The Asian EFL Journal

Volume 21, Issue 1

ASIAN EFL JOURNAL
The EFL Professional's Written Forum

Published by the English Language Education Publishing

Asian EFL Journal
A Division of TESOL Asia Group
Part of SITE Ltd Australia

http://www.asian-efl-journal.com

editor@asian-efl-journal.com
Publisher: Dr. Paul Robertson
Chief Editor: Dr. John Adamson
Production Editor: Dr. I-Chin Nonie Chiang

ISSN 1738-1460

ASIAN EFL JOURNAL
The EFL Professional's Written Forum

Table of Contents

Book Reviews

Forward

Welcome to the 1st edition of the Asian EFL Journal in 2019. This edition contains six research articles and two book reviews. In the first article, *Contribution of L2 Morphological Awareness to Passage-Level L2 Reading Comprehension*, Aydin and Yildirim investigate the impact of morphological awareness in adult Turkish EFL learners in passage-level reading comprehension through statistical analysis and conclude that the learners show moderate awareness of English word morphology, which might contribute to reading comprehension indirectly. Next, Lee and Gray, in *'Please, somebody talk': Incorporating Dialogical Critical Thinking into EFL Classes for Young Beginners* explore ways to incorporate critical thinking through discussion with a small class of young beginners. Through the extensive analysis of data, the authors offer five practicing principles for teachers seeking to include dialogical CT in young beginners' EFL classes. Then, in *Constructing, Reconstructing and Developing Teacher Identity in Supportive Contexts*, Kaya and Dikilitaş detail a language teacher's identity development through six-month triangulated case study data based on interview, observation, and reflective notes, and find the perspective change pattern of a novice teacher and the important role a supportive context plays in the teacher development. Then, to understand the factors which influence Korean jobseekers' perceptions of the TOEIC regarding employment, how these factors (de)motivate test preparation, and the statistical relations among the factors, Kim, Choi and Kim, in their *South Korean Jobseekers' Perceptions and (De)motivation to Study for Standardized English Tests in Neoliberal Corporate Labour Markets*, analyse data with exploratory factor analysis, correlation and regression. The results indicate that the participants consider the tests more as a job skill to help their work performance after they get a job than a proficiency measurement. Also, the participants' motivation does not lead to positive performance, while the neoliberal pressure negatively influences performance. In the following paper, *Young Children's L2 Vocabulary Learning Through Cooking: The Case Of Korean EFL Children*, Park, Choi, Kiaer and Seedhouse investigate and discuss the pedagogical benefits of learning vocabulary through real cooking tasks. These real-world experiences create not only significant better performance but also an enjoyable and sustainable learning environment for young learners. Then, Thirakunkovit and Chamcharatsri, after carrying out an extensive review and careful screening in the fields of composition studies and second language writing, find that teacher feedback produces larger effect size than peer feedback. Additionally, the authors argue that specific training on providing peer feedback must be provided in order to enhance the quality and effectiveness. The penultimate paper is the first book review. Irgin's strong recommendation on *Improving Foreign Language Teaching: Towards a Research-*

Based Curriculum and Pedagogy (2015) makes the book a must read for all related parties in the foreign language teaching field. Finally, Zohrehfard's insightful review of *Language, Identity and Education on the Arabian Peninsula-Bilingual Policies in a Multilingual Context* (2016) helps readers see the strengths and further expansions of relevant issues.

I-Chin Nonie Chiang
Production Editor of AEJ January 2019
National Chengchi University

Contribution of L2 Morphological Awareness to Passage-Level L2 Reading Comprehension

Fatma Aydin
Anadolu University, Turkey
Ozgur Yildirim
Anadolu University, Turkey

Bioprofile:

Fatma Aydin is an EFL lecturer at School of Foreign Languages, Anadolu University in Eskisehir, Turkey. She holds an MA degree on Teaching English as a Foreign Language. Her research interests include reading in foreign language, foreign language reading strategies, and morphological awareness. **Email:** f.aksoy@anadolu.edu.tr **(corresponding author)**

Ozgur Yildirim is an EFL teacher trainer at Faculty of Education, Anadolu University in Eskisehir, Turkey. He holds a PhD degree on Second and Foreign Language Education. His research interests include learner autonomy in language learning, EFL teacher training, strategy use in foreign language learning, international education, and morphological awareness. **Email:** oyildirim@anadolu.edu.tr

Abstract

The purpose of this study was to investigate to what extent intermediate-level adult Turkish EFL learners are aware of lexical and grammatical functions of certain suffixes in English. The study further aimed at exploring the unique contribution of L2 morphological awareness to passage-level L2 reading comprehension above and beyond such reading-related factors as vocabulary knowledge, grammar knowledge and reading strategy use. The data were collected from 87 EFL learners through a Morphological Awareness Test, a Survey of Reading Strategies, and tests of reading comprehension, vocabulary, and grammar. Data analysis included descriptive statistics, paired samples t-test, hierarchical multiple regression, and bivariate regression. The results of the study revealed that the participants show at least moderate awareness of English word

School of Foreign Languages, Anadolu University, Eskisehir, Turkey 26470

morphology, and they could judge whether a morphologically complex word comes from a simple word better than they identified the lexical and grammatical functions of certain derivational suffixes in English. The results further revealed that L2 morphological awareness does not make a direct contribution to L2 reading comprehension but it might make an indirect contribution to it via its significant contribution to L2 vocabulary knowledge.

Key w ords: Intra-word structure, word morphology, morphological awareness, L2 reading comprehension

Introduction

Background to the Study

To be able to read and comprehend texts in a foreign language is an integral part of knowing that language, particularly for those who are learning a foreign language for academic purposes, as they have to deal with written language most of the time. Accordingly, Carrell (1988) and Grabe and Stoller (2001) state that reading should receive special focus in foreign language teaching because many foreign language learners have reading as one of their most important goals. When it comes to the problems in reading, they may stem from a number of reasons such as word-recognition difficulties, weaknesses in grammar knowledge, not being able to apply reading strategies or even being unaware of them, lack of background knowledge, lack of knowledge of text structure and discourse cues (Koda, 2005). Moghadam, Zainal and Ghaderpour (2012) maintain that vocabulary learning is prevalent in language acquisition and a lot of unknown words that learners encounter while reading might cause difficulties in processing the text, and that reading comprehension breakdowns stem from problems/limitations concerning word recognition and lexical access. Similarly, Anjomshoa and Zamanian (2014) note that without the ability to recognize the meanings of words, both perceiving and producing the language would be impossible, and Mehrpour, Razmjoo and Kian (2011) indicate that even if students decode successfully and read fluently, knowledge of word meanings is critical to reading comprehension.

Considering the crucial role of vocabulary knowledge in reading comprehension, we have the following problem: there are always unknown words. Abundance of unknown words in a reading text may intimidate EFL learners. Therefore, they get bored, distracted and ultimately give up reading. For this very reason, EFL learners should be provided with the awareness that they are able to and had better survive reading a text in English without knowing the exact meaning of every single unknown word. Considering the unknown vocabulary and the challenges it poses to L2 reading comprehension, a possible solution may be lexical inferencing, which is defined as “making informed guesses as to the meaning of a word in light of all available linguistic cues in combination with the learners’ general knowledge of the world, her awareness of context and her relevant linguistic knowledge” (Haastrup, 1991, p.13). Intra-word knowledge or morphological awareness is a part of lexical inferencing process and is likely to be a practical tool for EFL learners to gain insights into the meaning of multi-morphemic unknown words regardless of contextual information. However, the number of studies investigating the role of morphological awareness in teaching and learning English as a foreign language is quite limited in the general EFL context (Jeon, 2011).

A considerable number of studies (ranging from 1970s to present) have yielded empirical findings that morphological awareness boosts L1 reading comprehension (Brittain, 1970; Carlisle, 1995; Carlisle, 2000; Katz, 2004; Kirby et al., 2012). Within the light of these findings, foreign/second language teaching practitioners and researchers attempted to find out whether morphology contributes to L2 reading as it does to L1 reading, and how much this contribution is when the effects of other reading and language-related variables are controlled. However, a review of literature lays bare that, despite the abundance of studies conducted to investigate the contribution of morphological awareness to reading comprehension in L1, the number of studies carried out to explore the contribution of morphological awareness to L2 reading comprehension is quite limited (Jeon, 2011). Besides, these limited number of studies tackle the issue from the point of view of language minority children who learn the target language as a second language. The number of studies conducted in the foreign language learning environments is even more limited.

The present study aims to explore to what extent EFL learners are aware of syntactic and lexical functions of certain suffixes in English, and to investigate whether morphological awareness in English contributes to EFL reading comprehension. With these purposes, the study addresses the following research questions: (1) To what extent are intermediate-level adult Turkish EFL learners aware of syntactic and lexical functions of certain suffixes in English?; (2) Does morphological awareness contribute to passage-level L2 reading comprehension of intermediate-level adult Turkish EFL learners above and beyond other reading and language-related factors such as grammar knowledge, vocabulary knowledge, and use of reading strategies? As the second research question lays bare, this study aims to control the effects of grammar knowledge, vocabulary knowledge, and use of reading strategies, which are the variables that are believed to have a strong and positive relationship with reading (Grabe, 2009; Koda, 2005), so that the unique contribution of morphological awareness to passage-level EFL reading comprehension can be explored.

Morphological Awareness as a Predictor of L2 Reading

Jeon and Yamashita (2014) investigate ten key reading component variables in their meta-analysis on L2 reading comprehension and its correlates. Results suggest that there are four high-evidence correlates: L2 decoding, L2 vocabulary knowledge, L2 grammar knowledge, L1 reading comprehension; and six low-evidence correlates: L2 phonological awareness, L2 orthographic awareness, L2 morphological awareness, L2 listening comprehension, working memory, metacognition. Additionally, a good number of both L1 and L2 reading researchers accept that effective reading strategy use supports reading comprehension (Akkakoson, 2013; Block & Pressley, 2007; Guthrie & Taboada, 2004; Hudson, 2007; Peng, Siriyothin & Lian, 2014; Qanwal, & Karim, 2014).

Koda (2005) maintains that word knowledge is formulated in specific contexts, from which extracting lexical information is an acquired competence that contributes to reading comprehension. However, it is crucial to understand the internal structure of the words to be able to extract lexical information. Koda (2005) further maintains, “in-depth awareness of such structural awareness can point up the essential competencies underlying word-knowledge accretion through reading” (p.71). In this sense, metalinguistic awareness, which refers to “the ability to identify, analyze, and manipulate language forms” (Koda, 2005, p.72), plays an important role in reading development. It is worth mentioning that metalinguistic awareness differs from linguistic awareness in the sense that it requires a general understanding of language, regardless of its specific details (Koda, 2005). There has been a boost in research on metalinguistic awareness recently. Morphological knowledge, a type of intra-word knowledge, and morphological awareness, a type of metalinguistic awareness, has gained importance accordingly. The term ‘morphological awareness’ refers to “the ability to reflect on and manipulate morphemes and word formation rules in a language” (Kuo & Anderson, 2006, p.161). In other words, it is the ability to recognize the internal structure of a word, identify the root and the affixes as well as being able to form new words applying the word formation rules in a particular language.

Several studies investigated the effect of morphological awareness on L2 reading comprehension in ESL environments. In Wang, Cheng, and Chen (2006) Chinese immigrant children living in the United States were assessed on phonological awareness, morphological awareness, vocabulary knowledge, word reading and reading comprehension in English. The study found that both derivational awareness and compound structure made a significant contribution to reading comprehension. Furthermore, morphology was found to have a strong positive correlation with vocabulary knowledge and reading comprehension.

Kieffer and Lesaux (2008), a longitudinal study, assessed Spanish-speaking English language learners’ ability to decompose derived words during reading with a specifically-designed task, and analyzed whether their accomplishment on this task makes a contribution to reading comprehension above and beyond such reading-related skills as word reading, phonological awareness and breadth of vocabulary knowledge. Findings revealed a significant relationship between derivational morphological awareness and reading comprehension.

Another study by Kieffer and Lesaux (2012), which tested the direct and indirect contributions of derivational morphological awareness to English reading comprehension in sixth graders of Spanish, Filipino and Vietnamese backgrounds, reported that morphological awareness made a significant, direct contribution to reading comprehension when the effects of reading vocabulary and word reading fluency were controlled. They also found that morphological awareness made a significant, indirect contribution to reading comprehension through reading vocabulary.

Kieffer, Biancarosa and Mancilla-Martinez (2013) also studied the direct and indirect roles of morphological awareness in L2 reading comprehension in 101 Spanish-speaking language minority learners. Results indicated that morphological awareness uniquely and significantly contributed to reading comprehension, controlling for phonemic decoding, listening comprehension, reading vocabulary, word reading fluency and passage reading fluency.

All the aforementioned studies were conducted with children learning English as a second language. Considering the differences between second and foreign language learning contexts, it is important to review the studies that explored the contribution of morphological awareness to foreign language reading comprehension in EFL context. Jeon (2011) investigated the contribution of L2 morphological awareness to foreign language reading comprehension assessing 188 South Korean tenth grader learners of English on six reading and language-related variables: phonological decoding, listening comprehension, vocabulary knowledge, passage level reading comprehension, metacognitive reading awareness, and morphological awareness. The results of ten sequential regression analyses showed that morphological awareness significantly contributed to L2 reading comprehension when the other variables were controlled.

Zhang and Koda (2012) tested the direct and indirect effects of morphological awareness and lexical inferencing ability on L2 vocabulary knowledge and reading comprehension among advanced Chinese EFL learners. Results revealed that morphological awareness contributed to L2 vocabulary knowledge both directly and indirectly. The indirect contribution was due to the mediation effect of learners' lexical inferencing ability. Inferencing ability was reported to refer to the integration of structural information (derivational morphology) and semantic information (morpheme meaning). However, it was found that morphological awareness failed to make a significant, unique or direct contribution to reading comprehension after adjusting for vocabulary knowledge. Instead, it was found to be a significant, but indirect contributor to reading comprehension, controlling for vocabulary knowledge.

Almost all of the previous studies report to have found that L2 reading comprehension benefits from morphological awareness either directly or indirectly, even though there is no standard among the methodologies they employed. This makes the relationship between morphological awareness and L2 reading comprehension significant for researchers to understand EFL learners' L2 learning. However, considering that the number of studies carried out to explore the contribution of morphological awareness to L2 reading comprehension is quite limited, the questions whether morphology contributes to L2 reading, and how much this contribution is when the effects of other reading-related variables are controlled remain only partially answered (Jeon, 2011). Besides, the data of relevance to foreign language learning (EFL environment), which is the case in Turkey, is even more limited. Therefore**,** the present study attempts to reduce a gap in the field by finding out whether adult Turkish EFL learners are aware of morphological structure of English and whether or to what extent this awareness is in line with passage-level L2 reading comprehension. As well as reducing a gap in the field, the present study is significant in the sense that it investigates the role of morphological awareness, which might be a practical tool for EFL learners to struggle with the challenge of unknown vocabulary during reading.

Methodology

Participants

With the purpose of assessing adult Turkish EFL learners' morphological awareness level and determine its relationship to passage-level L2 reading comprehension, 87 Turkish first year university students who were taking intermediate level English courses were selected as the participants of the study. The participants' proficiency level had been determined as a result of a proficiency exam administered at the beginning of the semester they were taking the English courses. The university where the participants of the study were studying adopts an integrated approach to foreign language learning; that is, English is taught in a way in which language skills and sub-skills are interwoven, and students are encouraged to learn these skills simultaneously.

The participants were between 18 and 22 years old. They were all monolinguals with their native language being Turkish, and no participant included in the current study had an extended stay in an English speaking country. They had not taken any standardized tests before, like IELTS or TOEFL. Their areas of studies varied (engineering, business administration and economics).

Materials

Considering that the main objective of the present study is to explore the contribution of morphological awareness to passage-level L2 reading comprehension, passage-level L2 reading comprehension is the dependent variable, while morphological awareness is the independent variable of the study. Additionally, L2 vocabulary and grammar knowledge, and L2 reading strategy use are the control variables in the study. In this regard, the participants of the study were given a Morphological Awareness Test, tests of reading comprehension, vocabulary and grammar, and a Survey of Reading Strategies.

Morphological Awareness Test

Morphological Awareness Test (MAT) (Appendix 1) consists of two sections: Sentence Completion and Word Relation. The Sentence Completion section consists of 27 multiple-choice test items. The purpose of this section is to understand whether the participants are aware of the grammatical and lexical functions of some common derivational suffixes in English. The target words used in this section are pseudo-words that are appropriate for the morphological structure of English. The derivational suffixes used in this test are -tion, -ist, -(i)ty (noun-making suffixes); -ate, -ize, -fy (verb-making suffixes); and -ous, -ive, -al (adjective-making suffixes). These suffixes have been selected for the test because they are among the very common suffixes in English (Fry & Kress, 2006; White, Sowel & Yanagihara, 1999); -tion, -(i)ty, -ous, -ive and -al are among the most common 20 suffixes that account for 93 percent of occurrences according to a research-based list by White, Sowel, and Yanagihara (1999); and -ist, -ate,

-ize and –fy are among other common suffixes that account for seven percent of the occurrences.

The test items in the Sentence Completion section have been taken from Mahony (1993) with minor adjustments. The participants were instructed to choose the best word out of four options formed with the same root (a pseudo-word) but different suffixes in order to complete a sentence with a blank. As can be seen in the following sample test item (Figure 1), once a learner knows what functions the derivational suffixes -ize, -ive, -al and -tion serve, as listed in the options respectively, they can easily complete the sentence with the correct form of the word, indicating that they have some insights into English derivational morphology.

His ___________________________ is greatly admired.

A. superfilize
B. superfilive
C. superfilial
D. superfilation (correct answer)

Figure 1 Sample Test Item from the Sentence Completion Section of MAT

In the Word Relation section, the participants were provided with a list of 20 word pairs (real words), and asked whether the second word in each pair comes from the first word or not. As the following sample test items displays (Figure 2), after reading each word pair the participants circle YES, if they think the second word comes from the first word, or NO if they think the second word does not come from the first word. Three teachers of the participants reviewed the words and confirmed that the participants were familiar with the meaning of all of the words that were used in this section of the test. Nine of the word pairs were morphologically related, while the remaining 11 pairs were morphologically unrelated. Some of the word pairs in this section were taken from Mahony (1993), while some others were designed accordingly by the researcher.

creative → creativity	*YES*	-	*NO*	(Related)
part → party	*YES*	-	*NO*	(Unrelated)

Figure 2 Sample Test Item from the Word Relation Section of MAT

To ensure the validity of MAT for this research purpose and context, expert opinion of four professors from the English Language Teaching Department of the same university were taken. For internal reliability, Cronbach's alpha coefficient of MAT was found to be .78 for this study.

Tests of Reading Comprehension, Vocabulary, and Grammar

The reading comprehension test used in this study consisted of 25 multiple-choice test items. It included authentic reading texts with an average length of 200-500 words. The texts were chosen from magazines, newspapers, or similar sources considering the proficiency level of the participants. These sources were preferred because they apply to a wider readership providing real-life texts that are likely to address to most peoples' interests. Question types were understanding the main idea and details, making inferences, guessing meaning from context and finding in-text references. The vocabulary test also consisted of 25 multiple-choice test items. Question types in the test included finding the meaning of a word, assessing intra-word structure, finding synonyms and/or antonyms of certain words, and assessing knowledge of word phrases. The grammar test consisted of 25 multiple-choice test items as well. Some of the test items required participants to choose the best answer with the appropriate structures to fill in the gaps embedded in sentences or paragraphs. In some other test items the participants were instructed to choose the ungrammatical word or phrase from among the underlined parts of a sentence. All the questions that were used in the tests of reading comprehension, vocabulary, and grammar were prepared by a group of experienced English instructors who were teaching English at the university the participants of the study were attending.

Survey of Reading Strategies

Reading strategy use was assessed by using Survey of Reading Strategies (SORS), which was adapted by Mokhtari and Sheorey (2002) from Mokhtari and Reichard (2002). SORS consists of 30 items and the participants answer the survey by using a Likert-type scale ranging from 1 (never) to 5 (always). Higher average scores from SORS indicate more frequent reading strategy use. In the present study, Cronbach's alpha coefficient of SORS was found to be .80.

Data Collection and Analysis

The participants were first given Morphological Awareness Test (MAT) and Survey of Reading Strategies (SORS). During the administration of MAT, the participants were not told that the words used in the Sentence Completion section were pseudo-words, they were told that the words were new to them, and they were reminded to use their knowledge of word endings (suffixes) to be able accomplish the test. Approximately ten days after taking MAT and SORS, the participants took the tests of reading comprehension, vocabulary, and grammar.

For the analysis of the data, first, descriptive statistics (mean scores, standard deviations, quartiles, percentages) were calculated for all the instruments of the study. Then, the first research question of the study was answered by using the descriptive statistics, and by comparing the mean scores of the two sections of MAT using a paired-samples t-test. As for the second research question, with the purpose of finding the unique contribution of morphological awareness to reading comprehension above and beyond such reading-related factors as vocabulary knowledge, grammar knowledge and reading strategy use, a series of multiple regression and bivariate regression analyses were conducted.

Results and Discussion

Morphological Awareness among Intermediate-level Adult Turkish EFL Learners

The first research question of the present study is as follows: To what extent intermediate-level adult Turkish EFL learners are aware of syntactic and lexical functions of certain suffixes in English? Table 1 below presents the minimum and maximum scores of the participants from Morphological Awareness Test (MAT) and its subsections along with the mean scores and standard deviations. In order to better understand the results, it is important to note that each correct answer in MAT corresponds to 1 point, making the highest possible score for the overall test 47, and highest possible scores for the Sentence Completion and Word Relation subsections 27 and 20, respectively.

