similar to the findings of Savignon (1991), Zekariya (2010), and Chang (2011).

The second category of CLT implementation difficulty in this study is related to students. The practice of CLT involves not only teachers, but also students. Students' resistance and low-English proficiency weaken the teachers' effort to use CLT. The results echo those of previous studies (Li, 1998; Liao, 2003; Tsai, 2007).

Thirdly, the educational system is the other source of difficulty in the implementation of CLT. For example, there is no obvious, efficient and effective assessment instrument to assess students' communicative competence in EFL, as has been stated by (Li, 1998). The other difficulty is that since CLT is developed as an ESL methodology, adjustments should be made before adopted in an EFL setting (Halliday, 1994; Hiep, 2007; Pennycook, 1989). This also has to be taken into account in Ethiopia, which is an EFL setting too. The educational system needs to adjust the educational values if CLT is to be implemented properly. For example, rather than focusing on the form-based approach, which aims at developing test-taking skills, students' communicative competence should be encouraged, emphasized and developed (Ozsevik, 2010).

Finally, regarding assessment difficulties, traditional grammar-based examination was reported as a major barrier to implementing CLT in this study and it was also identified by Li (1998), Mustafa (2001) and Gorsuch (2000). Since there has been an uncertainty about how to evaluate students' communicative competence, teachers strive to administer grammar /form-based exams. In line with this, Karim (2004) stated that, the Bangladeshi EFL teachers' opinions about the traditional grammar-based exams as a difficulty and their priority of being trained to assess students are logical as they may have the expertise to assess grammar from years of experience, but they may not know how to assess communicative abilities of students. Hence, form-based exams need to be modified in order to evaluate students' communicative competence. Unlike the form-based written tests, the assessment of CLT involves assessing students' communicative competence (Bachman,1990; Hymes, 1971; Li,1997). Nevertheless, there is no ready-made tool for speech assessment easily available; as a result, teachers face challenges to evaluate the

students' communicative competence. Based on the principles of CLT, exams should assess not only grammar and vocabulary, but also listening, speaking, conversation, and writing skills as well.

In conclusion, while most teachers 'beliefs and conceptions are on the right track about CLT, their classroom practice, which is aimed at developing students' communicative competence in the target language (English), is entangled with difficulties. In addition, CLT implementing difficulties in the four categories have been found to contain significant problems of the study that hinder its successful implementation. This was also attributed to practical implementing difficulties of teachers, grammar-based examinations, educational system, and the nature of CLT itself.

Pedagogical Implications of the Study

The results of the study could have different pedagogical implications to teachers, educational professionals, and researchers. Here are two major pedagogical implications of this research.

Since a number of factors have made it difficult for CLT to be integrated into English teaching classrooms in Ethiopia, the current examination system should be reformed. As there is too much reliance on grammar-based examinations in Ethiopia, most teaching practices also focus on the skills tested in exams, which are mainly grammar, vocabulary, and reading. To improve students' communicative competence in the target language (English), attention should be shifted towards testing other language skills such as listening, speaking, and writing. To bring about this desirable change, amendments should be made especially in national examinations, which determine students' academic success and have the power to change their attention towards skills. In doing so, it is possible to shift the attention of both teachers and students towards skill-based classroom activities, skill-based teaching and learning.

The other implication of the study is that, as many teachers have difficulties with the practical implementation and assessment of CLT, they strive to use grammar-based method. Hence, in Ethiopia, special attention should be given for teacher training and retraining. Markee (2001) puts forward that teachers can change their values more easily and help bring about better changes, if they understand why there is a need to change. Therefore, the most effective way to implement CLT is to provide teachers with opportunities to retrain themselves in CLT. Hence, teachers need to be

given adequate training on how to implement CLT in the classroom and assess their students in line to its principles. This in turn reduces teachers' practical classroom implementation difficulties of CLT.

Recommendations

The ministry of education, education policy makers, curriculum designers, examination experts, post-secondary institutions and secondary school teachers in Ethiopia need to consider these suggestions /issues for the successful implementation of CLT in our context.