Table 1 Descriptive Statistics: MAT & Subsections of MAT

	N	**Min.**	**Max.**	**Mean**	**SD**
Morphological Awareness Test	87	18	45	33.09	5.86
Sentence Completion Section	87	4	26	15.38	4.99
Word Relation Section	87	10	20	17.69	1.80

As shown in Table 1, the mean score of the participants from MAT is 33.09, while the minimum and maximum scores are 18 and 45 respectively. Lower quartile value of MAT was found as 29, which indicates that 75 percent of the participants of this study got a score of 30 or more from the test. All in all, the mean score of the participants and the percentage of participants who scored more than 30 may together indicate that they have moderate to high morphological awareness. This means that they have moderate to high "ability to reflect on and manipulate morphemes and word formation rules in English" (Kuo & Anderson, 2006, p.161). In other words, they are able to recognize the internal structure of a word, identify the root and the affixes as well as being able to form new words applying the word formation rules in a particular language to a moderate to high extent. The explanation that the majority of the participants in the present study have at least moderate awareness of morphological awareness might lie in their English proficiency and the time they have spent learning English so far. They are at intermediate level, indicating that they have made considerable progress towards being proficient EFL learners. At their proficiency level they can understand and talk about the main points of familiar topics or the topics of personal interest; they can deal with situations that arise while they are travelling in an English speaking country and they can describe their experiences, dreams and hopes giving reasons.

For further analysis, first, participants' scores from each test item in the Word Relation section, which consists of 20 test items, were multiplied with 1.35 to make them equal to the participants' scores from the Sentence Completion section, which consists of 27 test items. By doing so, statistical comparison was made possible between the mean scores of the subsections of MAT. Next, a paired samples t-test was carried out to find out if the participants' performance in the Sentence Completion section of MAT significantly differed from their performance in the Word Relation section. The results revealed a statistically significant difference between the Sentence Completion section ($M = 15.38$, $SD = 4.99$) and the Word Relation section (adjusted $M = 23.93$, $SD = 2.45$) ($t(86) = -17.000$, $p < .001$). This finding indicates that the participants in the present study performed significantly better in the Word Relation task than they did in the Sentence Completion task. In other words, intermediate-level adult Turkish EFL learners participating in the present study were considerably better at judging whether a morphologically complex word comes from a simple word than identifying the lexical and grammatical functions of certain derivational suffixes in English.

The items in the Word Relation section measure a basic facet of morphological awareness. These items interrogate solely whether or not EFL learners have a superficial knowledge of intra-word structure in English. Results indicate that the majority of the participants of the present study can recognize that the word 'sunny' comes from the word 'sun', and once they notice these two words are semantically related, they can conclude that the ending –y is able to form new words when it is added to a word. In other words, if they know that 'sun' is a noun and 'sunny' is an adjective, they can deduce that the ending –y is used to make adjectives from nouns. Then, they can multiply the examples such as rain - rainy, snow - snowy, oil - oily. However, they can also recognize that the word 'man' does not come from the word 'many' although this pair seems to have the same relationship with the 'sun – sunny' pair. Similarly, the participants of the present study can judge that the noun 'winner' comes from the verb 'win' but the noun 'letter' does not come from the verb 'let'. These findings point out that intermediate-level adult Turkish EFL learners participating in the present study can identify whether or not two apparently similar words are morphologically and thus semantically related or not.

Contribution of Morphological Awareness to Reading Comprehension

The second research question of the study seeks an answer to whether or not morphological awareness contributes to passage-level L2 reading comprehension above and beyond other reading and language-related factors such as grammar knowledge, vocabulary knowledge, and use of reading strategies among intermediate-level adult Turkish EFL learners. Table 2 presents mean scores and standard deviations regarding reading strategy use and tests of reading comprehension, grammar, and vocabulary. It is important to note that the highest possible score from the tests of reading comprehension, grammar, and vocabulary was 25; and the highest possible score from the Survey of Reading Strategies was 5.

Table 2 Descriptive Statistics: Proficiency Exam

	N	Mean	SD
Reading Comprehension	87	14.71	4.046
Vocabulary	87	14.37	4.029
Grammar	87	12.61	3.826
Reading Strategy Use	87	3.40	0.46

A hierarchical multiple regression analysis was conducted in order to understand the effect of morphological awareness on reading comprehension. In this regression analysis, vocabulary, grammar, and reading strategy scores were first entered in the regression model as a block in order to control their effects. Table 3 presents the results of this regression analysis.

As Table 3 indicates, vocabulary knowledge, grammar knowledge, and reading strategy use significantly explain 44.7 percent of the variance in reading comprehension among intermediate-level adult Turkish EFL learners. On the other hand, controlling for

Table 3 Hierarchical Multiple Regression Model (Dependent Variable: Reading Comprehension)

	R	R Square	Adjusted R Square	Standard Error	F Model	R Square Change	F Change
SORS Grammar Vocabulary	.669	.447	.427	3.062	22.385*	.447	22.385*
MAT	.677	.459	.432	3.049	17.365	.011	1.722

*. F is significant at the 0.01 level.

these three variables, morphological awareness does not make a significant contribution to reading comprehension because the unique contribution of MAT was not found to be significant in the regression model. In other words, this regression model clearly shows that morphological awareness makes almost no contribution to reading comprehension above and beyond reading strategy use, grammar knowledge, and vocabulary knowledge.

The present study differs from Droop and Verhoeven (2003), Wang et al. (2006), Kieffer and Lesaux (2008), Kieffer, Biancarosa and Mancilla-Martinez (2013), Kieffer (2014), Jeon (2011), Zhang and Koda (2013) and Zhang, Koda and Sun (2014) in the sense that its findings did not yield a statistically significant contribution of morphological awareness to passage-level L2 reading comprehension. This differentiation may be due to a number of differences between aforementioned studies and the present study. First of all, most of these studies (Droop & Verhoeven, 2003; Kieffer, 2014; Kieffer, Biancarosa & Mancilla-Martinez, 2013; Kieffer & Lesaux, 2008; Wang et al., 2006) seek the contribution of morphological awareness to reading comprehension among ESL learners whereas the target population in the present study is EFL learners. ESL learners might have higher awareness of morphological structure in English and accordingly they might make use of this awareness in reading academic texts to a greater extent than EFL learners.

Second, the different findings might also be attributed to variations in learners' age. In most of these studies, except Jeon (2011), the participants are third to eighth grade learners. However, in the present study, we tried to measure the effect of morphological awareness to reading comprehension among adult EFL learners whose ages range between 18 and 22. Such a difference indicates that the unique variance in L2 reading comprehension explained by morphological awareness might decrease as the learners grow up. To verify such a claim, however, further study on younger Turkish EFL learners is required.

Third, regarding the morphosyntactic differences between Turkish and English, the participants' L1 (Turkish) might also be playing a role in the contribution of morphological awareness to L2 reading comprehension (English). However, considering that some of the native languages of the participants in the studies that found unique contribution of morphological awareness to reading comprehension (Arabic, Turkish (L2-Dutch), Chinese, Korean, Spanish, Flipino and Vietnamese (L2-English)) have some morphosyntactic features in common with the target languages, while some others do not, it is beyond the scope of the current study to determine the potential role of L1 in the contribution of L2 morphological awareness to L2 reading comprehension. In addition, it is also necessary to bear in mind that the present study is assumed to be the first study investigating the contribution of morphological awareness to L2 reading comprehension among Turkish EFL learners, which means that for the Turkish context there is not enough evidence for L1 influence on the relationship between L2 morphological awareness and reading comprehension. In other words, L1 might play a role in the contribution of L2 morphological awareness to L2 reading comprehension, which can be better investigated through a comprehensive cross-linguistic study.

The fourth reason why the present study did not come up with unique contribution of morphological awareness to reading comprehension unlike the previous ones might lie in the measurement of morphological awareness. The morphological awareness task in the current study only dealt with learners' ability to identify lexical and grammatical functions of certain derivational suffixes in English and whether two words that seem to be related are in fact morphologically related or not. Measures tapping different levels of morphological awareness could have been employed. On the other hand, it is well worth noting that because currently no test of basic L2 morphological awareness is available, researchers design their own morphology test with some adaptation from previous research such as Carlisle's (2000) Test of Morphological Structure. Therefore, there might be variation in each context in terms of the testing instruments. In addition, focusing on one only category of derivational suffixes might have produced different findings. For instance, including only verb-making derivational suffixes in the test might have changed the results. Furthermore, focusing on prefixes as well as suffixes might have changed the findings of the present study as well.

On the other hand, the findings of the present study corroborate Zhang and Koda (2012), who explored the effect of morphological awareness and lexical inferencing ability on vocabulary knowledge and the contribution of morphological awareness, lexical inferencing and vocabulary knowledge to reading comprehension among adult Chinese EFL learners. It is worth noting that the two studies are similar in their English language learning context, namely learning English as a foreign language (EFL) and age group of the participants, namely adult learners. Additionally, the current study accords with Qian (1999), who found that morphological awareness did not make a significant contribution to reading comprehension among Canadian ESL learners controlling for vocabulary size and depth.

Indirect Contribution of Morphological Awareness to Reading Comprehension

Although the results of the present study did not reveal any direct contribution of morphological awareness to reading comprehension, it might be possible to talk about an indirect contribution by examining the effect of morphological awareness on vocabulary knowledge and grammar knowledge. Table 4 shows the results of the multiple regression analysis in which reading comprehension is the dependent variable, and the independent variables of Grammar and Vocabulary are entered into the model after reading strategy use and morphological awareness.

Table 4 Hierarchical Multiple Regression Model (Dependent Variable: Reading Comprehension)

	R	**R Square**	**Adjusted R Square**	**Standard Error**	**F Model**	**R Square Change**	**F Change**
SORS	.131	.017	.006	4.035	1.487	.017	1.487
MAT	.175	.030	.007	4.031	1.320	.013	1.150
Grammar	.570	.324	.300	3.385	13.285	.294	36.113*
Vocabulary	.677	.459	.432	3.049	17.365	.134	20.324*

*. F is significant at the 0.01 level.

The results presented in Table 4 indicate that, controlling for reading strategy use and morphological awareness, grammar and vocabulary make significant contributions to reading comprehension. Grammar significantly explains 29.4 percent of the variance in reading comprehension above and beyond reading strategy use and morphological awareness; and vocabulary significantly explains 13.4 percent of the variance in reading comprehension above and beyond other three independent variables. Considering these results, examining the effect of morphological awareness on grammar and vocabulary might indicate some indirect contribution of morphological awareness on reading comprehension.

Table 5 presents the results of the bivariate regression analysis in which the grammar knowledge test score is the dependent variable and the morphological awareness test score is the independent variable; and Table 6 presents the results of the bivariate regression analysis in which the vocabulary knowledge test score is the dependent variable and the morphological awareness test score is the independent variable.

Table 5 Bivariate Regression Model (Dependent Variable: Grammar Knowledge)

	R	**R Square**	**Adjusted R Square**	**Standard Error**	**F Model**	**R Square Change**	**F Change**
MAT	.178	.032	.020	3.787	2.786	.032	2.786

*. F is significant at the 0.01 level

Table 6 Bivariate Regression Model (Dependent Variable: Vocabulary Knowledge)

	R	**R Square**	**Adjusted R Square**	**Standard Error**	**F Model**	**R Square Change**	**F Change**
MAT	.364	.132	.122	3.775	12.966*	.132	12.966*

*. F is significant at the 0.01 level.

Table 5 reveals that morphological awareness does not have a significant effect on grammar knowledge. On the other hand, Table 6 points out that the R square of the regression model was found to be .132 and significant at the 0.01 level. This finding indicates that morphological awareness significantly explains 13.2 percent of the variance in vocabulary knowledge scores of the intermediate-level adult Turkish EFL learners.

Studies investigating knowledge of affixes, lexical inferencing skill and vocabulary knowledge also state that the three are related. Paribakht and Wesche (1999), for instance, purport that learners' knowledge of L2 word derivations, namely stems and affixes, and that of grammatical inflections such as –ed and –s together came out as the second most important knowledge source in inferring the meaning of unknown words. In addition, Schmitt and Meara (1997) found weak but still significant correlations between receptive derivative suffix knowledge and vocabulary size. They note that derivative suffix knowledge fits with vocabulary size, and once learners are familiar with derivative suffixes, they can have access to the other members in a word family. Mochizuki and Aizawa (2000), in a similar vein, explored the relationship between L2 learners' vocabulary size and their affix knowledge. They found that affix knowledge increased in proportion to vocabulary size. They point out that vocabulary expansion enhances understanding of affixes, which in turn extends vocabulary size. That learners who have better morphological awareness are likely to learn words better and thus hold a larger vocabulary size is also congruent with Nagy (2007), who, with regard to English monolingual children's vocabulary acquisition, maintains that there is a relationship between intra-word structure and vocabulary acquisition, indicating the effect of morphological awareness on vocabulary knowledge. Furthermore, Sandra (1994) proposes that learners' insights into stems, affixes and word structure, help them infer the meanings of unknown words, enhance the representation of complex words in their mental lexicon and enable them to retain vocabulary items.

Considering that vocabulary knowledge plays a significant role in one's reading comprehension and is likely to predict reading achievement, we can arrive at the conclusion that morphological awareness indirectly contributes to reading comprehension by means of the significant contribution it makes to vocabulary knowledge. Nagy (2007) states that morphological awareness may contribute to reading comprehension because it helps students interpret novel morphologically complex words that they encounter while reading, which he calls "on-the-spot vocabulary learning". Nagy further notes that morphological awareness may contribute to reading comprehension through the use of syntactic signals provided by suffixes to parse complex sentences. They exemplify this with the difference between the sentences "Observant investigators proceed carefully" and "Observe investigators' procedures carefully", which is signaled completely by suffixes (p.64). Tyler and Nagy (1990), with this regard, maintain that poorer readers are less likely to notice such signals, indicating the indirect contribution of morphological awareness to reading comprehension via syntactic parsing.

Pedagogical Implications

The findings of the current study point out that word formation rules and morphological awareness are important in EFL curriculum and pedagogy. Therefore, appropriate in-class activities and multimedia packages can be developed to draw EFL learners' attention to the lexical and grammatical functions of English derivational affixes and structure of English derivational words. Such activities may include segmenting affixes from derived words and identifying base morphemes, grouping derived words according to their suffixes, and inferring meanings of unfamiliar derived words with familiar base words (Zhang & Koda, 2013). Moreover word-sorts and word-relatives activities may also be used in EFL classrooms. Direct instruction of word roots might also provide benefits for EFL learners (Apel & Werfel, 2014). It is worth mentioning that teachers should consider the characteristics of the learner group that they address, especially age and language proficiency level. Morphological awareness/production tasks can be used with both young learners and adults. As for learners' proficiency level, teachers should decide on implicit or explicit morphological awareness/production tasks in accordance with what learners already know and need to learn more.

It should also be kept in mind that the primary objective of morphological awareness instruction is offering learners an array of opportunities to benefit from word morphology so as to boost their vocabulary knowledge, enhance their spelling skill and be more proficient in reading, writing and even speaking. Therefore, morphological awareness instruction should be integrated into vocabulary, reading, writing and speaking tasks. To iterate, for instance, subsequent to discovering the meaning of a specific affix and how it is spelled, learners may be requested to read a text, find the words to which this affix is attached, read them loudly and spell them. Dictating sentences for learners to write down might be another opportunity to practice the spelling of multi-morphemic words. To accomplish the goal of language production, learners may be asked to make their own sentences using multi-morphemic words that they form attaching the new affix they have learned. They may also be encouraged to find as many words as they can with the suffixes, prefixes and common roots they learn, which will contribute to their vocabulary size.

According to Metalinguistic Hypothesis, which is “that some of the variance shared by vocabulary knowledge and reading comprehension can be attributed to metalinguistic abilities that impact both” (Nagy, 2007, p.71), learning vocabulary and reading comprehension are metalinguistic processes. In this regard, vocabulary instruction, for instance, is not teaching words only but teaching about words, which refers to teaching how words are formed, learned and used. As for how words are formed, morphological awareness plays a crucial role without any doubt. Besides, in the sense that morphological awareness is an integral part of metalinguistic awareness, it is worth noting some implications of the metalinguistic hypothesis here. First, more attention should be given to the metalinguistic demands of vocabulary instruction. For instance, Nagy (2007) states that providing more user-friendly explanations of word meanings instead of dictionary definitions is one step. In the case of multi-morphemic words, introducing common affixes and root words in English to learners is the very way of providing user-friendly explanations. Second, metalinguistic awareness should be increased as an instructional goal. To exemplify, it can be good to teach word parts, how learners can identify word parts through suffixes, synonyms, antonyms, use of context, dictionary skills and figurative language instead of teaching word meanings only. Lastly, vocabulary instruction and comprehension instruction can be integrated. As an example of such integration, Nagy (2007) suggests teaching word-learning strategies as a part of comprehension strategies. Making use of word morphology and context clues, for example, might be taught as a part of comprehension strategies.

Conclusion

In contrast to some other studies, this study did not find unique contribution of morphological awareness to reading comprehension among intermediate-level adult Turkish EFL learners. The explanation of why the current study differs from some of the previous research might lie in the differences between the current study and the previous ones such as English language learning context (ESL versus EFL), age group or L1 background of the participants, and the way morphological awareness was measured. That the findings of the present study are congruent with those of Zhang and Koda (2012), which is the only study carried out with adult EFL learners to our knowledge, supports this explanation to some extent.

Further analyses revealed that morphological awareness significantly explains 13.2 percent of the variance in vocabulary knowledge. Therefore, it can be concluded that although L2 morphological awareness fails to contribute to L2 reading comprehension directly, it might make an indirect contribution to reading comprehension via its direct contribution to vocabulary knowledge. Considering the indirect contribution of morphological awareness to reading comprehension, it can be stated that word formation rules and morphological awareness are important in EFL curriculum and pedagogy. Therefore, EFL learners should be made aware of lexical and grammatical functions of English derivational affixes and structure of English derivational words, and trained accordingly.

Limitations and Suggestions for Further Research

One potential limitation of the present study is its measurement. Because there is no established test of morphology in L2, an L1 morphological awareness test created with pseudo words (Mahony, 1993) was adapted for the accomplishment of the present study. A part of this test assessed learners' ability to recognize lexical and grammatical functions of some common suffixes in English, while another part of the test assessed learners' ability to judge whether two seemingly morphologically related words are in fact morphologically related or not. It is likely that this limited morphological insight may be the reason of the lack of contribution of morphological awareness to reading comprehension in the current study. Apart from the measurement, another potential limitation lies in the number of the participants. Although the current study involved a significant number of participants, a larger sample would have been preferred to reach at more generalizable results that could apply to a wider readership. Third, instructors' opinions regarding benefits of word morphology and its relation to reading comprehension could have been asked so that we could have gained insights into practicality or impracticality of morphological awareness instruction in EFL classrooms and found out some potential pitfalls.

As the first of its kind to model the intricate relationship of morphological awareness to reading comprehension and vocabulary knowledge among adult Turkish EFL learners, the current study sets an example for future research. Therefore, the current limitations should be considered as suggestions for further research. In other words, further research had better tap different levels of morphological awareness including more measures or indicators of morphological knowledge/awareness. Further research may also consider focusing on one type of derivational suffix (verb making derivational suffixes, for instance) and including prefixes as well as suffixes. Besides, the number of participants may be increased. Finally, EFL instructors' opinions with regard to the potential benefits of word morphology and its relation to reading comprehension may be a part of related further research.

References

Akkakoson, S. (2013). The relationship between strategic reading instruction, student learning of L2-based reading strategies and L2 reading achievement. *Journal of Research in Reading, 36* (4), 422-450.

Anjomshoa, L., & Zamanina, M. (2014). The effect of vocabulary knowledge on reading comprehension of Iranian EFL learners in Kerman Azad University. *International Journal on Studies in English Language and Literature, 2* (5), 90-95.

Apel, K., & Werfel, K. (2014). Using morphological awareness instruction to improve written language skills. *Language, Speech and Hearing Services in Schools, 45,* 251-260.

Block, C., & Pressley, M. (2007). Best practices in teaching comprehension. In L. Gambrell, L. Morrow & M. Pressley (Eds.), *Best practices in literary instruction* (pp. 220-242). New York: Guilford Press.

Brittain, M. (1970). Inflectional performance and early reading achievement. *Reading Research Quarterly, 6* (1), 34-48.

Carlisle, J. F. (1995). Morphological awareness and early reading achievement. In L. B. Feldman (Ed.), *Morphological aspects of language processing* (pp. 189– 210). Hillsdale, NJ: Erlbaum.

Carlisle, J. F. (2000). Awareness of the structure and meaning of structurally complex words: Impact on reading. *Reading and Writing: An Interdisciplinary Journal, 12,* 169-190.

Carrell, P. L. (1988). Introduction: Interactive approaches to second language reading. In P. Carrell, J. Devine & D. Eskey (Eds.), *Interactive approaches to second language reading* (pp. 1– 7). Cambridge: Cambridge University Press.

Droop, M., & Verhoeven, L. (2003). Language proficiency and reading ability in first- and second-language learners. *International Reading Association*, *38* (1), 78-103.

Fry, E. B., & Kress, J. E. (2006). *The Reading Teacher's Book of Lists*. (5th ed.). San Francisco, CA: Wiley.

Grabe, W. (2009). *Reading in a Second Language: Moving from Theory to Practice*. New York, NY: Cambridge University Press.

Grabe, W., & Stoller, F. L. (2001). Reading for academic purpose: guidelines for the ESL/EFL teacher. In M.Celce-Murcia (Ed.), *Teaching English as a second or foreign language* (pp. 187-203). Boston: Heinle and Heinle.

Guthrie, J., & Taoada, A. (2004). Fostering the cognitive strategies of reading comprehension. In J. Guthrie, A. Wigfield, & K. Perencevich (Eds.), *Motivating reading comprehension: concept-oriented reading instruction* (pp. 87-112). Mahwah, NJ: L. Erlbaum.

Haastrup, K. (1991). *Lexical inferencing procedures or talking about words: receptive procedures in foreign language learning with special reference to English*. Tübingen: Gunter Narr.

Hudson, T. (2007). *Teaching second language reading*. New York, NY: Oxford University Press.

Jeon, E. H. (2011). Contribution of Morphological Awareness to Second-Language Reading Comprehension. *The Modern Language Journal, 95*, 217-235.

Jeon, E. H., & Yamashita, J. (2014). L2 reading comprehension and its correlates: A meta-analysis. *Language Learning, 64* (1), 160-212.

Katz, L. A. (2004). *An investigation of the relationship of morphological awareness to reading comprehension in fourth and sixth graders*. Unpublished doctoral dissertation, University of Michigan, Ann Arbor.

Kieffer, M. J. (2014). Morphological awareness and reading difficulties in adolescent Spanish-speaking language minority learners and their classmates. *Journal of Learning Disabilities, 47* (1), 44-53.

Kieffer, M. J., Biancarosa, G., & Mancilla-Martinez, J. (2013). Roles of morphological awareness in the reading comprehension of Spanish-speaking language minority learners: Exploring partial mediation by vocabulary and reading fluency. *Applied Psycholinguistics*, *34*, 697-725.

Kieffer, M. J., & Lesaux, N. K. (2008). The role of derivational morphology in the reading comprehension of Spanish-speaking English language learners. *Reading Writing, (21)*, 783-804.

Kieffer, M. J., & Lesaux, N. K. (2012). Direct and indirect roles of morphological awareness in the English reading comprehension of native English, Spanish, Filipino, and Vietnamese speakers. *Language Learning, 62* (4), 1170-1204.

Kirby, J. R., Deacon, S. H., Bowers, P. H., Izenberg, L.,Wade-Woolley, L., & Parrila, L. (2012). Children's morphological awareness and reading ability. *Reading Writing, 25*, 389-410.

Koda, K. (2005). *Insights into second language reading: a cross-linguistic approach*. Cambridge: Cambridge University Press.

Kuo, L., & Anderson, R. C. (2006). Morphological Awareness and Learning to Read: A Cross-Language Perspective. *Educational Psychologist, 41* (3), 161-180.

Mahony, D. L. (1993). *The role of sensitivity to word structure in the development of reading skill*. Unpublished doctoral dissertation, Department of Cognitive Sciences, University of California, Irvine.

Mehrpour, S., Razmjoo, S., & Kian, P. (2011). The relationship between depth and breadth of vocabulary knowledge and reading comprehension among Iranian EFL learners. *Journal of English Language Teaching and Learning, 53*, 97-127.

Mochizuki, M., & Aizawa, K. (2000). An affix acquisition order for EFL learners: an exploratory study. *System, 28*, 291–304.

Moghadam, S. H., Zainal, Z., & Ghaderpour, M. (2012). A review on the important role of vocabulary knowledge in reading comprehension performance. *Procedia-Social and Behavioral Sciences, 66*, 555-563.

Mokhtari, K., & Reichard, C. A. (2002). Assessing students' metacognitive awareness of reading strategies. *Journal of Educational Psychology, 94* (2), 249-259.

Mokhtari, K., & Sheory, R. (2002). Measuring ESL students' awareness of reading strategies. *Journal of Developmental Education, 25* (3), 2–10.

Nagy, W. (2007). Metalinguistic awareness and the vocabulary-comprehension connection. In R. K. Wagner, A. E. Muse, & K. R. Tannenbaum (Eds.), *Vocabulary acquisition: Implications for reading comprehension* (pp. 52–77). New York, NY: The Guilford Press.

Paribakht, T. S., & Wesche, M. (1999). Reading and incidental L2 vocabulary acquisition: an introspective study of lexical inferencing. *SSLA, 21*, 195-224.

Peng, S., Siriyothin, P., & Lian, A. P. (2014). Reading strategy use and reading proficiency of Chinese undergraduate students majoring in English. *International Journal of Academic Research, 6* (2), 69-74.

Qanwal, Karim, & Karim, S. (2014). Identifying correlation between reading strategies instruction and L2 text comprehension. *Journal of Language Teaching and Research, 5* (5), 1019-1032.

Qian, D. D. (1999). Assessing the roles of depth and breadth of vocabulary knowledge in reading comprehension. *Canadian Modern Language Review, 56*, 282–307.

Sandra, D. (1994). The morphology of the mental lexicon: internal word structure viewed from a psycholinguistic perspective. *Language and Cognitive Processes, 9*, 227–269.

Schmitt, N., & Meara, P. (1997). Researching vocabulary through a word knowledge framework: word associations and verbal suffixes. *Studies in Second Language Acquisition, 19*, 17–36.

Tyler, A., & Nagy, W. (1990). Use of derivational morphology during reading. *Cognition, 36*, 17–34.

Wang, M., Cheng, C., & Chen. S. (2006). Contribution of morphological awareness to Chinese-English bilitarecy acquisition. *Journal of Educational Psychology, 98*, 542-553.

White, T. G., Sowell, V., & Yanagihara, A. (1999). Teaching elementary students to use word-part clues. *The Reading Teacher, 42*, 302-308.

Zhang, D., & Koda, K. (2012). Contribution of morphological awareness and lexical inferencing ability to L2 vocabulary knowledge and reading comprehension among advanced EFL learners: testing direct and indirect effects. *Reading Writing, 25*, 1195-1216.

Zhang, D., & Koda, K. (2013). Morphological awareness and reading comprehension in a foreign language: a study of young Chinese EFL learners. *System, 41*, 901-913.

Zhang, D., Koda, K., & Sun, X. (2014). Morphological awareness in biliteracy acquisition: a study of young Chinese EFL learners. *International Journal of Bilingualism, 18* (6), 570-585.

'Please, somebody talk': Incorporating Dialogical Critical Thinking into EFL Classes for Young Beginners

Roxy Lee
Dankook University, South Korea
Stewart Gray
Hankuk University of Foreign Studies, South Korea

Bioprofile:

Roxy Lee is an English teacher working at Annyung Elementary School in Gyeonggi province, South Korea, and a graduate of the Dankook University TESOL MA program. Her research interests include critical literacy, critical pedagogies, and creativity in English education. **Email:** roxy369@naver.com

Stewart Gray is an Assistant Professor in the Foreign Language Education Center of Hankuk University of Foreign Studies. He completed his MA TESOL at Dankook University and is a Ph.D. student at the University of Leeds. His research interests include language and identity, reflective practice, critical thinking, and critical pedagogies. **Email:** ec_391@hotmail.com (**corresponding author**)

Abstract

This writing details a four-month exploratory action research project conducted by two teacher-researchers with five, early-elementary age, Korean EFL students of mostly beginner level to explore ways of promoting dialogical critical thinking through English-language discussions. Researchers employed Kolb's learning cycle to generate reflective notes throughout the project, then performed a conventional content analysis on these and on representative class audio transcriptions. This analysis led the researchers to the following conclusions: (1) Teachers can encourage participation by designing game-like discussions; (2) teachers can encourage participation by allowing students to engage in *translanguaging* when discussing; (3) teachers can facilitate students' English use in discussions by providing language support, particularly discourse markers.; (4) teachers may promote discussion by encouraging students to interact with peers; (5) teachers can encourage students to generate critical discussion questions themselves, and connect personally with even

Dankook University, 152 Jukjeon-ro, Suji-gu,
Yongin-si, Gyeonggi-do, South Korea, 16890
Hankuk University of Foreign Studies, 81
Oedae-ro, Mohyeon-myeon, Cheoin-gu,
Yongin-si, Gyeonggi-do, South Korea, 17035

serious topics.

Keywords: Young learners, dialogical critical thinking, discussion, EFL, Korea

Introduction

While many commentators highlight 'critical thinking' (CT) as an essential skill for life in the modern world (Ellison, 2010; Kettler, 2014), teachers interested in CT face many uncertainties. Definitions of CT are diverse (Moore, 2011), as are proposed approaches for including it in education (Daniel & Auriac, 2009; Khajavi & Shahvali, 2013). There have been calls for further research on incorporating CT into English as a foreign language (EFL) classes (Ketabi, Zabihi, & Ghadiri, 2013), but while such research has been conducted with older students (Curtis, 2001; Dewaelsche, 2015), students studying in their first language (Kettler, 2014), and 'gifted' students (Connerly, 2006; Kettler, 2014), research on CT in classes of young, beginner-level EFL students has, to our knowledge, not been conducted. Therefore, two EFL teachers in Korea (the researchers) conducted a four-month exploratory action research project with a private class of five early-elementary-age students of mostly beginner English proficiency with the goal of promoting students' use of CT through English-language discussion, and exploring the answer to the following question:

- What practical considerations may be relevant for teachers seeking to incorporate critical thinking through discussion into beginner-level, young learners' EFL classes?

Literature Review

Dialogical Critical Thinking

Definitions of CT are numerous, and include idea evaluation (Daniel & Auriac, 2009), logical defense of a point (Leming, 2010; Van Gelder, 2005), and information processing skills like sequencing, classification, and comparison (Bloom, 1956; Fisher, 2006). Beyond these, there is the contrast between *monological* and *dialogical* CT (Benesch, 1999). Monological CT is the use of the aforementioned thinking skills to develop logical thinking capacity for its own sake (ibid), while dialogical CT is the use of discussion and dialog to challenge ideas and ‘taken-for-granted assumptions’ not simply to ‘exchange ideas but also to promote tolerance and social justice’ (ibid, p.576) as part of a consciousness-raising, liberational pedagogy (Freire, 1970). The researchers selected dialogical CT as the basis for their project because of their shared solidarity with the pedagogical ideals of Benesch and Freire.

Dialogical Critical Thinking in the EFL Curriculum

Many commentators affirm the suitability of EFL classrooms as a context for critical discussions, because the EFL class naturally provides students with the opportunity to learn a language while exploring various ‘cultures and ways of thinking’ (Yang & Gamble, 2013, p.398). Thus, L2 study can be complementary to CT engagement (Benesch, 1999; Pishghadam, Zabihi, & Kermanshahi, 2012). Furthermore, many activities involving CT, such as debate and problem solving, are useful for language practice (Khodanshas & Farahani, 2014). As to whether specifically dialogical CT should be incorporated, critical educators have long called for teachers of EFL to take responsibility for addressing social issues with students in order to help them develop simultaneously into competent English speakers and ‘critical agents in society’ (Ketabi, *et al*., 2013, p.16), a responsibility that English teachers merit because of the potentially oppressive function of international EFL education (Craig & Porter, 2014). For these reasons, the two researchers decided to incorporate dialogical CT into an EFL class of their own.

Encouraging Young Learners to Engage in Critical Thinking

CT is important for all students, including the young (Edmonds, Hull, Janik & Rylance 2005; Ellison, 2010; Kettler, 2014; Lipman, 2003), and the argument for CT in elementary education is decades old (Facione, 1990, in Kim, 2006). However, as a child’s level of cognitive development is likely lower than an adult’s, it is questionable whether a child can engage in CT above the most basic level (Daniel & Auriac, 2009; Leming, 2010). However, Benesch (1999) asserts that judgments about the appropriateness of CT must be based on classroom data, not essentialized, *a priori* views about students’ CT inabilities. Furthermore, Vygotsky argues that the young need to practice CT to grow into CT-competent adults (Vygotsky, 1989, in Shin, 2015, p.70), and educators such as Quinn (1997, in Hare, 1998) have shown through in-class research that young children can engage in critical discussions given ‘opportunity and support’ (ibid, p.470). Goswami and Brown (1990, in Kim, 2006) suggest that children can engage in analytical thinking at perhaps eight or nine years old, and UNESCO notes that children are naturally disposed to ask ‘radical questions’ (UNESCO, 2011, in Shin, 2015, p.66).

Critical Thinking in Korean EFL

Though questions have been raised about the perceived failure of the Korean education system to adopt CT seriously (McGuire, 2007), national curricula in Korea reflect an increasing sense of CT’s importance, even in elementary-school English education (Ministry of Education Science and Technology, 2015). Meanwhile, educators and researchers in Korea have produced research in myriad classroom contexts exploring practical approaches to CT. Shin and Crookes (2005) demonstrated that junior and senior high-school Korean students can engage in critical discussions as part of their English development. Craig and Porter (2014) encouraged Korean high-school students to engage in dialogical CT by drawing graffiti on newspapers and expressing themselves politically through their graffiti, thus demonstrating that such students can engage with topics of social concern. Dewaelsche (2015) designed questions based on the upper three levels of Bloom’s (1956) taxonomy of CT (analysis, evaluation, and creation) and employed them with EFL students at a Korean university, demonstrating that students with sufficient English can engage in critical dialog when encouraged. Huh and Suh (2015) used graphic novels with elementary-age, high-English-proficiency Korean students, demonstrating that such students can challenge ‘dominant ideolog(ies)’ (p.123).

These studies confirm the suitability of CT and of dialogical CT specifically for Korean students, that it is possible regardless of student age, and that it can fit within EFL education. However, none of these studies involve elementary-age, beginner-level EFL students. To the researchers’ knowledge, no such studies exist. This may be because, as Dewaelsche (2015) notes, one must have sufficient English skills to discuss in English. However, Canagarajah (2007) asserts that students capable of only basic English can discuss, so long as they use all of their linguistic resources. Furthermore, since discussion classes for young, beginner EFL learners have not been researched, and data is needed to make a judgment (Benesch, 1999), the researchers decided to conduct this project to explore the feasibility of doing dialogical CT in classes for such students and any significant practical considerations for doing so.

Methodology

Action Research

The researchers conducted a collaborative, exploratory action research project (Kemmis & McTaggart, 2005) in a small, private class of their own. With the goal of encouraging students to engage in dialogical CT, they planned and taught classes, and reflected on their in-class experiences following Kolb's (1984) experiential learning cycle: planning a change/intervention, enacting that change, observing the results, reflecting on the value of the intervention (Farrell, 2007), planning further changes, and sharing and discussing all of this with each other to aid the reflective process (Kemmis & McTaggart, 2005).

Participants and Research Site

This project involved five students (Table 1) and two teachers (the researchers, Table 2). Parents of participating students signed consent forms to allow classes to be recorded and for data gathered to appear in this writing. All participating students have been given pseudonyms. The classes were conducted in the communal room of the apartment complex at which one of the researchers resides. The room included no desks or chairs; students and teachers sat on sofas or the floor; a whiteboard was affixed to one wall. Student sampling was based on 'convenience' (Etikan, Musa, & Alkassim, 2016); initially, no age or level criteria were stated. Two students, Yeona and Somin, are the daughters of one of the researchers. The other three students lived locally and were recruited through online advertisements and word-of-mouth appeals for participants.

Four of the five students were in the third grade of elementary school during the project. Yeona was in first grade. Of the five, only Heeseon had spent significant time outside of Korea (in Australia). Her spoken English was the strongest. All other students were beginners. Somin and Jeongbin possessed limited English vocabulary and an overall positive attitude towards English practice; they often expressed ideas in English enthusiastically, but could not do this in grammatical sentences. Yeona and Jaeseok were early-beginner English students who expressed frustration in class more frequently than others. Students were of similar average age (elementary) and level (beginner), with some variation (Heeseon's level and Yeona's age), but classes with outlying students are a reality for EFL teachers (Chien, 2012), so the researchers judged it reasonable to proceed.

The two teachers/researchers were Roxy (Korean) and Stewart (British). Each teacher had around six years' experience teaching EFL in Korea mainly with young learners, and chose to collaborate on this research due to a shared belief, echoed in comments they received at academic conferences in Korea, that students do not currently receive sufficient encouragement to discuss critically in English classes, and a shared curiosity about how to remedy this issue. Both researchers were competent in English and Korean.

Table 1 Participating Students

Name	Gender	Born	Nationality	English learning experience/level
Yeona	Female	2009	Korean	No prior English education; Early beginner.
Somin	Female	2007	Korean	No private English education, public elementary after-school English class only; Vocabulary for simple communication.
Jaeseok	Male	2007	Korean	No private English education, public elementary after-school English class only; Early beginner.
Heeseon	Female	2007	Korean	Six months in Australia visiting family; One year private English education; Confident verbal communication.
Jeongbin	Male	2007	Korean	No private English education, public elementary after-school English class only; Vocabulary for simple communication.

Table 2 Participating Teachers

Name	Gender	Nationality	Professional background
Roxy	Female	Korean	Five-six years EFL teaching experience in Korea (four years in private schools for young learners, one-two years in public elementary school).
Stewart	Male	British	Five-six years EFL teaching experience in Korea (two years in public elementary school, two years in private schools for young learners, one-two years in university).

Designing Classes for Dialogical CT

The literature on CT in EFL education suggests several principles for designing beginner-level classes to promote dialogical CT: *Students' contribution of ideas should be maximized.* To create a comfortable learning environment for discussing ideas, teachers must not treat students as empty vessels apt to receive wisdom (Freire, 1970); they have thoughts and opinions worthy of respect (Freire, 1970; Khajavi & Shahvali, 2013). Therefore, discussion topics should relate to students' lives (Yang & Gamble, 2013) so they can contribute experiences and perspectives. Relatedly, to maximize topic relevance, *discussions should be based on suitable literature* (Yang and Gamble, 2013) containing content students find interesting and/or humorous. Furthermore, *students should be encouraged to discuss with peers.* Benesch (1999) describes a dialogical CT class as a 'balance between extended student contributions and gentle challenges by the teacher' (p.578). Also, encouraging students to discuss amongst themselves (Yang & Gamble, 2013) and assess each other's ideas (Ubaque Casallas & Pinilla Castellanos, 2016) may encourage them to actively defend their ideas, and thus encourage participation (Khodashenas & Farahani, 2014). Finally, *students should be encouraged to employ all of their linguistic resources in discussion.* Mixing of multiple languages by students to communicate complex ideas should be accepted/tolerated/promoted (Ellison, 2010; Lin & Wu, 2015). This mixing (*translanguaging*, Creese & Blackledge, 2010) enables students of all levels to discuss in EFL settings (Canagarajah, 2007; Lin & Wu, 2015). Teachers may support English use in discussion by providing language support (Ellison, 2010) such as particular discourse markers ('In my opinion...', etc.) to help them structure and express arguments in English (Ubaque Casallas & Pinilla Castellanos, 2016).

Project Curriculum

The researchers conducted classes over four months (Table 3), two 90-minute classes per week. Classes were designed around four English-language children's books (two-four weeks per book). The first three of these were taken from the Scholastic Brain Bank series (2007). The researchers judged these three books and the fourth, Anthony Browne's Piggybook (1990), to have content suitable for age-appropriate critical discussion (Yang & Gamble, 2013). English language to be practiced in class was chosen based on the books, and on researchers' predictions about students' language needs (Ellison, 2010). The two researchers each taught one class per week, and frequently discussed ways to (re)structure and improve classes. Throughout the project, activities to promote student engagement in discussion were designed by the researchers based on their judgment and adapted upon reflection of in-class experiences.

Data Collection

Multiple forms of qualitative data were collected for triangulation in analysis (Farrell, 2007):

- Complete audio recordings of all classes. Every class included discussion activities, which allowed for ongoing assessment (Shirkhani & Fahim, 2011), and constant data collection.
- In-class observation notes; taken in real time whenever an experience stood out to researchers.
- Reflective notes: made by both researchers immediately after every class based on recollection and observational notes following Kolb's (1984) learning cycle.
- Students' opinions: collected in class several times, and in a post-course member-checking session.

Table 3 Four-Month Curriculum Summary

Date	Book: Author (Year)	English focus	Critical themes	Example activities	Example discussion questions
1.14 -1.27	N/A	Basic conversation	Introducing discussion	Class discussion	Is it okay to eat meat?
2.2 - 2.11	Fat Fred: Evelyn Davidson (2007)	(1) I think… because… (2) Asking questions (five Ws)	(1) Express opinions (2) Make judgments	Team debate (questions raised by students)	(1) Do you think it's fair to have snacks all together except Fred? (2) Why did the two kids go to hospital with their dog?
2.18 -3.11	Molly's Friend Ben: Janine Scott (2007)	(1) I (don't) think so. (2) She/He can…	Consider the lives of people with disabilities	Discussing topic to personalize	(1) Would it be easy to live with being deaf? (2) Do you know anyone who is deaf? (3) What jobs can deaf people do?
3.15 - 4.8	The Yard Sale: Mary Atkinson (2007)	Simple past tense	(1) Make inferences (2) Assign value	Peer discussion	(1) What can you see in the picture? (2) What can/can't you sell?
4.12 -5.6	Piggybook: Anthony Browne (1990)	Household vocabulary	Consider family roles	Role-play: Argument	Mom and Dad are having an argument about housework.

Data Analysis

For analysis, the researchers pooled their reflective notes (two notebooks-worth), and transferred these into an Excel document. To identify and summarize key insights accumulated in the notes, the researchers, working separately, conducted an inductive, *conventional content analysis* (Hsieh & Shannon, 2015). This involved reading the reflective data through, re-reading it while identifying keywords and highlighting/coding these, re-reading it again to refine codes, then comparing examples of coded data and combining them to create categories. Once both researchers had individually identified categories of insights within the notes, they compared results with each other to identify similarities and differences, and thus improve the credibility of the categories (Lincoln & Guba, 1985). This final comparison produced five categories:

- Teachers encourage participation by designing discussions to be game-like.