1. As the findings signify, teachers generally have good conceptions and positive beliefs towards CLT. Nonetheless, they are not in a position to practice these conceptions and beliefs in the actual EFL classroom effectively. This in turn results in students' 'poor communication skills' in the target language, which is quite contrary to the goal of CLT. Therefore, retraining teachers on the practical aspects of CLT has to be designed and they should be committed to implement CLT properly as well.

2. The National Organization for Examination (NOE) should change highly centralized grammar-based examinations to exams that measure students' communicative competence.

3. The contents and objectives of assessments should be in line with the contents and objectives of the curriculum. Thereby, exams would have a positive backwash effect on the curriculum, which in turn will strengthen implementation of CLT in classroom.

4. Assessment standards should be designed for CLT, which strengthens its implementation and develops students' communicative capacity.

5. Since CLT is basically an ESL methodology, not EFL, it should be amended and/or adapted to Ethiopian existing situation before use.

6. Further research needs to be conducted on the issue so as to bring desirable change on students' communicative competence in the target language (English) and confirm the consistency of findings.

References

Ajzen, I. (1988). *Attitudes, personality, and behaviour.* Chicago: Dorsey Press.

Bachman, L. F. (1990). *Fundamental Considerations in Language Testing.*

    Oxford: Oxford University Press.

Bataineh, R., Bataineh, R., & Thabet, R. (2011*). C*ommunicative language teaching in the Yemeni EFL classroom: Embraced or merely lip-serviced? *Journal of Language Teaching and Research,* 24, 859-866.

Bogale, B. (2007). Language determination in Ethiopia: What medium of instruction. Proceedings Paper: 16th International Conference of Ethiopian Studies, Trondheim

Breen, M. P., & Candlin, C. (1980). The essentials of a communicative curriculum in language teaching. *Applied Linguistics, 1*(2)*,* 89-112.

Breen, M. P., & Candlin, C. N. (2001). The essentials of a communicative curriculum in language teaching. In D. R. Hall & A. Hewings (Eds.)*, Innovation in English language teaching: A reader (pp. 9-26).* London: Routledge.

Burnaby, B., & Sun, Y. (1989). Chinese teachers' views of western language teaching: Context informs paradigms. *TESOL Quarterly, 23*(2), 219-237.

Chang, M. (2011). EFL teachers' Attitudes towards Communicative Language Teaching in Taiwanese College. *Asian EFL Journal Professional Teaching Articles*, 53, 17-34.

Deckert, G. (2004). The communicative approach: addressing frequent failure. *English teaching Forum, 42*(1), 12-17.

Ellis, G. (1996). How culturally appropriate is the communicative approach? *ELT Journal,50*(3), 213-218.

Fishbein, M., & Ajzen, I. (1975). *Belief, attitude, intention and behavior: An introduction to theory and research.* Reading, MA: Addison-Wesley Publishing Company.

Geoff Petty, G. (2004). *A Practical Guide Teaching Today*. Oxford: Oxford University Press.

Gorsuch, G. (2000). EFL educational policies and educational cultures: Influences on teachers' approval of communicative activities. *TESOL Quarterly, 34*(4), 675-710.

Gray, C., & Klapper, J. (2009). Key aspects of teaching and learning in languages. In H. Fry, S. Ketteridge & S. Marshal (Eds.), *A hand book for teaching learning in higher education: Enhancing academic practice (3rd ed.)*. New York: Rutledge.

Halliday, A. (1994). *Appropriate Methodology and Social Context*. Cambridge: Cambridge University Press.

Harmer, J. (2001). *The Practice of English Language Teaching.* England: Pearson Education Limited.

Hiep, P. (2007). Communicative Language Teaching: Unity within Diversity. *English Language Teaching Journal*, 61(3), 193-201.

Hughes, A. (1989). *Testing for Language Teachers.* Cambridge: Cambridge University Press.

Hymes, D. (1971). *Competence and Performance in Linguistic Theory*. London: academic Press.

Karim, K. M. R. (2004). *Teachers' perceptions, attitudes, and expectations about communicative language teaching (CLT) in post-secondary education in Bangladesh*. Unpublished master thesis, University of Victoria, Canada.