- Teachers encourage participation by allowing students to mix languages when discussing.
- Teachers facilitate students' English use in discussions by providing language support, particularly discourse markers.
- Teachers promote discussion by encouraging students to interact with peers.
- Teachers can encourage students to draw personal connections with topics, and/or to generate discussion topics themselves.

Next, researchers selected excerpts from class audio recordings for each of the five categories, with excerpt selection based on statements made in the reflective notes. Having identified relevant excerpts, researchers then listened to and transcribed these in Excel. After transcription, researchers again performed a *conventional content analysis* (Hsieh & Shannon, 2015) to explore the ways in which students and teachers had used language, verbally interacted, expressed and challenged ideas (Benesch, 1999) and become engaged or disengaged in discussion during the class. From among these analyzed transcriptions, illustrative examples were selected for presentation in the findings section. Finally, the researchers took the five categories to participating students for discussion, again for credibility's sake (Lincoln & Guba, 1985). Statements students made in this post-course member-checking session broadly supported researchers' conclusions. A selection of these statements is included in the findings section.

It must be noted that because the data used for this project is observational and reflective, and because the process of data analysis involved selection and interpretation, the findings given should not be viewed as objectively representative of the events that transpired (Ellis & Bochner, 2000). Rather, the findings reflect the researchers' understandings of their experiences, and the interpretive reconstruction of those experiences through writing (Pavlenko, 2007). It is inevitable that the researcher's own positionality will influence a project with narrative dimensions (ibid). In this particular project, the researchers' focus on and desire to promote CT should be taken into consideration in understanding the data presented below. To the greatest extent possible, the researchers have striven to record, select and analyze data rigorously and transparently, and to make the presence of their own voices clear in the findings that follow.

Findings

The following data was originally collected in English unless otherwise stated. Korean data has been translated by the researchers, and is *italicized* and annotated with *(KOR)*. Some Korean expressions appear in the 'Translanguaging' section to show students' language mixing. Besides translation, data examples are unaltered; Students' statements have not been edited/corrected.

Game-like Discussions

> Stewart (Reflective note - January 14): By the end (of the discussion) Yeona was completely off task and unresponsive to instructions. I finished class with 'Animal charades.' Yeona said, 'It's okay if it's a game.'

Frequently, the researchers observed students asking to play games. As a result of observations of student disengagement in conversational discussion early on (January 14, above), the researchers attempted to reframe critical discussion as game-like and challenging to promote engagement; they added elements of competition and victory criteria. One such activity was team debate. In the following example, Roxy showed students an image from the book Fat Fred (Image 1) in which the family is eating, but the dog (Fred) may not eat because he is fat. Roxy asked 'Do you think it's fair to eat in front of Fred?'; Heeseon and Somin said it was not fair, Yeona, Jeongbin and Jaeseok said it was fair; Roxy divided students into two teams (fair/not fair), and instructed them to raise their hand if they wished to pitch arguments at one another; She found they were so enthusiastic that they were difficult to manage. This extract begins as Somin was making the case that the character 'Dad' is also fat, and therefore should not eat:

Debate (February 2):

110	Somin:	Dad is fat, *but I don't think it's fair...(KOR)* (IS INTERRUPTED)
111	Yeona:	(TO SOMIN, INTERRUPTING) *No! The adults take good care of the dog.*
112	Roxy:	(LAUGHING) Okay.

113 Jeongbin: (STAMMERING, TRYING TO SPEAK FAST) *Uh uh uh, because they're adults…*

114 Roxy: What?

115 Jeongbin: *Because they're adults…(KOR)* (STOPS MID-SENTENCE)

116 Yeona: (INTERRUPTING JEONGBIN) *They can choose to go to hospital if they need to (KOR).*

117 Roxy: So, *it's okay for the dad to eat because he's an adult?(KOR)*

118 Jeongbin: (INSTANTLY) *Yes, he can take care of his own health! (KOR)*

119 Yeona: (LOUDLY) *Then, the dog also...(KOR)* (IS INTERRUPTED)

120 Roxy: (INTERRUPTING) Ok, wait a minute, raise your hand. Yeona, yes.

121 Yeona: *You can see the dad is eating vegetables. (KOR)*

122 Somin: *But there are adults who don't take care of themselves. (KOR)*

123 Jeongbin: *Oh yeah, he is eating vegetables.(KOR)*

124 Yeona: (EMPHATICALLY) *There are vegetables in his sandwich! (KOR)*

Image 1 Fat Fred (Evelyn Davidson, 2007, p.9)

Positioning students in such a way that they had to justify their ideas to each other fostered enthusiasm to participate, to argue and reason. As can be seen above, students argued energetically. Yeona raised her voice and interrupted other speakers (lines 111, 116, 119 and 124), and Jeongbin's effort to speak caused his voice to break (line 113).

Roxy had to intervene to calm things down (line 120). Also, despite Yeona's emotion, she attempted to support her point logically (line 116) and using evidence from the picture (lines 121 and 124); this while participating in a discussion on the topic of fairness, a dialogical CT activity (Benesch, 1999).

The issue with the above debate was that it involved little English. Therefore, researchers experimented with English-use 'rules'. In the following example, Roxy initiated a role-play/discussion in which Somin and Jaeseok portrayed the mother-character from Piggy Book and challenged the father-character (Jeongbin and Yeona) about failure to do housework ('home job'). Thus, students engaged in discussion of the book's critical theme, domestic life, from the characters' perspectives. The scenario was explained, roles assigned, and students were instructed to argue freely in character. Crucially, the students were told to 'win' the role-play by using English. Mid-argument, Roxy intervened only to guide, ask clarifying questions (following Benesch's (1999) description of dialogical discussions) or discourage Korean use:

Role-play (May 3):

133	Jaeseok:	Why no home job?
134	Yeona:	*Because I want to play (KOR).*
135	Roxy:	No Korean.
136	Jeongbin:	I play.
137	Jaeseok:	Why play?
138	Jeongbin:	Your is why home job?
139	Somin:	(EMPHATICALLY) No home job and it is ugly home!
		(140-173: BACK AND FORTH IN ENGLISH [REDACTED])
174	Somin:	And said Mrs. Piggott (BOOK CHARACTER) is, 'No, you are not washing dishes, and clothes, and ironing.'
175	Yeona:	(SURRENDERING) Job I'll help. I'll help.
176	Somin:	Ok, come here.
177	Jaeseok:	Why help job?
178	Jeongbin:	Because I'm hungry, and brain is head black.

Although this role-play happened late in the project, the volume of English spoken including by the lowest-level students, stunned the researchers. The mother-characters challenged the father-characters (line 133). Jeongbin attempted to turn the argument around (line 138), but Somin shut him down (line 139). This exchange continued, in English, for some time, before Yeona gave in and agreed to do housework (line 175). The high volume of English students used in this role-play, and their enthusiastic participation, impressed upon the researchers the utility of a game-like organization of discussions, with rules, competition and an aspect of challenge. It is worth noting, however, that after the disappointing start of January 14, researchers observed students engaging well in simple (non-game-like) discussions also, an observation confirmed by post-course member-checking. Students were, they observed, engaged when they were given permission to speak partially in Korean, appropriate English language support, the space to discuss amongst themselves, and opportunities to personalize discussion topics, as described below.

Translanguaging

> Roxy (Reflective note - March 15): Somin was about to say something in Korean, I said, 'English, please,' and they all became passive. They stopped participating well, and I said, 'Please, somebody talk.' Only Heeseon talked. And I said, 'You should participate more to learn more, and it's okay to mix (English) with Korean.' And then, students started to talk.

Although students used English well when doing so was a game-rule (see May 3, above), researchers otherwise observed students becoming stressed and reticent when pressured to use English. After Roxy's experience on March 15 (above), the researchers began to allow mixing of Korean and English to facilitate participation. Shortly thereafter, the following observation was made during a discussion in which Roxy posed to students a question derived from the book 'The Yard Sale' to encourage them to discuss the value of things: 'What can't you sell?':

Discussion (March 22):

215 Jaeseok: I can't sell life.

216 Others: (IMPLYING AGREEMENT) I think so

217	Jeongbin:	(EMPHATICALLY) Life *palmyeon* (if sell) die!
218	Somin:	Life *if you sell it you become a slave (KOR).*

It was Jeongbin's language use in line 217 that caught researchers' attention; one Korean word amidst two English words; Jeongbin used the English he had, then completed his thought with Korean, resulting in successful communication. Similar language use was observed in the following exchange from the mother vs. father role-play (see also 'Game-like Discussion' section) even though this activity explicitly prohibited Korean. Somin challenged the father characters over their excessive eating of supermarket food:

Role-play (May 3):

230	Somin:	Every morning is super food eat, and every morning home food, uhm, every morning super food?
231	Jeongbin/ Yeona:	Yes, yes.
232	Somin:	You are very, very sick.
233	Roxy:	Why did you say you are very, very sick?
234	Somin:	Super food is *jomiryo* (artificial ingredients) big.
235	Roxy:	If you buy superfood, you will get sick. I think that's what she meant.
236	Jeongbin:	Super.. Eh…
237	Roxy:	Discuss.. Discuss.. If you eat ramen every day you will get sick. So she said if you eat… (IS INTERRUPTED)
238	Jeongbin:	(INTERRUPTING) I'm big money, plane go to ocean, eat water and fish.

In line 234, Somin used a Korean word to complete an argument about healthy eating that was otherwise English. The word was 'artificial ingredients,' a term beyond her English level. Thus, her first-language resources supplemented and facilitated her use of English. Accommodating *translanguaging* (Creese & Blackledge, 2010) may facilitate participation in discussions by students whose L2 is insufficient for the activity. Moreover, *translanguaging* need not preclude English practice. As can be seen in the examples, students employed it for words and forms they did not know in English, whilst endeavoring to use English where possible. In the post-course member-checking session, students mostly expressed agreement with this idea, although Yeona hedged slightly: *'What we're learning is English, so using Korean is strange.' (KOR)*

Language Support and Discourse Markers

Discussion (February 16):

441 Roxy: Would it be easy to live with being deaf?

442 Somin: No, because Mom sad. *We'd need to learn sign language (KOR).'*

To support English use, researchers taught discourse markers (Ubaque Casallas & Pinilla Castellanos, 2016) to facilitate/guide idea expression, including 'I think so,' 'I don't think so,' 'I think…,' 'Why,' 'Because…,' and 'What do you think?'; Researchers then frequently encouraged students to use these phrases, as in the following example. Taking inspiration from the book The Yard Sale, Stewart asked 'What can and can't you sell?' and directed students to use the discourse markers to express themselves and question each other. This excerpt takes place after Heeseon, Yeona and Jeongbin have already spoken:

Discussion (March 18):

350 Jeongbin: (TO JAESEOK) What do you think?

351 Jaeseok: I(…) think(…) you(…) can(…) sell, (ELONGATED) uh~(...) socks.

352 Somin: (DISGUSTED SOUND)

353 Stewart: Socks?

354 Jaeseok: Yes.

355 Yeona: (EMPHATICALLY) No! I don't think so!

356 Stewart: Because?

357 Jaeseok: Because, *I have plenty of socks at home (KOR).*

358 Others: (LOUDLY) I don't think so!

359 Stewart: Somin, I think so? I don't think so?

360 Somin: I don't think so.

361 Stewart: (LAUGHING) Wow! No, no, no.

362 (MANY STUDENTS TALKING LOUDLY AT ONCE)

363 Yeona: (LOUDLY) Really, really, really smell!

Although Jaeseok used Korean to express his thoughts (line 357), he was able to use the provided structure to make his initial statement in English (line 351). The other students clamored energetically to respond in disgust using the phrases ‘I think so’ and ‘I don’t think so’. Even the lower-level students were thus able to use English to participate (line 355). Students later affirmed the usefulness of such phrases in member-checking. Furthermore, the researchers found that, having taught discourse markers and permitted students to mix languages, it became easier to encourage them directly to express their thoughts at least partly in English. For example, on one occasion Stewart showed a series of images to students to provoke interpretations of what they could see, including one of a seemingly unhappy couple, and another of school students standing in a hall:

Picture reading (March 25):

381 Stewart: What can you see?

382 Somin: *Couple fight (KOR).*

383 Stewart: In English?

384 Heeseon: (TO SOMIN, IN SUPPORT) I can see a…

385 Somin: I can see a(…) mom and dad(…) fighting.

386 (STEWART SHOWS PICTURE: SCHOOL STUDENTS STANDING IN CORRIDOR)

387 Stewart: Why are they standing here? Somin?

388 Somin: *People(…) (KOR)*

389 Stewart: (TO SOMIN) People.

390 Somin: Uh(…) *they were going to the hall, but they were very noisy, and the teacher got angry (KOR).*

391 Stewart: Can you give me an “I think”?

392 Somin: I think they(…) *went to the hall, but they were chatting, so they got*

		punished (KOR).
393	Stewart:	Wait. English, please.
394	Somin:	*I can't (KOR).*
395	Yeona:	(TO SOMIN) *Just mix English and Korean (KOR).*
396	Somin:	*What is hall in English (KOR)?*
397	Stewart:	Uh, hall.
398	Somin:	Hall go, to the walking, teacher and friends, many talking and teacher angry, *punishes them (KOR).*

Both Stewart and Heeseon supported Somin by offering English vocabulary (lines 384 and 397), Yeona recommended that Somin employ *translanguaging* (line 395), and Stewart encouraged Somin to use a discourse marker (line 391). The result of this was Somin (in line 398) expressing herself, despite her hesitancy (line 394), almost entirely in English. From this, researchers concluded that encouragement, support, and tolerance of language mixing and grammatical imperfection were conducive to English use, and by extension to English learning, as the process of negotiation to successfully output English contributes to language acquisition (Swain, 1985).

Peer Interaction

> Stewart (Reflective note - February 25): Students changed their opinions often (in discussion). Usually this seemed to be because their initial declaration had been hasty, and after some thought they arrived at a different conclusion. Not always though, sometimes it seemed they had their minds changed by their peers.

The researchers often interpreted student disengagement from discussion as being the result of the researchers' own excessive management of the conversation. In a discussion conducted on January 14 about eating meat, Stewart positioned himself at the center things, asking students questions and giving answers; students were unenthusiastic. Stewart's initial interpretation was that the topic had been age-inappropriate. However, soon after, Roxy observed students begin a spontaneous, Korean-language discussion on the necessity of learning English. Roxy had commented on Heeseon's enviable English level, which was the highest in the class.

Jeongbin spoke up, declaring that he was not envious, and that English was not important to learn. In response, other students ventured their own thoughts:

Spontaneous discussion (February 2):

406	Jeongbin:	*Even if I'm good at English, it's not useful (KOR).*
407	Roxy:	*Why isn't it useful?(KOR)*
408	Jeongbin:	*Where can I use it? (KOR)*
409	Roxy:	(INAUDIBLE)
410	Jeongbin:	No.
411	Others:	Yes.
412	Yeona:	*If you go to other countries.(KOR)*
413	Heeseon:	*If you go to Paris, in other countries English is basic, so even though you say it's useless.(KOR)*
414	Yeona:	*And if you go to other countries, you have to know English, they speak English (KOR).*
415	Jeongbin:	*When you go to other countries, you can ignore what people are saying to you (KOR).*
		(416-422: DISCUSSION BACK AND FORTH [REDACTED])
423	Roxy:	Somin, what do you think? *They are talking about whether we have to learn English. Heeseon thinks you have to learn English, so (KOR),* what do you think?
424	Somin:	Heeseon. (IMPLYING AGREEMENT WITH HEESEON)
425	Jeongbin:	*I was born in Korea (KOR).*
426	Roxy:	Yes, you're right.

Jeongbin did not resent studying English normally, so Roxy found his assertions surprising. Instantly, Yeona and Heeseon defended English as useful for travel (lines 412-414), which Jeongbin dismissed (line 415). Jeongbin persisted, though unsupported, and made reference to his being Korean to defend his point (line 425). Observing the interpersonal tensions felt even by a young child as a result of having to compete with his peer, Heeseon, the researchers decided that students should be encouraged to discuss directly with each other to foster engagement, generate argumentation, and promote participation. To this end, the researchers had students engage in team debates (see ‘Game-like Discussion’ section). In debate, the researchers observed that students would defend their ideas and occasionally question them after hearing their peers’ arguments. In the following example, Stewart conducted a debate on the question ‘What jobs can deaf people do?’ based on the book Molly’s Friend Ben, to exploring the topics of deafness and disability. The extract begins with Somin giving one possible answer:

Debate (February 25):

452 Somin: *A house cleaner (KOR).*
453 Stewart: Deaf people can be house cleaners?
454 Jeongbin: I think so!
455 Jaeseok *No, no, no, no(…) I say, 'Clean here, please,' but they can’t hear me, so they clean somewhere else (KOR).*
456 (SHORT SILENCE)
457 Heeseon: Maybe.
458 Jeongbin: *Oh, that’s right (KOR).*
459 Yeona: (LOUDLY) *Oh! They can’t be house cleaners! (KOR).*

Students’ engagement with Jaeseok’s argument (lines 457, 458 and 459) was an example of dialogical CT (Benesch, 1999). Students challenged each other’s ideas and evaluated their own through discussion (though the conclusion students reached, in this case, was not a point of view the researchers would have chosen to promote). However, while students sometimes modified their ideas in response to argumentation, researchers also observed students bowing to peer pressure rather than logic. In a later discussion of the above question, Stewart observed the following interaction. All students but Jaeseok declared that deaf people with hearing aids could work as school teachers, a position with which Jaeseok disagreed:

Discussion (March 4):

475 Stewart: Jaeseok, what do you think?
476 Jaeseok: *If they use a hearing aid, they can’t be a teacher (KOR).*

477 Stewart:	Oh! So we've got one for, 'No.' Why?
478 Jaeseok:	*Uh, because they may get teased by the students, and won't be able to come to school because of the shame (KOR)*
	(479-483: ALL PEERS DISAGREE AND GIVE VARIOUS REASONS [REDACTED])
484 Jaeseok:	(TO STEWART) *I want to change my vote (KOR).* (FROM 'NO' TO 'YES')
485 Stewart:	Oh yeah? Why?
486 Jaeseok:	*Just because (KOR).*
487 Stewart:	*Did your opinion change (KOR)?*
488 Jaeseok:	*No (KOR).*
489 Stewart:	*Well then, why (KOR)?*
490 Jaeseok:	*Because all my friends said, 'Yes' (KOR).*
491 Stewart:	*But they could be wrong (KOR).*
492 Jaeseok:	*Even so (KOR).*

Thus, while researchers appreciated for the potential benefits of student-student interaction for dialogical discussion and tried to encourage it, they also observed that it was not a panacea, and that CT could not be reliably expected in peer discussion.

Topic Personalization

> Roxy (Reflective note - April 5): I feel that they engaged more with the conversation because it started from their perspective. It's good to have more perspectives reflected in the questions we use, not just the teacher's. Having them make the questions puts them at the center of the class.

Researchers frequently directed students to create English questions about the books. These questions were then used as discussion questions. In the following example from April 5, Roxy asked all students to select one page from either of two books, The Yard Sale and Fat Fred, and directed them to make questions for discussion ('without an easy answer') based on the picture on that page. She then helped them to formulate the questions in English. Roxy then used these questions to initiate discussion. Somin and Jeongbin's questions referred to an image in The Yard Sale (Image 2), showing a band on stage, and at the bottom of the page some hands of varying skin tone. The students' questions concerned the characters' genders:

Students questions (April 5):

523	Somin:	How many girls and boys in the show?
524	Yeona:	Girl two, boy one.
525		(ALL STUDENTS AGREE)
526	Roxy:	How do you know that?
527	Heeseon:	Long hair.
528	Roxy:	If somebody has a long hair, is it a girl?
529	Somin:	*Not all long-haired people are girls (KOR).*
		(530-532: STUDENTS DISCUSS LONG-HAIRED MALE TV PERSONALITY [REDACTED])
533	Jeongbin:	Is this boy's hand and girl's hand? (REFERRING TO HANDS, BOTTOM OF PAGE, MIDDLE OF THREE)
534	Others:	It's girl's.
535	Roxy:	Why do you think so?
536	Somin:	Color is *light (KOR).* Boy is *dark (KOR),* girl is *light (KOR).*
537	Roxy:	Everybody show me your hands. Is it true girl's hand is lighter than boy's hand?
		(STUDENTS EXAMINE EACH OTHER'S HANDS)
538	All students:	No.

Because students' questions (lines 523 and 533) focused on gender, Roxy was able to encourage the students to question their assumptions about gender appearance (lines 528 and 537), which they did (lines 529 and 538). This affirmed the potential for the students to decide topics for themselves, and to generate questions for dialogical discussion with the teacher's support. Roxy's impression (above, April 5 reflective note) was also that students were enthusiastic in these discussions because they had chosen the questions themselves. In an effort to maintain this enthusiasm in

Image 2 The Yard Sale (Mary Atkinson, 2007, p.15)

cases where the researchers themselves posed discussion questions, they asked supplementary questions to help students make personal connections with topics. In the following example, Roxy initiated a discussion about blindness and deafness based on the book Molly' Friend Ben. After confirming that students understood the terms, she asked them if they knew anyone who was blind or deaf:

Discussion (February 16):

502 Somin: *Upstairs,* (REFERRING TO THE HOUSE ABOVE HERS) *the Mom can't hear (KOR).*

503 Heeseon: *I had a friend in first grade who couldn't hear (KOR).* I was sad

504 Jeongbin: Yes, watching TV, EBS. *Deaf people in hospital*. *It was interesting (KOR).*

505 Jaeseok: *I saw someone on the subway with sunglasses and a stick.*

Although deafness and disability are serious topics, the above statements show that the young students in our class had experiences and feelings to contribute (Khajavi & Shahvali, 2013). In the follow-up discussion, students talked, unprompted, about sign language and hearing aids, further demonstrating their pre-existing familiarity with discourses of deafness. This confirmed for researchers the appropriateness of such topics for young learners.

Discussion

This research supports the possibility of including dialogical CT (Benesch, 1999) in beginner-level, young-learner EFL classes. Even the lowest-level students in this project discussed at least partly in English when teachers took steps to promote and support participation. As to those steps, this research highlights five relevant principles for teachers seeking to include dialogical CT in young beginners' EFL classes. Firstly, students may discuss, express opinions and argue points enthusiastically if the discussion is game-like and challenging. Researchers generated this sense of challenge with confrontational role-plays and debates, confirming the usefulness of role-plays to address critical topics (Connerly, 2006), and the value of debate for fostering competition (Khodashenas & Farahani, 2014).

Secondly, this research indicates language mixing for communication should be allowed. The young, beginner-level EFL students in this project discussed using rudimentary English (Canagarajah, 2007) if they were allowed speak Korean for expressing complex ideas (Ellison, 2010). Permitting *translanguaging* (Creese & Blackledge, 2010) facilitates participation in discussion by students whose English is insufficient to be the sole language of participation, and in supporting their L2 with their L1 students may develop L2 skills through conversational practice that might otherwise be closed to them. Examples of *translanguaging* in this paper resemble those in Lin and Wu's (2015) study of Hong Kong high-school students. This research supports Lin and Wu's argument for the use of students' varied language resources for discussing in EFL contexts, and extends this argument to elementary-level, Korean contexts.

Thirdly, this research supports the teaching of *discourse markers*. Researchers in this project taught several ('I think so,' 'because,' etc.), and encouraged students to use them to frame arguments. This increased English use by students in discussion, and, as in Ubaque Casallas and Pinilla Castellanos's (2016) project, gave students structure for expressing opinions and justifications. While Ubaque Casallas and

Pinilla Castellanos taught discourse markers to older students, this research shows their applicability for younger students.

Fourthly, this study supports encouragement of students to discuss and debate with peers. This promotes proactive justification of ideas (Khodashenas & Farahani, 2014) and presents students with opportunities, relevant to dialogical CT (Benesch, 1999), to challenge the ideas of others and revise their own. The value of social interaction and peer-judgment for promoting CT are affirmed in the literature (Ubaque Casallas & Pinilla Castellanos, 2016; Yang & Gamble, 2013), and this project, too, confirms that students may be encouraged to defend and evaluate their ideas when in direct discussion with peers. However, as a cautionary counter-example, an instance is provided of a student seeking to conform to their peers' ideas. Social pressure should be accounted for when planning dialogical discussions.

Fifthly, this paper demonstrates that young students can, with teacher support, create questions that can be used for dialogical discussion. Doing so may ensure that discussion topics are relevant to them (Yang & Gamble, 2013) and create opportunities for students to express their thoughts, experiences and perspectives (Khajavi & Shahvali, 2013). Likewise, this research confirms that even young students can contribute perspectives and experiences relevant to serious discussions. The summary argument of this research is optimistic: young, beginner-level EFL students can engage in dialogical CT and discussion in English class. EFL teachers hoping to encourage this may do well to design game-like discussions, and to provide students with English language support, freedom to use their varied language resources, and the space and opportunity to share and negotiate their ideas together. Given all of this, they may discuss quite well.

Conclusion

Students in this project numbered only five. The results of this project may be applied to larger classes in conventional classrooms, but this may have unique challenges/benefits. A similar project conducted with a larger group of young, beginner-level EFL students in different contexts may confirm or modify this paper's arguments, and illuminate factors unique to other settings. Furthermore, while the researchers partially came upon the five principles highlighted in this research through the trial-and-error reflective process of action research, a project that was designed from its inception with these principles in mind might be valuable to further explore their potential applicability to various students, teachers and settings. Finally, the researchers selected books and, often, discussion topics, but future studies might allow students to select the books for reading and discussing, in order to prioritize students' thoughts and feelings, and facilitate student-led identification of critical issues (Khajavi & Shahvali, 2013).

The results of this project support the view that the inclusion of critical discussion in classes for young L2 learners is possible (Ellison, 2010; Kettler, 2014; Lipman, 2003), even for beginners. Our young students demonstrated willingness and capacity for discussing topics relating to fairness and justice, which is to say for engaging in dialogical CT (Benesch, 1999). Further, they did so partially in English, despite their modest levels (Canagarajah, 2007). The opinions that students expressed, the arguments they made, and the English they used show that critical discussion can have a place in young learners' EFL education. The researchers hope this project will interest practicing teachers of young, beginner-level EFL learners, and those designing curricula for such learners, as it supports the feasibility of incorporating dialogical, critical discussion into their English classes. This project may also inform the development of teacher training programs focused on CT, highlighting as it does a number of teachable instructional strategies, questions, and activities that pre-/in-service teachers can use to facilitate student participation in fruitful, critical, L2 discussion. The five students in this study all possessed ideas, perspectives and some natural capacity for arguing critically. What the researchers provided for these students were English classes in which they were encouraged to explore critical topics in an enjoyable fashion, and to bring their natural abilities, varied linguistic resources, and unique thoughts into the conversation.

References

Benesch, S. (1999). Thinking critically, thinking dialogically. *TESOL Quarterly, 33(3)*, 573-580.

Bloom, B. S. (1956). *Taxonomy of educational objectives.* Boston, MA: Allyn and

Bacon.

Canagarajah, S. (2007). Lingua franca English, multilingual communities and language acquisition. *The Modern Language Journal, 91,* 923-939.

Chien, C-W. (2012). Differentiated instruction in an elementary school EFL classroom. *TESOL Journal, 3*(2), 280-291.

Connerly, D. (2006). *Teaching critical thinking skills to fourth grade students identified as gifted and talented* (MA dissertation). Graceland University, Iowa.

Craig, M., & Porter, C. (2014). "Speaking back" from the English periphery: Art-work in a South Korean high school English classroom. *English Teaching: Practice and Critique, 13*(2), 35-54.

Creese, A., & Blackledge, A. (2010). Translanguaging in the bilingual classroom: A pedagogy for learning and teaching?. *The Modern Language Journal, 94*(1), 103–115.

Curtis, A. (2001). Hong Kong secondary school teachers' first experiences of action research. *Journal of the Pan-Asian Consortium, 1*(1), 65-78.

Daniel, M. F., & Auriac, E. (2009). Philosophy, critical thinking and philosophy for children. *Educational Philosophy and Theory, 43*(5), 415-435.

Dewaelsche, S. A. (2015). Critical thinking, questioning, and student engagement in Korean University English courses. *Linguistics and Education, 32,* 131-147.

Edmonds, M., Hull, J. A., Janik, E. L., & Rylance, K. (2005). *History and critical thinking: A handbook for using historical documents to improve students' critical thinking skills in the secondary grades.* Wisconsin, US: Wisconsin Historical Society.

Ellis, C., & Bochner, A. P. (2000). Autoethnography, personal narrative, reflexivity: Researcher as subject. In N. K. Denzin & Y. S. Lincoln (Eds.), *Handbook of qualitative research, second edition* (pp. 738-768). Thousand Oaks, CA: SAGE.

Ellison, M. (2010). Make them think! Using literature in the primary English language classroom to develop critical thinking skills. *E-F@BULATIONS.* Retrieved from ler.letras.up.pt/uploads/ficheiros/8664.pdf

Etikan, I., Musa, S. A., & Alkassim, R. S. (2016). Comparison of convenience sampling and purposive sampling. *American Journal of Theoretical and Applied Statistics, 5*(1), 1-4.

Farrell, T. S. C. (2007). *Reflective language teaching: From research to practice.* London, UK: Continuum.

Fisher, R. (2006). Thinking skills to thinking schools: Ways to develop children's thinking and learning. *Early Child Development and Care, 153*(1), 51-63.

Freire, P. (1970). *Pedagogy of the oppressed.* New York, NY: Herder and Herder

Hare, W. (1998). [Review of the book *Critical thinking in young minds*, by V. Quinn].

Canadian Journal of Education, 23(4), 468-470.

Hsieh, H. F., & Shannon, S. E. (2015). Three approaches to qualitative content analysis. *QUALITATIVE HEALTH RESEARCH, 15*(9), 1277-1288.

Huh, S., & Suh, Y. (2015). Becoming critical readers of graphic novels: Bringing graphic novels into Korean elementary literacy lessons. *English Teaching, 70*(1), 123-149.

Kemmis, S. & McTaggart, R. (2005). Participatory action research: Communicative action and the public sphere. In N. K. Denzin & Y. S. Lincoln (Eds.), *Handbook of qualitative research, 3rd edition* (pp. 271-330). London, UK: SAGE.

Ketabi, S., Zabihi, R., & Ghadiri, M. (2013). Critical thinking across the ELT curriculum: A mixed methods approach to analyzing L2 teachers' attitudes towards critical thinking instruction. *International Journal of Research Studies in Education*, *2*(3), 15-24.

Kettler, T. (2014). Critical thinking skills among elementary school students: Comparing identified gifted and general education student performance. *Gifted Child Quarterly, 58*(2), 127-136.

Khajavi, Y. & Shahvali, N. (2013). Promoting critical thinking in the language classroom. *Humanising Language Teaching, 15*(3). Retrieved from http://www.hltmag.co.uk/jun13/sart05.htm

Khodashenas, M. R., & Farahani, S. K. (2014). The effect of critical thinking on Iranian EFL learners' speaking ability. *Modern Journal of Language Teaching Methods, 4*(4), 196-202.

Kim, J. S. (2006). Use of fables to teach critical thinking to elementary students. *Sagogaebal*, *2*(1), 41-59.

Kolb, D. (1984). *Experiential learning: Experience as the source of learning and development.* Englewood Cliffs, NJ: Prentice Hall.

Leming, J. S. (2010). Some critical thoughts about the teaching of critical thinking. *The Social Studies, 89*(2), 61-66.

Lin, A. M. Y., & Wu, Y. (2015). 'May I speak Cantonese?' – Co-constructing a scientific proof in an EFL junior secondary science classroom. *International Journal of Bilingual Education and Bilingualism*, *18*(3), 289-305.

Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic inquiry*. Beverly Hills, CA: Sage.

Lipman, M. (2003). *Thinking in education.* Cambridge, UK: Cambridge University Press.

McGuire, J. M. (2007). Why has the critical thinking movement not come to Korea?. *Asia Pacific Education Review, 8*(2), 224-232.

Ministry of Education Science and Technology. (2015). *English education curriculum* (Department publication number 2015-74). Seoul, Korea.

Moore, T. (2011). Critical thinking: Seven definitions in search of a concept. *Studies in Higher Education, 38*(4), 506-522.

Pavlenko, A. (2007). Autobiographic narratives as data in applied linguistics. *Applied Linguistics, 28*(2), 163-188.

Pishghadam, R., Zabihi, R., & Kermanshahi, P. M. (2012). Educational language teaching: A new movement beyond reflective/critical teaching. *Life Science Journal, 9*(1), 892-899.

Shin, H. J., & Crookes, G. (2005). Exploring the possibilities for EFL critical pedagogy in Korea: A two-part case study. *CRITICAL INQUIRY IN LANGUAGE STUDIES: AN INTERNATIONAL JOURNAL, 2*(2), 113–136.

Shin, J. Y. (2015). A study on teaching critical thinking skills in elementary schools. *Kyoyukhwahakyeongu, 17*(1), 65-88.

Shirkhani, S., & Fahim, M. (2011). Enhancing critical thinking in foreign language learners. *Procedia – Social and Behavioral Sciences, 29*, 111-115.

Swain, M. (1985). Communicative competence: Some roles of comprehensible input and comprehensible output in its development. In S. Gass & C. Madden (Eds.), *Input in second language acquisition* (pp. 235–253). Rowley, MA: Newbury House.

Ubaque Casallas, D. F., & Pinilla Castellanos, F. S. (2016). Argumentation skills: A peer assessment approach to discussions in the EFL classroom. *PROFILE Issues in Teachers' Professional Development*, *18*(2), 111-123.

Van Gelder, T. (2005). Teaching critical thinking: Some lessons from cognitive science. *College Teaching, 53*(1), 41-46.

Yang, Y-T, & Gamble, J. (2013). Effective and practical critical thinking-enhanced EFL instruction. *ELT Journal*, *67*(4), 398-412.

Constructing, Reconstructing and Developing Teacher Identity in Supportive Contexts

Mehmet Haldun Kaya
İzmir University of Economics, Turkey
Kenan Dikilitaş
Bahçeşehir University, Turkey

Bioprofile:

Mehmet Hâldun Kaya is currently a PhD student at Bahçeşehir University and works as an instructor and a CELTA tutor at İzmir University of Economics. He is an experienced teacher trainer offering workshops on technology-driven courses and teaching language skills. His research interests include teacher development and integration and use of technology in teacher education programs. **Email:** mehmethaldunkaya@gmail.com

Kenan Dikilitaş is Associate Professor in the ELT department at Bahçeşehir University, Turkey. His primary research interests include English language teacher education and teacher professional development. He has published articles and books on action research and teacher research, and conducted several teacher research projects with local and international teachers. **Email:** kenan.dikilitas@es.bau.edu.tr **(corresponding author)**

Abstract

Teacher identity has been one of the recent topics that has created a perspective to investigate teachers' development by reflecting on who they are and how they do what they do in a certain context, namely an Asian context in this paper. This qualitative research exploring a language teacher's identity development of a novice English language teacher was carried out as a case study. A triangulation model using semi-structured interviews, observations, narratives, and/or reflective writing brought together data from a longitudinal perspective involving 6-month monitoring of Derviş' knowledge, beliefs, and practices. Derviş' practices and reported beliefs were documented and analysed over six months. The research concluded that a novice language teacher constructed and developed his identity from behaviourist perspective into socio constructivist one within a supportive context where professional development is key part of the language programme.

Izmir Ekonomi Üniversitesi, Sakarya Caddesi,
No: 156, Balçova – İzmir, Turkey 35330
Bahçeşehir University Çırağan Caddesi,
Osmanpaşa Mektebi Sokak, No: 4 – 6,
Beşiktaş, İstanbul, Turkey 34353

Keywords: Teacher identity, teacher development, identity construction, qualitative research

Introduction

For many years, language teachers were regarded as merely the practitioners of approaches or theories (Varghese, Johnson & Johnson, 2005). The nature of the teaching and learning environment, however, has proved that relegating language teacher to technicians is insufficient to define their role inside or outside the classroom (Ng & Boucher-Yip, 2016). Understanding teachers' characters has turned out to be a prerequisite to understanding language teaching and learning (Varghese *et al.*, 2005) since learning to teach is not related to finding out new skills and knowledge in language teaching but it has roots in being a language teacher as well (Richards, 2010). Yet, it remains unclear or whether educators can help teachers (re/co) construct their identity. This paper considers that a teacher's professional knowledge, beliefs regarding learning and teaching, along with the classroom practices, are in fact the reflections of identity. Thus, an understanding of a teacher's identity along with the factors contributing to the identity construction, reconstruction and development is of paramount importance in teacher education and continuous professional development. To this effect, teacher identity has emerged as the key focus for many researchers.

This case study aims to explore a novice English language teacher's (Derviş) identity development from a longitudinal perspective which involves a six-month of monitoring of his knowledge, beliefs and practices in the contexts of a training course and a foundation university.

Literature Review

Identity is "a being in continuous becoming" (Roth, 2003 p.8); thus, teachers are generally in the process of developing new identities. Different learning theories define the nature of teacher identity in their own particular terms. Mayer (1999, p.8) argues that one can become a practitioner with expertise through teaching but becoming a true professional may require 'intellectual dimension". From a cognitivist view, teachers can shape and re-shape their own identities through cognitive and meta-cognitive abilities. Zacharias (2010) also argues that teacher identity is related to beliefs, attitudes and metacognitive processes. Hence, cognitivists claim that identity is constructed, not ready-formed. In the process of construction, experimenting and experiencing new roles help teachers be “free from unthinking reliance on received knowledge” (Johnson & Golombek, 2016, p.35). However, constructivism, like behaviourism and humanism, focuses exclusively on the psychological and mental dimensions, and neglects the social role of teaching, the relations of a teacher, and the expectations of others. This directs us to the socio-constructivist theory, since teacher identity is a social phenomenon that can be constructed, reconstructed or developed in a social environment. For example, Coldron and Smith (1999) describe being a teacher as a matter of evaluating the professional role not only by himself but also by other people. They also underscore the social acceptability and changing nature of the teacher identity. Varghese *et al.* (2005), similarly, regard the formation, negotiation, and growth of teacher identity as essentially social process taking place in institutional settings. Developing and evolving in a professional context, a teacher’s identity is then shaped and reshaped in interaction with others (Beauchamp & Thomas, 2009) and ‘negotiated through experience’ (Sachs, 2005, p.15). Hence, it can be argued that identity is socially constructed, as it is usually defined considering a person’s social relations and participation within the community. In the process of becoming a member of a community, the teacher can take on a great many sub-identities, based on the context and relationships. Before focusing on what influences the formation or reconstruction of a teacher’s identity, it is appropriate to pay attention to the different aspects of teacher identity.

Components of Teacher Identity

Though the common features of teacher identity have been determined, it is important to clearly understand how an identity is formed and changes, what triggers this change, and what happens in this shifting process. The development of identity development is inevitably affected by personal, institutional, cultural and global factors (Ivanova & Skara-Mincane, 2016). Figure 1 shows that a teacher’s past and present, personal and professional lives are inter-related; how a teacher’s self is interaction with the context; and “how each teacher component mediates (and is mediated by) the others” (Olsen, 2008, p.25).

The formation and development of teacher identity take place in contextual settings (Ivanova & Skara-Mincane, 2016). These contexts can be characterized by providing continuous professional development, teacher-sensitive mentoring, and constructive feedback. Such characteristics are the determining factors which might help assess whether teachers can develop or change their identities as they teach in a context.

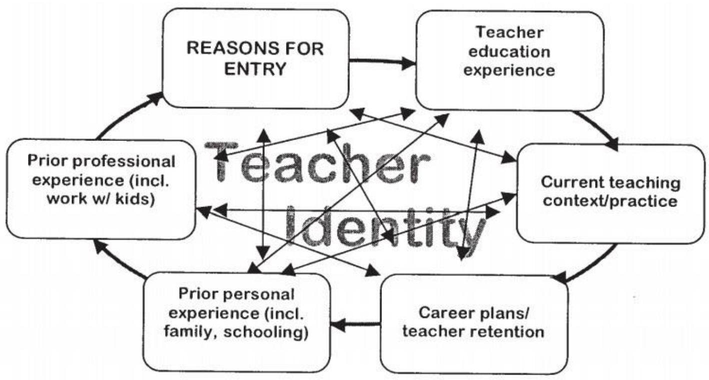


Figure 1 Components of teacher identity and their interactions. Taken from Olsen (2008)

Korthagen (2004, p.85) claims that "fundamental changes in teacher identity do not take place easily: identity change is a difficult and sometimes painful process, and often there seems to be little change at all in how teachers view themselves". Teachers do not have the same attitude towards shaping their identity, as Livingston (2016, p.401) emphasizes "some teachers feel comfortable about re-forming their identity as a teacher and others find it a very difficult process". This highlights the link between emotions and identity. Roth (2007) suggests emotions and motives are central to understanding the formation of identity. Teachers' emotions can become the source of resistance to identity change (Zembylas, 2003). As a teacher's identity changes, a teacher's confidence may also be shaken, potentially leading to feelings of vulnerability, confusion or even incompetence (Heidemann *et al*., 2016). The next section focuses on how teachers overcome the obstacles in the development of identity.

Previous Studies

A spate of studies have been conducted in relation to formation and development of teacher identity. Various aspects have been investigated, including what influences the (re-)construction of teacher identity in terms of the factors affecting identity (Duff & Uchida, 1997), identity negotiation (Arvaja, 2016; Barkhuizen, 2009), the relationship between identity and context (Botha & Onwu, 2010), the effects of a teacher training course on teachers' professional identity (Abednia, 2012), the relationship between identity and teaching experience (Lee, 2013), the role of mentoring in identity (re-)construction (Izadinia, 2015), and the identity development in pre-service teachers and teaching practicum (Ivanova & Skara-Mincane, 2016; Ruohotie-Lyhty & Moate, 2016). These studies report a number of factors that influence identity construction, reconstruction and development. Furthermore, the studies were conducted in a wide variety of contexts, highlighting how identity formation varies according to environment.

Personal background and contextual factors are key elements in shaping teacher identity. In a reflective journal and questionnaire–based study, Duff and Uchida (1997) explore how teachers' sociocultural identities, understandings, and practices are inter-connected and change over time, and what factors are associated with those changes. The study involved four teachers at a language center in Minato City in Japan. It was found that the teachers' perceptions of their sociocultural identities were based on their personal background, including their educational, professional, and (cross-) cultural experiences. Inevitably, they reshape their identity due to the dynamic contextual elements, such as the classroom/institutional environment, instructional materials and reactions from students and colleagues. In another study, Botha and Bonwu (2010) focused primarily on how first year teachers form their identities in different school contexts. They interviewed and observed two beginner primary level teachers in early mathematics science and technology teaching in South Africa, and underscored "the mediating influence of the personal and professional sides and teachers' work environment in shaping the professional identities" of these teachers (Botha & Bonwu, 2010, p.113).