Lakachew, M. (2003). *Teachers' attitudes towards communicative language teaching and practical problems in its implementation.* Unpublished MA thesis, Addis Ababa University, Ethiopia.

Larsen-Freeman, D. (2000). *Techniques and principles in language teaching*. Oxford: Oxford University Press.

Li, D. (1998). "It's always more difficult than you plan and imagine": Teachers' perceived difficulties in introducing the communicative approach in South Korea. *TESOL Quarterly*, 32 (4), 677-703.

Liao, X. (2003). *Chinese secondary school teacher's attitude toward communicative language teaching and their classroom practice.* Unpublished doctoral dissertation, the University of Auckland, Australia.

Littlewood, W. T. (1981). *Communicative Language Teaching. An introduction*. Cambridge: Cambridge University Press.

Liu, G. (2005). The trend and challenge for teaching EFL at Taiwanese universities. *Regional Language Centre Journal*, 36(2), 211-221.

Markakis, J. (1974). *Ethiopia: Anatomy of Traditional Polity.* Oxford: Oxford University Press.

Markee, N. (2001). The diffusion of innovation in language teaching. In D. R. Hall & A. Hewings (Eds.), *Innovation in English language teaching: A reader* (pp. 118-126). London: Routledge.

Mengistu, L*. (1996). Autobiography [Amharic text].* Addis Ababa: Mega
Publishing.

Ministry of Education. (2005). *Education sector development program III
(ESDP-III) 2005/2006 – 2010/2011*. Unpublished document, Addis
Ababa, Ethiopia.

Mustafa, B. (2001). Communicative language teaching in Indonesia: Issues of
theoretical assumptions and challenges in the classroom practice.
Journal of Southeast Asian Education, 2(2), 1-10.

Nunan, D. (2003). The impact of English as a global language on educational
policies and practices in the Asian-Pacific region. *TESOL Quarterly*,
37(4), 589-613.

Ozsevik, Z. (2010). *The use of Communicative Language Teaching (CLT):
Turkish EFL Teachers' Perceived Difficulties in Implementing CLT*.
Unpublished Master's thesis. University of Illinois at Urbana-
Champaign, Turkey.

Pennycook, A. (1989). The Concept of Method, Interested Knowledge, and the
Politics of Language Teaching. *TESOL Quarterly*, 23(4), 589-618.

Phillipson, R. (1992). *Linguistic Imperialism.* Oxford: Oxford University Press.

Richards, J. C. (2006). *Communicative language teaching today.* New York:
Cambridge University Press.

Savignon, S. J. (1991). Communicative Language Teaching: State of the art.
TESOL Quarterly, 25 (2), 261-277.

Savingnon, S. J. (2007). Beyond communicative language teaching: What's
ahead? *Journal of Pragmatics*, 39, 207 – 220.

Savignon, S. J., & Wang, C. (2003). Communicative language teaching in
EFL context: learner attitudes and perceptions. *IRAL,*41*,* 223-249

Schulz, R. A. (1996). Focus on form in foreign language classroom: Students'
and teachers' views on error correction and of grammar. *Foreign
Language Annuals*, 29(3), 343-346.

Stoddart, J. (1986). *The Use and Study of English in Ethiopian Schools.*
Unpublished report. Addis Ababa, Ethiopia.

Tsai, T. H. (2007). *Taiwanese educators' perspective on the implementation of the
new English education policy.* Unpublished doctoral dissertation, Alliant
International University.

Yu, L. (2001). Communicative Language Teaching in China: Progress and
Resistance. *TESOL Quarterly*, 35(1), 194-198.

Weigle, S. C., & Jensen, L. (1997). *Changing Language Teaching Through
Language Testing: A Washback Study*. Cambridge: Cambridge
University
Press.

Appendixes

Survey Questionnaire Filled by Grade-Nine English Teachers & Students

Bahir Dar University

Faculty of Humanities

Department of English Language and Literature

Post-Graduate Program

Dear Respondents,

The basic objective of this questionnaire is to gather information concerning the perceived- difficulties that grade-nine EFL teachers and learners face in the implementation and assessment of CLT in EFL setting. Genuine response to this questionnaire will enable the researcher to gather reliable data which would be useful to pinpoint the major difficulties in implementing and assessing CLT in EFL context. This again serves as a benchmark to forward sound and reliable solution to the issues and which would be useful to the field of EFL too. So, you are kindly requested to give your genuine and accurate information as it greatly determines the research finding.