Negotiating the current and future roles also plays a major role in identity formation. Barkhuizen (2009) highlights the importance of providing spaces for teachers to examine their imagined roles within a community of peers and shows that holding those roles up to the mirror of current theory and practice allows them to evaluate and rethink their ideas about teaching, and their imagined roles. He investigated a pre-service English teacher's emerging teacher identity by eliciting interview data and reflective journals to find the answer to the question of "*Who am I*?" He employs a positioning analysis of a small story in which a migrant pre-service teacher imagines a better future life. In his case study, the Tongan pre-service teacher first described the characters and the events (level 1). Then, she constructed the story interactively (Level 2). Finally, she positioned herself in the story and determined her possible identities. Arvaja (2016) also focused on identity negotiation; and he examined how a teacher constructs his identity through the process of positioning in the context of a one-year programme at a university in Finland. Using narratives and interviews as the major data collection tools, the study demonstrated that teacher identity construction is the result of various positions that the teacher negotiates with others.

The relationship between teacher identity and teaching experience is another focal point for many studies. Lee (2013) explored how EFL teachers construct and negotiate their identities in their work context as they learn to become teachers of writing. He worked with 4 EFL teachers in a 20-hour part-time in-service writing teacher education course offered by a Hong Kong university. The interviews with the teachers demonstrated that identity is not a static but a changing concept. He also claims that "teachers' identities are discursively constructed (i.e. identity-in-discourse), and reflected in the ways they talk about their work (i.e. narrated identity-in-practice)".

Other studies have focused on the effects of a teacher training course on teachers' professional identity. Gathering data via reflection and interviews, Abednia (2012) showed in his research that even a short teacher training course can influence the professional identity of teachers. This shows that changes in teacher identity can be triggered by interaction with tutors and peers in the immediate environment.

The role of mentoring relationships in creating changes in the professional identity is another focus in teacher identity-related studies. Izadinia (2015) worked with seven preservice teachers in the Graduate Diploma of Education-Secondary (GDE-S) Course, in the School of Education at a university in Western Australia. Her study involving classroom observations revealed that interactions are of paramount importance in identity formation. She also noted that when there was no problem in mentoring relationships and expectations were fulfilled, pre-service teachers gained confidence, but for teachers "who experienced a partially negative mentoring relationship, their confidence declined and they felt they did not improve" (Izadinia, 2015, p.1).

The relationship between teaching practicum and identity development has attracted many researchers' attention. Ivanova and Skara-Mincane (2016) investigated the identity development in 66 prospective teachers at the University of Latvia, Faculties of Education, Psychology and Art. Their questionnaire-based study concluded that how teachers' professional identity greatly depended on their previous professional and personal experiences. Student teachers do not regard themselves as teachers until they actually start to teach. Similarly, Ruohotie-Lyhty and Moate (2016) focused on how pre-service teachers develop their professional identities and what kinds of experiences they draw on in this process. Their study involving classroom observations of 12 participants in the Department of Teacher Education in Finland revealed that "student teachers are active within their own development and draw on experiences as they exert their identity-agency" (p.318). Identity-agency is linked to the teaching context, which provides feedback in the form of changes in the student teacher and the others.

This case study is different in that it is a longitudinal investigation of identity development and it includes a combination of methods and triangulated data collection tools. The study follows a process-oriented data collection methodology to capture the intricacies during the teachers learning and teaching experiences. The identity development of the novice teacher was monitored not only in a training course but also in a professional setting throughout a six-month period with a special emphasis on the developments not only in beliefs, but also in knowledge and practices. As an alternative to reflective journal and questionnaire-based, interview-based studies, common data collection tool in identity studies, we used observation (classroom and video-enhanced) of teacher practices combined with narratives and written reflective interview in the current study in order to gain a deeper understanding of the process of identity development.

Research Questions

1. What are the stages Derviş has gone through in developing his teacher identities?
2. What factors have played a role in Derviş' identity development?

Material and Methods

Context of study

This case study involves two different settings Derviş experienced teacher learning and teaching. The first setting was a full-time four-week teacher training course run in July 2016. In the programme, the participants receive a 120-hour contact with their tutors and peers. They receive 35 input sessions on a wide variety of topics including language learners, authentic materials, and lesson objectives. They are required to hand in 4 written assignments and successfully deliver 8 lessons at two distinct levels. The second setting, from September, 2016 to January, 2017 was a private university in Izmir, Turkey. This institution has a well-established Teacher Development Unit, which provides sustained support for its instructors through regular professional development activities. Derviş taught general English in A2 level to a group of 22 young adults 20 hours per week in the university preparatory programme. During this time, he attended three workshops given by the Teacher Development Unit and held regular meetings every two weeks with his training course tutor who was also a teacher trainer and instructor at the university.

Participant

The single participant of this study, Derviş, had a BA in English language teaching and 9 years of teaching experience before the teacher training course. He had worked in different language schools, mainly delivering grammar lessons, as he reported to us. He had also given private lessons for a few years. On successful completion of the teacher training course, he was recruited as a part-time instructor at the university. The first author was one of the tutors for Derviş during his short course.

Data Collection Tools and Procedures

To draw a full picture of the identity development of Derviş, over a period of six months, we collected a triangulated set of data including recordings of interviews, reflective writing, and observation, in order to increase the credibility of the results obtained. The prolonged engagement with his experience of teaching also allowed us to build evidence of change and development. The data collection procedures were as follows:

Tool for Examining Knowledge

Knowledge test: To examine the knowledge development in Derviş, we developed a professional knowledge test based on the syllabus of the training course and the priorities given in the private university. The questions involved the basic terminology, concepts and practices related to English language teaching such as ICQs, peer-checking, PPP and jig-saw reading in line with the certified content of an international teacher training course. The test was given to Derviş three times in order to record how cognizant he was of the theoretical knowledge of the basics of ELT at the beginning of the training course, and whether he had improved in the middle and at the end of the study period.

Tools for Examining Beliefs

Written interview: Derviş was also asked to give three written responses related to his teaching philosophy and beliefs, at the beginning, in the middle and at the end of the study. The questions were designed to display his perception of teaching and learning, and the roles of the teacher and students.

Narratives: Derviş took notes of his important experiences and opinions in the format of narration throughout the study, and shared them with his mentor at the middle and at the end of the study period.

Tools for Examining Classroom Practices

Classroom observations and video-enhanced observations: To corroborate the ideas in the written interviews, we decided to observe his lessons with specific focus on his areas of concern and the activities he implemented. The aim was to provide evidence of change and development in his teacher identity. To investigate the practices that Derviş conducted, we observed three live lessons (Observations 1, 3, and 5) and three videoed lessons (Observations 2, 4, and 6). The points noted in these classroom practices concerned four issues: the types questions; activities; teacher and student talking time; and interaction patterns and the teacher's stance. These common points were determined based on the criteria of the teacher training course and the Teacher Development Unit Teacher Observation Checklist. After each lesson observation, Derviş received personalized feedback from his mentor.

Data Analysis

The collected data were subjected to content analysis (Krippendorf, 2004) by both authors to ensure the compatibility of the categorization of the themes that emerged. This resulted in a high level of intercoder reliability.

Knowledge test: The tests were assessed by two different raters each time. The answers given in the different tests were compared in order to notice any changes and progress Derviş had made up.

Written interview: The written interviews were first read by both authors and the meaning units were coded in a way that would represent Derviş' teaching beliefs. This initial analysis of the pre-training data revealed a recurrent philosophy of teaching, which was in line with behaviouristic principles. The post-training interview was analysed by the two authors who selected extracts that underlined the developments in his philosophy of teaching. The same process was followed for the written interview collected at the end of the six months of monitoring. The emerging categories were immediately cross-compared with the earlier set to identify the developmental patterns, which had the potential to reveal a broader insight into identity development.

Observations: The observations were carried out both by classroom visits and video-enhanced observations. The lessons were carefully watched and the units for each category were noted and counted. Any changes in the number of units at each stage would provide evidence for the cognitive and practical changes, which could also be evidence for his developing identity. In each observation, the focal points were the same four issues: the types of questions, activities, teacher and student talking time, interaction patterns and the teacher's stance. We categorised the questions in each lesson according to Bloom's taxonomy (Bloom, *et al.*, 1956), in order to determine whether or not the questions triggered the higher order thinking skills. The activities were evaluated as to whether they required meaningful communication, or simply involved controlled mechanical exercises. The duration of teacher and student talking time was recorded to reveal the extent to which the lesson was teacher dominated. The parts of the classroom used in teaching was another point of focus, as his stance in the lessons revealed his attitude towards monitoring, participation in activities and mediation.

Narratives: The reflective writings that Derviş created throughout the 6 months were also analyzed by using content analysis. The two authors selected extracts that would either support or contradict the emerging themes in the interview data.The narratives were unstructured; and there was no deadline, so Derviş was given the freedom to describe the important points in his teaching experiences and the changes in his identity. The quotes selected through analysis were then brought together to provide evidence of identity development. The comments related to the changes in his teacher identity and teaching philosophy are reproduced in the relevant parts of this study.

Trustworthiness

The first and foremost priority was given to the ethical dimensions of this study. We informed Derviş about the study before we began to collect the data. He was willing to participate as this process, he said, would also offer him a chance to reflect on his own learning and development process. The participant gave us consent to using his real name.

This qualitative research aiming to explore the identity development of a novice English language teacher was carried out as a single case study (Kazdin, 1982). The study followed a triangulation model, bringing together data from a longitudinal perspective, involving 6 months of monitoring Derviş' knowledge, beliefs and practices. The identity developed over the six months was documented and analysed in relation to the context he was supported with a one-month initial teacher education and a 5-month period of teacher development support. To gain a detailed insight into the identity development of Derviş, we collected a triangulated set of data including interview, reflective writing, and observation, with the aim of increasing the credibility of the results obtained. The prolonged engagement in with the participant's experience of teaching also allowed us to strengthen
our evidence of change and development.

Results

In this section, we present qualitative findings to address the research questions. First, we examine the stages in the development of Derviş' teacher identity. Throughout the study, Derviş was through 3 phases: pre-existing identity, developing identity and liberating identity, which correspond to behaviourist, cognitivist and socio-constructivist theories respectively. Each phase was evaluated using a set of subheadings.

Research Question 1: What are the stages Derviş has gone through in developing his teacher identities?

Derviş' pre-existing identity

Professional knowledge

Derviş was a given a set of 10 questions on the basics of ELT on the first day of the training course, July 09, Sunday, 2016. He could not answer four; and his response to the last question was completely irrelevant. The rest were correct but quite limited. Derviş did not seem knowledgeable on the first day of his initial training.

Teaching beliefs

On the first day of the training course, Derviş was also given a set of four questions to reveal his beliefs about teaching and learning process and the roles of teachers and learners. Below are the extracts from his responses at the beginning of the training course:

> "Teaching for me can be described as an art of transferring one's knowledge to another. To put it simply, teaching is transmitting the information to the students."
>
> "Learning is a process where you wait for knowledge from someone."
>
> "Actually, there is no specific quality a teacher should have. Teachers should support their students positively and they should make their students achieve their goals. To attract the students' attention, a teacher should talk and act like a magnet. They have to be patient and a good model for their students."
>
> "Analysing language and focusing on grammar are the backbone of language teaching. I feel satisfied when my students express that they have understood the target structure. Most of the activities are for my students to use the language accurately."

Teaching practice

Questions and Activities: In Observation 1, Derviş asked 51 questions, 24 of which were Y/N questions; and the rest were repetitive ones and they required short responses. The students asked each other 10 questions after being nominated to ask those specific questions. (Teacher: Bermal read the first question and pick one of your friends). Therefore, there was no freedom to choose which question to ask or when to ask them. The communication was also initiated and directed by the teacher. He started his lesson with a power point presentation with 13 slides and showed 5 pictures on the board. Apart from the controlled practice, 'forming questions' section in the course book, the rest of the lesson was simply asking and responding to questions. The students were not provided with a task requiring group work in order to find a solution or conclusion. They did not support each other or have to listen to each other, as there was no real cause, which demonstrated the lesson was not satisfactory in terms of social interactions.

Also, in Observation 2, the students did not have a second exchange i.e. an opportunity to transform the question-answer exercises into a dialogue. As there was no second turn, it was difficult to check whether all students heard the questions and answers. There was no reason for most of the students to listen to the speakers, who were nominated by the teacher or their friends. Communication was limited to the sentence-level.

Interaction Patterns: In Observations 1 and 2, the type of the questions, who asked the questions, and the lack of exchanges demonstrated the interaction patterns in the lesson. Throughout the lesson, all talk was teacher-initiated. In the first two observations, the interaction patterns were usually limited to T→S and T→Ss. S→T could be observed only when the students responded to questions.

Teacher Talking Time vs. Student Talking Time: Although Observation 1 was planned as a speaking lesson, Derviş dominated the talking time. He asked the questions, parroted responses, focused on the grammar points, talked about his hobby and his success in logo creating, all of which minimized the student talking time. Derviş acted "like a magnet", and always in the spotlight in his lesson. Similarly, in Observation 2, the activities and instructions reflected that the teacher's main priority was the language itself, not communication.

Teacher's Stance: In Observations 1 and 2, the teacher was near the board or his desk for 35 of 45 minutes of the lesson. When he was asking questions or listening to the students, he wandered aimlessly in front of the board. He spent only 6 or 7 minutes close to the learners, checking whether they could form the questions accurately.

Derviş' developing identity

Professional knowledge

On August 03, Sunday, 2016, in the last week of the training course, Derviş answered the same 10 questions related to the basics of ELT that he received on the first day. He answered all concisely and correctly, reflecting his development in the theoretical aspect of ELT. Such an improvement occurred due to the reading and writing assignments, and the sessions and workshops he attended during the course.

Teaching beliefs

Towards the end of the training course, there was a dramatic change in his teaching philosophy and beliefs. He explicitly stated that "in the course, I experienced a clash between the teacher's role I had already adopted and the one that is required." Before elaborating on the difference, it would be appropriate to point out the reasons for such a change:

> "The ages of the learners in the training course (24 - 57) and the students whose mother was not Turkish urged me to make a change in the way I taught English."

> "I believe that rather than presenting the target structure, the students need to discover the rules by themselves."

> "Observing my peers and experienced teachers, the feedback I received from my tutors and the mistakes that I made in my teaching practices helped me self-question myself and reshape 'the teacher' in my mind."

> "I used to think that students should work with their teacher only, not with their peers since it is the teacher who can make instant corrections and help them use the target language accurately."

Teaching practice

Questions and Activities: In Observation 3, the main purpose for the questions Derviş asked was to encourage the students to make inferences and justify their responses. The students were shown a picture from Paris and they were asked where the picture was taken and how they knew. They also looked at pictures of people; and the teacher asked the students to make inferences about their nationalities. After setting the context, Derviş gave the students a sample email and asked them to underline the important features of an email, employing an inductive rather than a deductive method to teach email conventions. However, this lasted so long that there was not enough time for the students to write and exchange emails. Derviş had similar timing problems in Observation 4. Focusing on the grammar point of "Have you ever…?", he asked the students the following concept checking questions:

- Is it past, present or both?
- Do you know when the action happened?
- Is it still possible to….?

The students listened to the song including the target structure, underlined the verbs and tried to understand the meaning of the present perfect tense in pairs. Unfortunately, the length of the guided discovery part and the following controlled practices left no time for freer oral or written production for meaningful language use.

Interaction Patterns: In Observations 3 and 4, although the students were encouraged to work in pairs and groups, the teacher still played a key role in the interactions, initiating most of the talk. He was the one with the answers (accuracy of the language and conventions of an email), which, in fact, maintained student dependency on the teacher. The main reason was that the focus of the pair and group work was language itself, rather than meaningful communication.

Teacher Talking Time vs. Student Talking Time: In Observations 3 and 4, the TTT was far less than previous observations, as the teacher asked the students to make predictions and give justifications for inferences in pairs and groups. However, the students were forced to use their mother tongue as no task language was provided.

Teacher's Stance: Derviş' position in the first halves of Observations 3 and 4 was similar to his stance in the first two observations, mainly around the board for whole class work, but in the second half of the lessons, he moved around the groups to check the students' inferences and written work.

Derviş' liberating identity

Professional Knowledge

Derviş seemed to have greatly improved his theoretical knowledge due to his interest in teaching and his experience, which is reflected in his answers to the questions on December 19, 2016, almost 6 months after the beginning of this study.

Teaching beliefs

When the training course was over, Derviş started to work at the preparatory programme of a private university and had the opportunity to apply all the theoretical and practical knowledge gained in the training course. This represented another transformation process. Before describing the process, it would be helpful to identify the causes of the transformation.

> "I noticed that though the teaching and learning environment was genuine in the course, being a full-time teacher that gives 20 hours a week is quite different from being a trainee who delivers a total of 8 lessons in a month"

> "Even in the training course, I realized that I had never spent time on the production part of a lesson. All my lessons were about presenting the language and related controlled practice"

> "My priority is now creating a learning environment in which my students can communicate with each other. I have turned into a participant in their interaction (not an instant feedback giver on their language use). I have developed my listening skills"

Teaching Practice

Questions and Activities: In Observation 5, Derviş was not the only source of information in the lesson; in fact, he was not the source or in the spotlight at all, and not the source of the majority of questions. The students worked in groups. They first had their group debate and then open-class debate. Meanwhile, Derviş monitored, joined group debates and moderated the class debate. Despite a more passive role, students were aware that the teacher was listening to their questions and responding to their contributions to the debates. Most questions that Derviş asked students required higher order thinking skills:

- Why is the café uncomfortable? (Analyze)
- How can you make the café a good one? (Synthesize)
- How would you feel in such a café? (Evaluate)
- Do you think changing the colours of the walls important? Why? (Evaluate)
- Can you justify the changes you recommend? (Evaluate)

In Observation 5, Derviş showed two pictures (holiday and café) related to the text read in the previous lesson, and asked students to write as many as possible related words in pairs in one minute. After eliciting the words, he formed groups of 4 (SA, SB, SC, SD) and asked the students to write why the café in the holiday resort was unsatisfactory. The students were also required to justify their answers. Then, he re-grouped the students (SAs in group 1, SBs in group 2, SCs in group 3, SDs in group 4). The task was to have a debate on the factors that made the café uncomfortable and how to improve it. After the group work, the teacher turned the activity into a whole class debate.

In Observation 6, the students first analysed some advertisements in their course book and expressed their opinions. The main features of advertisements were highlighted and then students worked in groups to create their own advertisement. In Observation 6, Derviş asked questions similar to those in Observation 5:

- Why do women wear make-up? (Analyse)
- Which advertisement affected you most? Why? (Analyse)
- How can you make an affective advertisement? (Evaluate)

Interaction Patterns: In Observation 5, the wide variety of interaction patterns is clearly noticeable. For about four minutes of the lesson, the teacher and the students talked to each other (T→Ss, S→T); however, throughout the rest of the lesson, the students spoke and listened to each other. (S→Ss, Ss→S, Ss→Ss). Similar interaction patterns were observed in Observation 6. The students were asked to work in groups to create an effective advertisement. Another important point about their interaction was that the purpose for their talk was not on the accuracy of the language usage only, their focus was on the meaning conveyed, which made their interaction meaningful and real-life.

Teacher Talking Time vs. Student Talking Time: In Observations 5 and 6, the high ratio of STT showed student-centred lesson. Apart from grouping the students and setting the tasks, the teacher did not dominate the lesson. His talk to the students during group work was participating rather than directive.

Teacher's Stance: In Observations 5 and 6, as the nature of the group work activities required, the teacher closely monitored the students and walked around the groups throughout the lesson. There was no dominant location, but it can be clearly seen that he did not anchor himself around the board as in earlier lessons.

Research Question 2: What factors have played a role in Derviş' identity development?

Once a considerable shift in his identity had been noted, we asked Derviş to describe the factors that supported his development in the context of initial teacher training and teaching. As key factors that facilitated his identity development, he highlighted the training course, stating *"The training course was the turning point in my professional life. This caused drastic changes in my teaching beliefs and classroom practices"*. He also stressed the critical role of observation and feedback by saying *"Observations and feedback sessions created the backbone of this new Derviş"*. He strongly emphasized his transformation as a new teacher. Derviş also stressed the importance of context and professional development policies followed in the institution. He clearly stated *"If I did not work here, I would not be the teacher that I am now"*, which once again indicates that the identity shift was enabled by the contextual facilities. The close mentor support was evidently a factor in his developing identities. He explained this as *"The more I teach, the better I have become. The feedback I have received from my learners and my self-reflection helped me reshape myself as a teacher"*. Derviş also highlighted the importance of a context-specific facility, the systematic video recordings:

> "At university, we have special Panopto cameras which I can observe myself or the lesson itself right after the lessons. If you ask me, I like them as they act like observers. Thanks to them, I could watch my lessons and focus on specific parts in the lesson and I could take some notes about the lesson. They are also helpful for self-reflection and seeing what you have done correctly or wrongly actually play a crucial part in your development as a teacher" (Derviş)

Derviş viewed the recorded lessons as an opportunity that helped him develop his understandings, and in turn, his teacher identity. In addition, he highlighted the role of the participant-sensitive mentoring provided by in-house trainers. He explicitly said *"My tutor always helped me to improve myself. Thanks to the regular meetings we had, I could evaluate and renew myself"*. Derviş was aware of facilitative factors that led him to develop new identities.

Discussion

The research first investigated, through qualitative data tools the stages Derviş has gone through in developing his teacher identities? From a more general perspective, the analysis of the data revealed three major stages including pre-existing, developing, and liberating identities as Derviş engaged in learning to teach and teaching in actual classrooms while simultaneously supported by systematic professional development activities. The initial characteristics of his identities showed that he was reporting traditional practices and behaving in the classroom as a transmitter and authority in front of the classroom with limited knowledge of what he knows about teaching English.