This questionnaire is composed of three parts. Part I asks for personal information; Part II asks questions pertaining to English language teaching methods; part III asks for grade- nine English teachers perceptions, beliefs towards CLT and the perceived difficulties in implementing and assessing CLT in an EFL setting. Finally, I would like to assure you that your responses will be kept confidential and will be used only for the research purpose.

Thank you in advance for your cooperation!

Part I: personal information

1. Gender   Female ____   Male ____

2. School: Debre Markos _____   Menkorer _____   Gozamin _____

Part II: Questions pertaining to language teaching methodology, particularly communicative Language Teaching (CLT) and Learning (1-12)

Read each item carefully and put (x) mark parallel to each item that express your understanding about CLT and its implementation in EFL setting.

Note: The meaning of the degree of agreement is indicated as follows:

4=Agree   3=Slightly Agree   2=Uncertain   1=Disagree

No.	Item	Responses			
		4	3	2	1
1	CLT is student/learner-centered approach.				
2	CLT emphasizes fluency over accuracy.				
3	CLT emphasizes communication in a second language (L2).				
4	CLT relies heavily on speaking and listening skills.				
5	CLT requires teachers to have high proficiency in English.				
6	CLT involves only group work or pair work.				
7	CLT requires higher knowledge of the target language culture.				
8	CLT involves no grammar teaching.				
9	CLT involves teaching speaking only.				
10	CLT is basically an ESL methodology, not EFL.				
11	CLT requires special training programs.				
12	CLT requires verities of teaching approach.				

Part III: Questions pertaining to perceived difficulties in adopting CLT

The following are some difficulties that other EFL teachers encountered in adopting CLT. Did you come across these difficulties or do you think they might be difficulties for you in implementing CLT in your school?

Please indicate how big an issue of these challenges is by putting ("√") mark under the following response scale. The meaning of the scale value is:-

4=Major Challenge   2=Mild challenge
3=Challenge   1=Not a Challenge at all

I.   Teacher-related difficulties

No.	Item	4	3	2	1
1	Teachers' proficiency in spoken English is not sufficient.				
2	Teachers lack the knowledge about the appropriate use of language in context.				
3	Teachers lack the knowledge about the target language (English) culture.				
4	There are few opportunities for teachers to get CLT training.				
5	Teachers have little time to develop materials for communicative classes.				
6	Teachers have student misconceptions about CLT.				

II.   Student- related difficulties and challenges

No.	Item	4	3	2	1
1	Students have low-level of English proficiency.				
2	Students have passive style of learning.				
3	Students resist participating in communicative class activities.				
4	Students lack motivation for developing communicative competence.				
5	Students have poor background knowledge about CLT.				
6	Students do not have opportunity to use the target language (English) outside classroom.				

III.   Difficulties and challenges related to educational system

No.	Item	4	3	2	1
1	There is lack of enough support from administration.				
2	Teachers lack authentic materials such as newspapers, magazines, movie etc.				
3	Traditional view on teachers' and learners' role is not compatible with CLT.				
4	Classes are too large for the effective use of CLT.				
5	The objectives of the syllabus and assessment don't hand in hand.				
6	Grammar-based examinations have negative impact on the use of CLT.				

IV. Difficulties related to Assessment

No.	Item	4	3	2	1
1	Grammar-based examinations discourage the use of CLT.				
2	It is reading, vocabulary and grammar that are most frequently included in both teacher made and standardized tests.				
3	Many teachers lack the skill and experience of assessing the above skills.				
4	Facilities for assessing listening, speaking and writing skills (for instance, tape recorder, video cassette etc.) are absent in the school.				
5	The intentions of the syllabus don't coincide with the objectives of the assessment.				
6	Formative assessment is not emphasized as of the summative one in terms of CLT implementation.				

Please list any other potential difficulties you might encounter in adopting CLT in your school.

_____

_____

_____

_____