However, later he developed his identity due particularly to professional support provided simultaneously for learning teaching. We found that classroom practice played a role in the identity formation and development, a relationship also argued by Kanno and Stuart (2011, p.250). Similarly, Lee (2013, p.332) relates classroom practices to developing identities and the emerging identities to the altering classroom practice. This was an emerging relationship in our research as Derviş revisited and rethought his teacher identity during the active engagement in teaching. However, only classroom practices may not lead to new and developing identities to emerge. Rather, there is need for professional support, which was often provided after the classroom observations, where the first author, had the opportunity to collect evidence of the identity construction and development. The observations we conducted in his classroom functioned as critical means for identifying the developing aspects including his identity (O'Leary, 2014). Some lessons were videotaped for further observational analysis, which also helped him reconstruct a sense of professional identity and positive self-image (Van Es, 2011, p.136), the benefits are also reported by Scida and Firdyiwek (2011). Derviş constructed and developed his beliefs based on these observations as well as on the reflections conducted with him after each. Such post-observation dialogues create opportunities to rethink and self-assess classroom practices without forgetting the pedagogical nuances and verbalize the procedures to the mentor who could further deepen the meanings that are constructed. Walsh and Mann (2015, p.356) argue that reflection does not have to take place "in isolation, but in discussion with another practitioner" as this dialogic professional talk paves the way to "cooperative development". How reflection could shape identity development is also argued by Beauchamp and Thomas (2009, p.182) who conceptualized reflection "as a key means by which teachers can become more in tune with their sense of self and with a deep understanding of how this self fits into a larger context which involves others; in other words, reflection is a factor in the shaping of identity". In this research, reflection on the observed and videotaped lessons played a key role in shaping the identity of Derviş. Towards the end of the study, the data Derviş provided regarding his beliefs and practices emerged as strong indicatives of his identity change as a teacher. He was more liberated in the sense that he started to ask more critical questions that were thought provoking as well as diversified the interaction patterns he used. He also reduced his own talking time and offered more opportunities for students to talk, and he was observed to be more around the students, which indicated to us that he was taking on a liberating identity by freeing himself for more dialogues, sharing, and co-learning with his students.

The second question investigated the factors which have played a role in Derviş' identity development. We found that identity construction and development is more effective when the teacher works in a supportive environment, one of the contextual factors which plays a major role in the reconstruction (Duff & Uchida, 1997) of Derviş' identity. The role of mentoring in identity (re) construction (Izadinia, 2015) in the workplace acted as a catalyst in this case in that Derviş was provided with sustained mentoring by experienced mentors during the training course and his part-time work. The effects of a teacher training course combined with the follow- up part time teaching post influenced the teachers' professional identity (Abednia, 2012). On the other hand, the series of classroom and video-enhanced observations offered Derviş opportunities to receive constructive feedback, which helped shape the identities Derviş brings to the classroom and his real experience of teaching (Sumara & Luce-Kapler, 1996, p.81). The observations and individual feedback sessions are highlighted to play a critical role in shaping teaching practices (Ivanova & Skara-Mincane, 2016; Ruohotie-Lyhty & Moate, 2016). Kanno and Stuart (2011, p.240) argue that there is a "mutually constitutive relationship between identity and practice". They change almost simultaneously (Lee, 2013), which is also evident is also evident in this study in that Derviş performed teaching in different ways as shown in the analysis of the observations.

Another critical finding was that we gathered and analysed ample evidence regarding how Derviş developed new identities by developing more constructivist ways of teaching and taking on different roles. Although Korthagen (2004) argues that significant changes in teacher identity are challenging and can involve a distressing transformation, Derviş underwent distinct phases within a relatively short time, and never described the process as uncomfortable. Livingston (2016) also states that not all teachers have the same attitude towards developing identity. This is directly linked to the degree of openness or resistance to identity change (Zembylas, 2003). Derviş was a volunteer in the study, and intrinsically motivated to improve himself. The identity transformation clearly reflects such an attitude. This improvement also reflected itself in Derviş' professional knowledge. He deepened his pedagogical and content knowledge in teaching. This enabled him to hone his instructional skill of what to teach when and how, which is directly linked to identity development because "having sufficient competence in basic teaching skills so as not to make obvious blunders in front of students is a necessary condition for one's claim as a legitimate teacher" (Kanno & Stuart, 2011, p.245).

Conclusions

In the study, we identified three different approaches that Derviş implicitly held in succession. The one reflected in the pre-existing stage coincided with several principals underlying behaviourism, and although unaware of this, he shows evidence for this in his teaching beliefs and classroom practices. Downgrading the teacher's identity to the "information transmitter" and "role model" reveals that he used to emphasize attention to the accurate use of the language (Jimerson, Cho, & Şimşek, 2014; Wayman, 2016). Thus, he stated several times in the verbal and written interviews that a teacher should be "patient" enough to correct the students' mistakes. Specifying a learner as someone "waiting for knowledge" implies that he considers students as passive recipients, and a preference for deductive teaching. A main focus on the analysis of the language implies that he did not regard meaningful communication as a priority for language learning. The practices he initially conducted included controlled mechanical drills, and yes/no questions that were not aimed at promoting thinking or creating free responses. In the developing identity stage, Derviş seemed to abandon his behaviourist identity and embarked on a more cognitivist approach, giving importance to 'discovery learning' and 'exploring the rules using pair work'. He noticed the benefits of inductive methods of presenting target language. By this stage, Derviş no longer regarded himself as the only source of knowledge, and his role was not limited to the feedback giver on the accuracy of student language. In the pre-existing identity stage, Derviş was unaware of his previous teaching beliefs; however, at this stage, he could reflect on himself and increase awareness of his own teacher identity. This also provides evidence that he developed his metacognitive skills. In the liberating identity stage, Derviş started to adopt the characteristics of a socio-constructivist approach. He underscored the importance of student-to-student communication for learning; and rather than focusing on his speaking, he aimed to improve his listening skills as thus far, he had habitually initiated and dominated classroom talk. This change reflected itself in his classroom practices as well, since in his last observations, his learners were able to fully participate in group and class debates, and to create their own advertisements. We believe that the institution also played a key role in the identity development and change Derviş experienced, providing a wide variety of professional development activities, including classroom observations followed by one-on-one feedback sessions as well as participant-sensitive mentoring, which supports individual novice teachers.

Since language teachers in Asian contexts enter the profession with their own beliefs and practices influenced by their own cultural background including the perception of the teacher as a central and omniscient figure, the following practical implications can be drawn. Teacher education programs in Asian context could

- allow opportunities for teachers to discover themselves through self-reflection and mentor supported exploration
- provide ongoing support for development particularly by engaging in tasks that require multiple perspectives
- mediate teacher learning and development process with participant-sensitive mentoring which aims to strengthen the teacher’s mindset from static nature of teacher capacities to a developmental mindset.
- develop ways of identity development over a period of time and monitor identity development as a way of teacher development.

The study has some limitations that should considered for further interpretation. First, we focused on one participant who has unique characteristics as a teacher, though we provided detailed information regarding his background. Secondly, we could not include Derviş’ students’ views about his development, which could have helped us triangulate the data about his teaching and learning as part of identity development. Further research could consider including more participants and students’ views about their teachers development, which could provide identity-related issues regarding how Derviş is seen by others too.

References

Abednia, A. (2012). Teacher’s professional identity: Contributions of a critical EFL teacher education course in Iran. *Teaching and Teacher Education. 28*(5), 706-717.

Akkerman, S. Meijer, P. C. (2011). A Dialogical Approach to Conceptualizing Teacher Identity. *Teaching and Teacher Education 27*, 308-319.

Arvaja, M. (2016). Building Teacher Identity through Positioning. *Teacher and Teacher Education*. *59*, 392-402.

Barkhuizen, G. (2010). An extended positioning analysis of a pre-service teacher’s better life small story. *Applied Linguistics*, *31*(2), 282-300.

Beauchamp, C. Thomas, L. (2009). Understanding Teacher Identity: An Overview of Issues in the Literature and Implications for Teacher Education. *Cambridge Journal of Education. 39*(2), 175-189.

Beijaard, D, Paulien C. Meijer, and Verloop, N. (2004). Reconsidering research on teachers' professional identity. *Teaching and Teacher Education, 20*(29), 107-128.

Bloom, B.S. (Ed.). Engelhart, M.D., Furst, E.J., Hill, W.H., Krathwohl, D.R. (1956). *Taxonomy of Educational Objectives, Handbook I: The Cognitive Domain*. New York: David McKay Co Inc.

Botha, M. & Onwu, G. (2013). "Beginning Teachers' Professional Identity Formation in Early Science Mathematics and Technology Teaching: What Develops?" *Journal of International Cooperation in Education*, *15*(3), 3-19.

Boud, D., R. Keogh, and D. Walker. (1985). What is reflection in learning? In D. Boud, R. Keogh, and D. Walker (eds.). *Reflection: Turning Experience into Learning*. London: Kogan Page Ltd.

Coldron, J., Smith, R. (1999). Active location in teachers' construction of their professional identities. *Journal of Curriculum Studies*, *31*(6), 711 – 726.

Duff, P. A., Uchida, Y. (1997). The Negotiation of Teachers' Sociocultural Identities and Practices in Postsecondary EFL Classrooms. *TESOL Quarterly, 31*(3), 451 – 486.

Heidemann, S., Meninga, B., Chang, C. (2016). *The Thinking Teacher: A Framework for Intentional Teaching in the Early Childhood Classroom.* Goldon Valley, Minnesota: Readleaf Press.

Ivanova, I. Skara-Mincane, R. (2016). Development of Professional Identity during Teacher's Practice. *Social and Behavioural Sciences*. *232*, 529 – 536.

Izadinia, M. (2015). A Closer look at the role of mentor tea hers in shaping pre-service teachers' professional identity. *Teaching and Teacher Education*, *52*, 1-10.

Jimerson, J.B., Cho, V., Wayman, J.C. (2016). Student-involved data use: Teacher practices and considerations for professional learning. *Teaching and Teacher Education 60*, 413-424.

Johnson and Golombek (2016). *Mindful L2 Teacher Education*. New York: Routledge.

Kanno, Y. and Stuart, C. (2011). Learning to become a second language teacher: Identities- in-practice. *The Modern Language Journal*, *95*, 236-252.

Kazdin, Alan (1982). *Single-Case Research Designs*. New York: Oxford University Press.

Krippendorf, K. (2004). *Content analysis: An introduction to its methodology* (2nd ed.). Thousand Oaks, CA: Sage.

Korthagen, F. A. J. (2004). In search of the essence of a good teacher: Towards a more holistic approach in teacher education. *Teaching and Teacher Education, 20*, 77-97.

Lee, I. (2013). Becoming a writing teacher: Using "identity" as an analytic lens to understand EFL writing teachers' development. *Journal of Second Language Writing, 22*, 330-345

Livingston, K. (2016). Developing teachers' and teacher educators' professional identity in changing contexts. *European Journal of Teacher Education, 39*(4), 401-402.

Mayer, D. (1999). *Building teaching identities: implications for pre-service teacher education*, paper presented to the Australian Association for Research in Education, Melbourne.

Ng, P.C.L. and Boucher-Yip, E.F. (2016). *Teacher Agency and Policy Response in English Language Teaching*. New York: Routledge

Olsen, B. (2008). How Reasons for Entry into the Profession Illuminate Teacher Identity Development. *Teacher Education Quarterly, 35*(3), 23-40.

O'Leary, M. (2014). *Classroom Observation: A Guide to the Effective Observation of Teaching and Learning*. New York, NY: Routledge.

Oruç, N. (2013). Early teacher identity development. *Procedia- Social and Behavioral Sciences, 70*, 207-212.

Richards, J. C. (2010). Competence and performance in language teaching. *RELC Journal, 41*(2), 101-122.

Rogers, C. (1951). *Client-centred Therapy*. Boston, MA: Houghton, Mifflin.

Roth, W. M. (2003). Review essay: Culture and identity. *Qualitative Research, 4*(1), 1-34.

Roth, W. M. (2007). Emotion at Work: A contribution to third-generation cultural-historical activity theory. *Mind, Culture, and Activity, 14*, 40-63.

Ruohotie-Lyhty, M., & Moate, J. (2016). Who and how? : Preservice teachers as active agents developing professional identities. *Teaching and Teacher Education, 55*, 318-327.

Sachs, J. (2005). Teacher education and the development of professional identity: Learning to be a teacher. In P. Denicolo & M. Kompf (Eds.), *Connecting policy and practice: Challenges for teaching and learning in schools and universities* (pp. 5-21). Oxford: Routledge. Schempp, P.G.

Scida, E. & Firdyiwek, Y. (2011). Video Reflection in Foreign Language Teacher Development. In Allen W. H. and Maxim H. H (eds). *Educating the Future Foreign Language Professoriate for the 21st Century*. Heinle: Cengage Learning

Sumara, D. J & Luce-Kpler, R. (1996). (Un)Becoming a Teacher: Negotiating Idenitites while Learning to Teach. *Canadian Journal of Education, 21*(1), 65-83.

Şimsek, M.R. (2014). Examining pre-service teachers' conceptions of language and the development of their teaching beliefs over the practicum. *Procedia-Social and Behavioral Sciences*, *152*, 937-942.

Taylor, L. A. (2017). How teachers become teacher researchers: Narrative as a tool for teacher identity construction. *Teaching and Teacher Education*, *61*, 16-25.

Van Es, E. A. (2011). A framework for learning to notice student thinking. In M. Sherin, V. Jacobs, & R. Philipp (Eds.), *Mathematics teacher noticing: Seeing through teachers' eyes* (pp. 134-151). London, UK: Routledge.

Varghese, M. Morgan, B. Johnston, B. Johnson, K. (2005). Theorizing Language Teacher Identity: Three Perspectives and Beyond. *Journal of Language, Identity and Education*, *4*(1), 21- 44.

Vygotsky, L. S. (1978). *Mind in Society*. Cambridge, MA: Harvard University Press.

Walkington, J. (2005). Becoming a teacher: Encouraging development of teacher identity through reflective practice. *Asia-Pacific Journal of Teacher Education*, *33*(1), 53-64.

Walsh, S. & Mann, S. (2015). Doing reflective practice: a data-led way forward. *ELT Journal*, *69*(4), 351-362

Zacharias, N.T. (2010). The Teacher Identity Construction of 12 Asian ES Teachers in TESOL Graduate Programs. *The Journal of Asia TEFL*, *7*(2), 177-197.

Zembylas, M. (2003). Emotions and teacher identity: A post-structural perspective. *Teachers and Teaching: theory and practice*, *9*(3), 213-238.

South Korean Jobseekers' Perceptions and (De)motivation to Study for Standardized English Tests in Neoliberal Corporate Labor Markets

Miso Kim
The Pennsylvania State University, U.S.A.
Duk-In Choi
Chung-Ang University, South Korea
Tae-Young Kim
Chung-Ang University, South Korea

Bioprofile:

Miso Kim is a Ph.D. Candidate in the Dept. of Applied Linguistics at the Pennsylvania State University. Email: miso.kim.09@gmail.com

Duk-In Choi is a Ph.D. Candidate in the Dept. of English Education at Chung-Ang University. Email: education@cau.ac.kr

Tae-Young Kim, Ph.D. is a Professor and CAU Distinguished Scholar (2014-2016) in the Dept. of English Education and Director of Graduate School of Education (English Major) at Chung-Ang University. Email: tykim@cau.ac.kr **(corresponding author)**

Abstract

The study examines 252 South Korean jobseekers' perception of three standardized English tests, their motivation and demotivation to study for the tests, and the relations among their perceptions, (de)motivation, and test scores. Exploratory factor analysis found four perception factors: the tests as a job skill, a gate, a qualification, and a scale of diligence. Instrumental motivation, competitive motivation, and ought-to L2 self were higher than other motivators. Correlation analysis revealed that all demotivation factors negatively correlated with the scores, whereas no motivation factors had statistically significant, positive correlations with the scores. Sequential regression showed that English

The Pennsylvania State University
201 Old Main, University Park, PA 16802, U.S.A.
Chung-Ang University, Heukseok-Ro 84,
Dongjak-Gu, Seoul, 156-756 South Korea

test stress, one of the demotivation factors, was the only significant negative predictor of the scores. The results indicate that the jobseekers perceived the tests more as needed for a job than as a measure of English proficiency, and their motivation did not result in a meaningful increase of the score, while demotivation caused by neoliberal pressure lowered the score.

Keywords: Jobseekers, Standardized English Tests, Neoliberalism, Motivation, Perception, South Korea, TOEIC, TOEIC Speaking, OPIc

Introduction

The global language, English, has been appropriated in different forms to a wide range of local contexts. In Kachru's concentric circles, South Korea (hereafter, Korea) belongs to the expanding circle, where English does not hold official status (Kachru, 1985). Even though it is not used in daily interactions or official settings, English holds strong social power to sort out people who are unqualified for access to valued opportunities. For instance, entrance to privileged colleges requires a high score on English in the College Scholastic Ability Test; college entrance and graduation and even university-funded scholarships require a certain score on standardized tests of English for eligibility. More importantly, the white-collar job market has become almost inaccessible without English, even when the position does not actually involve English (Park, 2009). In other words, English serves as a gatekeeper for decent employment opportunities. Because of its prominent role, the size of the private English education market is approximately 1.8 trillion KRW (1.7 billion USD) (Shim, 2015). Park (2009) calls this phenomenon in Korea *English frenzy*.

In Korean labor markets, English frenzy reveals itself through a form of *spec*, an abbreviation of specifications. Spec is now a widespread neologism referring to one's capabilities for a certain job, just as electronic products list their capacities under specifications. English has constituted the very essence of specs in labor markets since the government-led globalization drive and neoliberal economy reform of the late 1990s. High scores on standardized English tests, such as the Test of English for International Communication (TOEIC), TOEIC Speaking, and Oral Proficiency Interview – computer (OPIc), have since become essential resume-building activities. Cho (2015) called the jobseekers the “spec generation,” overeducated but underemployed young adults in a neoliberal job market.

English is conceived as a commodity yielding economic returns in the linguistic market (Kim, Choi, & Kim, 2018). Although there have been notable studies investigating the macro structure of the linguistic market, only a few studies have dealt with individuals living under the ideology. Young jobseekers in Korea are pressured to procure this symbolic commodity as reified by TOEIC scores for its crucial importance in the competitive labor market. To use the tests in a socially meaningful way, it is fundamental to understand how Korean jobseekers themselves perceive the tests and motivate themselves to prepare for the tests. While the tests are originally developed to measure test takers’ English proficiency, the tests are used for indicating the jobseekers’ spec and stratifying them in the era of *English frenzy*. The “spec generation” jobseekers may perceive the standardized tests as a tool for upward mobility. Depending on how they perceive the tests, their preparation for the tests may vary. Therefore, it is essential to understand how the test is taken up by the jobseekers, so that the stakeholders of the tests (i.e., employers, test developers, and instructors of English) could be informed of social uses of the tests.

This study examines South Korean entry-level jobseekers’ perceptions of the socio-economic importance of the TOEIC and speaking tests (i.e., TOEIC Speaking and OPIc), the motivation and demotivation to study for the tests, and the statistical relations between the perceptions of TOEIC, motivation, demotivation, and TOEIC scores. The research questions are: (1) What factors reflect jobseekers’ perceptions of the socio-economic importance of the TOEIC in the Korean job market? (2) What are the factors of the jobseekers’ motivation and demotivation in studying for the TOEIC for employment? and (3) What are the statistical relations between the factors extracted from (1), (2), and TOEIC scores?

Literature Review

Jobseekers' Perception of English and Neoliberalism in Labor Markets

Language functions as a symbolic commodity within the global labor market, where competence in speaking one language comprises one essential component of human capital (Bernstein, Hellmich, Katznelson, Shin, & Vinall, 2015; Duchêne & Heller, 2012; Holborow, 2015; Park & Wee, 2015; Shin, 2015; Shin & Park, 2015). The commodification of language has been exacerbated by the spread of neoliberalism, an economic paradigm that promotes free markets and free trade, liberates individual freedom, and minimizes regulation (Harvey, 2005). Neoliberalism manifests its impact on language education at multiple levels, including the emergence of a language teaching industry, language learners as consumers of language education, and language teachers as knowledge workers (Bernstein *et al.*, 2015). In this global market, a language is sought for its exchange value – the profitability of the language in the market – rather than its use value.

To account for the commodification of language, scholars have drawn on Bourdieu's (1991) notion of a linguistic market in order to account for the commodification of language in a neoliberal market. Bourdieu theorized that languages, or codes, carry particular symbolic value in relation to the prescribed and legitimized form of variety in the linguistic market. The linguistic competence prescribed by dominant institutions "function[s] in a particular market as linguistic capital capable of imposing the law of price formation which is the most favourable to its products and of procuring the corresponding symbolic profit" (Bourdieu, 1991, p.70). In other words, the legitimized code produced by the dominant class influences the symbolic price of the code in their society. Moreover, "the constitution of a linguistic market creates the conditions for an objective competition in and through which the legitimate competence can function as linguistic capital, producing a profit of distinction" (Bourdieu, 1991, p.55). Competence in speaking the legitimate variety is a symbolic commodity that yields a distinctive value in a linguistic market.

To date, several studies have addressed the commodification of language under the neoliberal economic regime, drawing on Bourdieu's notion of a linguistic market. For instance, Holborow (2015) argued that Bourdieu's framework is apt for describing inequalities and social power embedded in the workings of language, by metaphorizing language as linguistic capital. At the same time, she pinpointed that the framework is limited by explaining the role of human capital on increasing the linguistic capital's value. Park and Wee (2015) analyzed how English is sold for profit and what ideologies are projected onto English in the linguistic market. As English enters into the neoliberal market, it is imagined as a profitable skill that everyone can benefit from. The authors contended that the belief is based on false premises because unequal distribution of power and resource already dominate the competition in the market.

Although they illuminate the ideological commodification of languages, the aforementioned studies largely focus on macro-level discourse, overlooking the individuals living and learning languages under the neoliberal ideology. The studies approach the issue of language commodification and neoliberalism from a top-down perspective, analyzing the issue as a social phenomenon. Although the studies provide an explanatory framework to investigate the phenomenon, they do not provide accounts of the individuals living under the ideology and seeking language for its value as a symbolic commodity. There are only a few studies that report individuals' perceptions of the ideologies and struggles (cf. Block & Gray, 2015; Jang, 2015; Kubota, 2011).

The present study extends the discussion of how individuals perceive the ideology of English, motivate themselves to study languages in the linguistic market, get demotivated while preparing for a job, and how all of these relate to their TOEIC scores. The study mainly discusses the TOEIC test because it is the most commonly referenced test in corporations (Presidential Committee on Younger Generation, 2014). The jobseekers' perception of the socio-economic importance of English is essential for understanding how the ideologies of instrumentalism and neoliberalism that are spread through the dominant discourse affect individuals' study of English. For instance, a well-known Korean news broadcast reported how jobseekers' perceptions of the TOEIC lead to various social phenomena (JTBC News, 9 March 2016). The news said that jobseekers perceive the

TOEIC more as a test of diligence and perseverance than of English proficiency. From their perspective, getting a high score to demonstrate their diligence is more important than improving their proficiency. Paradoxically, this led them to prefer test-taking skills taught in private education, expanding the size of private English education market to 1.8 trillion KRW (1.7 billion USD) (Shim, 2015). The news portrayed extreme cases, such as attending Spartan TOEIC boarding schools and taking the TOEIC in South Asian countries to avoid close proctoring and to better cheat on the test. As shown, the individuals' perception of the TOEIC influences their preparation, the English education industry, and society as well. Therefore, examining jobseekers' perception of the TOEIC will inform other stakeholders how to develop meaningful strategies to teach for the TOEIC test, how to use the test in employment procedures, and how to reform the test. Furthermore, investigating the relation between perception, motivation, demotivation, and TOEIC scores will provide richer data to account for the instrumentalism and neoliberalism of English in a linguistic market.

Frameworks of L2 Motivation for Korean Jobseekers

Since Gardner and Lambert (1959) initiated studies of L2 motivation, it has largely been studied as a learner-internal affective trait for explaining individual differences in language learning achievement. In Gardner's (1985) view, L2 motivation consists of "the combination of effort plus desire to achieve the goal of learning the language plus favourable attitudes toward learning the language" (p. 10). The tradition used attitude as a construct of L2 motivation, since it assumed learners will be more motivated to study an L2 to the extent they are favorable to the L2 community. Following this trend, studies on L2 motivation adopted psychometric and positivistic methods to analyze relations between L2 motivation, attitude, and achievement.

These models of motivation and attitude have been criticized for conflating second and foreign language learning environments and omitting other informative theories of motivation in psychology (Oxford & Shearin, 1994). These criticisms led to a surge of other models. Dörnyei and Ottó (1998) proposed a process model of L2 motivation, which provides motivation a temporal dimension to account for fluctuations of motivation over time. The L2 motivational self system proposed by Dörnyei (2009) expanded the idea of possible selves (Markus & Nurius, 1986) in psychology to L2 motivation. According to

Dörnyei, the ideal L2 self is one's desired future self in the L2 and the ought-to L2 self refers to a set of attributes one must acquire to meet the expectations of others and avoid negative consequences. The will to reduce the gap between the two selves and the current actual L2 self arouses the subject's motivation. More recently, sociocultural and dynamic approaches to L2 motivation have been active areas of discussion. Sociocultural approaches understand L2 motivation in relation to the subject, object, goal, and community (Kim, 2009, 2017). Dynamic approaches reject linear cause-effect relationships and conceptualize L2 motivation as a developing system (Waninge, Dörnyei, & de Bot, 2014).

To examine jobseekers' motivation to prepare for standardized tests under the neoliberal regime and language commodification through statistical analyses, this study uses an exploratory design to examine different motivators derived from different models. This study's research purpose does not align with one specific approach to L2 motivation. Gardner's model may explain their motivation through an instrumental orientation, but it does not fully address the complexity of motivation in social context. Although Gardner (1985) admitted that language is "something more than a neutral linguistic code" (p.7), he attributed attitude to individual sentiment toward the target language community, not to the social context and ideologies influencing the commodification of language (cf. Coetzee-Van Rooy, 2006). A process-oriented model is more apt for explaining long-term motivation than the jobseekers' relatively short-term motivation in studying for the tests. The L2 motivational self system may explain jobseekers' motivations in terms of the desire to reduce the gap between their actual selves and their future images of getting a job by obtaining an appropriate L2 test score. Sociocultural and dynamic approaches may not be appropriate for the current research purpose, because sociocultural approaches are grounded on interpretive and qualitative epistemology and dynamic approaches situate the motivation in time to observe changes in and stabilities of motivation as the system changes (Waninge *et al.*, 2014). Therefore, the present study uses an exploratory design that tests motivators drawn from different models to find which are most salient among Korean jobseekers.

Method

Participants

A total of 252 university juniors and seniors who were on the job market or preparing for the job market at the time of data collection were recruited. Participants' majors were in the humanities and social sciences. Other majors were excluded because the TOEIC is considered more important for humanities and social science majors than others. We limited the pool of participants to jobseekers planning to get a job in the private sector in Korea, excluding those who will take public servant exams, go to graduate schools, or get a job abroad. Although the public servant exams, graduate schools, and jobs abroad require English proficiency, they use different tests (e.g., the Test of English Proficiency developed by Seoul National University [TEPS] or the Test of English as a Foreign Language [TOEFL]) or follow different application-employment procedures.

Data Collection

To recruit participants and collect questionnaires, we personally visited 300- and 400-level courses in the humanities and social sciences. We briefly explained the purpose of the study, asked them to fill out the questionnaire only if they were preparing for a job in the private sector, and assured them that participation in the study was totally voluntary.

The questionnaire consisted of five sections: (1) The profile of the participants; (2) 12 items on their perception of the socio-economic importance of the TOEIC in the job market; (3) 18 items on motivation and 6 items on demotivation to study for the TOEIC for employment; (4) the same items from Section 2 measuring the importance of oral tests (e.g., TOEIC Speaking and OPIc); and (5) the same items from Section 3 measuring motivation and demotivation to study for oral tests. The items in Sections 2 to 5 used a five-point Likert scale (i.e., 1 = strongly disagree, 2 = disagree, 3 = neither agree nor disagree, 4 = agree, and 5 = strongly agree). The first section asked for the general profile of the participants, including years in school, major, gender, months of job preparation, and TOEIC scores. Sections 2 and 4 consisted of items newly developed for this study. Since jobseekers' perceptions of the tests are a new construct, we developed the items by capturing themes emerging from pilot interviews with jobseekers. The items in Sections 3 and 5 were incorporated from previous studies on Korean jobseekers' motivation and demotivation (Hyun & Kim, 2013) and the socio-political analysis of Korean students' motivation (Kim, 2010). We chose to investigate six motivational constructs (intrinsic, integrative, instrumental, competitive, ideal L2 self, and ought-to L2 self) and two demotivational constructs (test stress and compulsoriness of the tests), based on these studies. All items were piloted with 52 jobseekers to ensure validity and reliability, and went through a refining process to clarify the wording and exclude items with low validity or reliability. Because of limited space, Sections 2 and 3 are presented in Appendix.

Data Analysis

The data from the questionnaire were analyzed using the Statistical Package for the Social Sciences (SPSS) 18.0 software. Five main statistical analyses were employed on the data: (1) exploratory factor analysis, (2) descriptive statistics, (3) the Mann-Whitney U test, (4) correlation, and (5) sequential multiple regression. The statistical significance level was set at .05. First, we used exploratory factor analysis to identify constructs reflecting college students' socio-economic ideas influenced by contemporary neoliberalism in Korea when they were preparing for the standardized English test. Maximum likelihood extraction and the direct Oblimin rotation method were adopted. Items with factor loadings lower than .4 were excluded. Second, we used descriptive statistics to calculate the mean and standard deviation of the extracted factors. Third, the Mann-Whitney U test was used to investigate the difference between the jobseekers' perceptions of the TOEIC and speaking tests. Fourth, correlation analysis was performed to examine the relationships between the factors extracted in (1), (de)motivational factors, and the TOEIC score. Last, sequential multiple regression analysis was conducted to examine the relations between all factors and the TOEIC scores. In the fourth and fifth statistical analyses, the factors for speaking tests were not used because jobseekers tend to take only one of the two oral tests, but as the two tests use different grading scales their scores cannot be directly compared.

Results

Jobseekers' Perceptions of the Socio-Economic Importance of the Tests

To extract the factors that reflect young South Korean jobseekers' perceptions of the socio-economic importance of the TOEIC and speaking tests in the job market, factor analysis was performed. Tables 1 and 2 present the results for the TOEIC and the speaking tests, respectively. The Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy is above .7 in both analyses, showing that the data is adequate for factor analysis. Bartlett's test of sphericity also showed that the samples have equal variances. Using a scree plot and setting the eigenvalues at 1, four factors were extracted for both the TOEIC and the speaking test datasets. The four factors accounted for 56.29% and 58.06% of the cumulative variances of the TOEIC and the speaking test datasets, respectively.

Four factors were extracted from the two datasets: (1) a job skill, (2) a gate, (3) a qualification or *spec*, and (4) a scale of diligence. The first factor reflects the jobseekers' belief that studying for the standardized tests will help them perform their job responsibilities with English after they get a job. The second factor, a gate, is a metaphorical label for jobseekers' belief that the tests are only a gate to pass through. The items indicate that they need to score only at a certain level to safely pass the threshold: The applicants do not need to get extremely high scores to pass others, since they will have to compete in other respects after they cross the threshold of the English tests. The third factor, a qualification or *spec*, indicates jobseekers' perception that test scores are a crucial component of the resume distinguishing them from other applicants. In this case, the scores must be as high as possible to outdo competitors and compensate for other weaknesses of their *spec*. The last factor, a scale of diligence, shows that the jobseekers consider the English tests a scale for measuring one's diligence and perseverance in preparing for a job, not an indicator of English proficiency.

Table 3 is a summary of the descriptive statistics and Shapiro-Wilk test results for each perception factor, motivation, and demotivation. For the TOEIC test, the descriptive statistics indicate that jobseekers perceived it most as a gate and least as a job skill. In other words, the jobseekers primarily perceived the TOEIC as an essential prerequisite for job hunting but as less likely to be helpful for performing job duties. On the speaking tests, the factor of a gate ranked lowest, opposite to the TOEIC. Rather, the jobseekers perceive the speaking tests as proving their diligence in preparing for the job and the competitiveness of their *spec* in the job market. That is, the speaking tests are perceived as more valid for demonstrating one's preparedness and competitiveness, while the TOEIC is more of a prerequisite for passing through a gate. Motivators and demotivators are largely similar for both the TOEIC and speaking tests. Instrumental motivation, ought-to L2 self, and competitive motivation are the primary motivators, while intrinsic motivation, integrative motivation, and ideal L2 self remain low.

Because the questionnaire adopted a five-point Likert scale and the sample size was limited, the normality assumption could not be guaranteed. Using the Shapiro-Wilk test, we found that most of the factors do not satisfy the normality assumption for parametric statistical tests. Therefore, non-parametric tests were used for further analyses.

Table 4 presents the results of the Mann-Whitney U test comparing the factors of the TOEIC and the speaking tests. All four perception factors show significant differences. This means that jobseekers perceive the socio-economic importance of the two types of tests differently – the TOEIC test is perceived as more of a gate and a spec, whereas the speaking tests are perceived as more important for performing job duties and demonstrating one's diligence. Among motivators, competitive motivation and ought-to L2 self are the only two factors that do not show a significant difference, which indicates that jobseekers are motivated to study the tests because they need to win the competition for jobs and meet external expectations by acquiring good test scores regardless of test type. Likewise, there is no significant difference in demotivators between the two test types, which suggests that jobseekers feel demotivated to study for tests because of test stress and the compulsory nature of the tests regardless of test type.

Table 1 Factor Analysis of the Jobseekers' Perceptions of the Socio-Economic Importance of the TOEIC

Factor name	Item No.	Items	Factor loading	Eigenvalue	Cronbach's α
A job skill	8	The content of the TOEIC is helpful for getting work done in the workplace.	.878	3.072	.839
	12	Studying for the TOEIC will be useful for performing job responsibilities.	.846		
	4	Studying for the TOEIC hard will help one use English after employment.	.679		
A gate	10	It is okay to have a good enough score on the TOEIC because it is the basic requirement for employment.	.924	2.297	.801
	6	If an applicant achieves a safe score in TOEIC, it is okay to stop studying for it.	.751		
	2	It is okay to have a TOEIC score just higher than the minimum score required for the initial screening process.	.630		
A spec	5	To defeat competitors in the job market, the TOEIC score needs to be as high as possible.	.838	1.842	.777
	9	The higher one's TOEIC score is, the more competitive one becomes in the job market.	.756		
	1	Applicants with low GPAs need to secure higher TOEIC scores.	.607		
A scale of diligence	3	The TOEIC score is an indicator that reflects how diligently one has prepared for a job.	.847	1.180	.630
	7	The TOEIC score is a tool to disqualify those who are not diligent.	.549		

Note. N = 252, KMO = .712, Bartlett p = .000

Table 2 Factor Analysis of the Jobseekers' Perceptions of the Socio-Economic Importance of the TOEIC Speaking and OPIc

Factor name	Item No.	Items	Factor loading	Eigenvalue	Cronbach's α
A job skill	8	The content of the TOEIC Speaking/OPIc is helpful for getting work done in the workplace.	.940	3.179	.867
	12	Studying for the TOEIC Speaking/OPIc is helpful for performing job responsibilities.	.820		
	4	Studying for the TOEIC Speaking/OPIc hard will help one use English after employment.	.678		
A scale of diligence	3	The TOEIC Speaking/OPIc score is an indicator that reflects how diligently one has prepared for a job.	.907	2.781	.793
	7	The TOEIC Speaking/OPIc score is a tool to disqualify those who are not diligent.	.766		
	11	The TOEIC Speaking/OPIc score reflects the effort put into job preparation rather than English proficiency.	.533		
A gate	10	It is okay to have a good enough score on the TOEIC Speaking/OPIc because it is the basic requirement for employment.	.809	1.345	.800
	6	If an applicant achieves a safe score on the TOEIC Speaking/OPIc, it is okay to stop studying for it.	.776		
	2	It is okay to have a TOEIC Speaking/OPIc score just higher than the minimum required for the initial screening process.	.693		
A spec	9	The higher one's TOEIC Speaking/OPIc score is, the more competitive one becomes in the job market.	.792	1.156	.710
	5	To defeat competitors in the job market, the TOEIC Speaking/OPIc score needs to be as high as possible.	.719		

Note. N = 252, KMO = .749, Bartlett p = .000

Table 3 Descriptive Statistics of the Perceived Socio-economic Importance and (De)motivation for the TOEIC and TOEIC Speaking/OPIC

Factor		TOEIC (N = 252)			TOEIC Speaking & OPIC (N = 249)		
		M	SD	Shapiro-Wilk	M	SD	Shapiro-Wilk
Perception of the tests	A job skill	2.23	.88	.000	3.30	.90	.000
	A gate	3.86	.81	.000	3.09	.84	.041
	A spec	3.58	.91	.000	3.56	.80	.000
	A scale of diligence	2.76	.95	.170	3.81	.79	.000
Motivation	Intrinsic motivation	2.43	.96	.000	2.67	.93	.000
	Integrative motivation	1.78	.78	.000	2.50	1.01	.000
	Instrumental motivation	3.68	.80	.000	3.58	.87	.000
	Competitive motivation	3.26	.91	.000	3.37	.87	.000
	Ideal L2 self	2.23	.84	.000	2.58	.88	.000
	Ought-to L2 self	3.29	.92	.000	3.29	.94	.000
Demotivation	Test stress	2.36	1.02	.000	2.54	1.02	.000
	Compulsoriness of the tests	2.84	1.11	.000	2.75	1.03	.000

Note. Min = 1.00, Max = 5.00

Table 4 Mann-Whitney U Test for Comparing the TOEIC and the Speaking Tests

Factor		Mann-Whitney U	Z-value	Significance
Perception of the tests	A job skill	12904	-11.543^{***}	.000
	A gate	24360	-4.503^{***}	.000
	A spec	26874.5	-2.951^{**}	.003
	A scale of diligence	24613	-4.337^{***}	.000
Motivation	Intrinsic motivation	26440	-3.064^{**}	.002
	Integrative motivation	18150	-8.254^{***}	.000
	Instrumental motivation	28109	-2.038^{*}	.042
	Competitive motivation	29438.5	-1.204	.228
	Ideal L2 self	24046.5	-4.556^{***}	.000
	Ought-to L2 self	31138.5	-.147	.883
Demotivation	Test stress	28339	-1.884	.060
	Compulsoriness of the tests	29676.5	-1.054	.292

Note. N = 252, *** means $p<.001$, ** means $p<.01$, * means $p < .05$

The Relationship between Perception, Motivation, Demotivation, and TOEIC Scores

After extracting and examining the factors, we ran correlation and regression analyses to investigate the relations between the factors and the TOEIC scores. As mentioned earlier, speaking tests were excluded because of the discrepancy between the scoring scales of TOEIC Speaking and OPIc.

Table 5 shows the result of Spearman's rho test. The jobseekers' perception of the TOEIC and motivators do not bear any statistically significant relation with the TOEIC score except in the case of ought-to L2 self. In contrast, both demotivators are significantly correlated with the score. The jobseekers who study the TOEIC to fulfill what they are supposed to become tend to have lower scores. In addition, the scores tend to be lower for those who are stressed by the TOEIC and those who feel obliged to study for it. The two findings could be explained by the washback effect of the TOEIC, which discourages the jobseekers by exerting high social pressure. Table 6, the result of sequential regression analysis, shows that TOEIC stress is the only factor that significantly predicts the TOEIC score. The more stress jobseekers feel about the test, the lower the score they are likely to get. None of the perception and motivation factors has any statistical relation with the score in Table 6, indicating that the jobseekers' perceptions of the TOEIC and motivation to study for it may not influence their scores.

Table 5 Spearman's Rho of the Factors and the TOEIC Score

Factor		Spearman's rho
Perception of the tests	A job skill	-.076
	A gate	-.017
	A spec	.037
	A scale of diligence	-.017
Motivation	Intrinsic motivation	.140
	Integrative motivation	.001
	Instrumental motivation	-.111
	Competitive motivation	-.019
	Ideal L2 self	-.032
	Ought-to L2 self	-.178*
Demotivation	Test stress	-.612***
	Compulsoriness of the test	-.377***

Note. N = 178, *** means p<.001, * means p<.05, Dependent variable: the TOEIC score

Table 6 Sequential Regression of the Perception, Motivation, Demotivation Factors and the TOEIC Score

Factor		R	R^2	ΔR^2	t
Perception of the tests	A gate	.047	.002	-.003	.537
	A job skill	.053	.003	-.008	-.515
	A scale of diligence	.070	.005	-.012	.503
	A spec	.074	.005	-.017	.337
Motivation	Intrinsic motivation	.065	.004	-.001	1.439
	Ideal L2 self	.117	.014	.003	-1.468
	Competitive motivation	.130	.017	.000	.506
	Ought-to L2 self	.133	.018	-.004	-.457
	Instrumental motivation	.135	.018	-.010	.272
	Integrative motivation	.135	.018	-.015	.128
Demotivation	Test stress	.360	.130	.125	-4.526^{**}
	Compulsoriness of the test	.363	.132	.122	.675

Note. N = 252, *** means $p<.001$, ** means $p<.01$, * means $p<.05$

Discussion

Washback Effect of the TOEIC Tests in the Jobseekers' Lives

Although standardized English tests are meant to measure test takers' English proficiency, the jobseekers perceive them more of a scale of diligence, a gate or a spec, as in indicated in Table 3. As for TOEIC, a job skill is the lowest (*M* = 2.23) and a gate (*M* = 3.86) and a spec (*M* = 3.58) are the highest. In Table 1, the items grouped under 'the job skill' factor represent the relevance between preparing for TOEIC and performing job responsibilities. Since the TOEIC self-claims that it "has set the standard for assessing English-language listening and reading skills needed in the workplace" (Educational Testing Service, 2016, para. 1), the job skill factor is meant to be the highest, if the jobseekers perceive it as a valid test. However, the analysis yields rather a contradictory result. The jobseekers perceive the TOEIC to be a gate to pass through (*M* = 3.86), a spec that distinguishes them from others (*M* = 3.58), and a scale of diligence which measures their preparedness for a job (*M* = 2.76). It means that jobseekers are pessimistic of the TOEIC's role for measuring business-related English skills.

The skepticism toward the TOEIC needs to be understood as washback effects of the TOEIC in jobseekers' lives. Young Korean adults encounter the TOEIC in every aspect of their lives. Korean universities require TOEIC scores to be eligible for scholarship, to be exempted from the mandatory English curriculum, to apply for extracurricular activities, and to graduate and earn a degree (YBMNET, 2015). After graduation, they face pressure to obtain high scores on the TOEIC, not only to apply for jobs but to outdo their competitors, especially in an unstable and competitive job market. Faced with such pressure, jobseekers opt for a relatively easy shortcut to get the test over with: rote memorization and test-taking skills, a way of learning they feel the most comfortable with (Cho, 2015). Examples of test-taking skills include choosing past participial over present progressive forms, not selecting adjectives if confused, and choosing “I don’t know” option from the multiple-choice questions in the listening comprehension section (Jeong, 17 August 2015). They finish the test using the skills, but not necessarily improve their English proficiency. Therefore, jobseekers’ scores on the tests reflect how hard they prepared for the job, rather than indicating their English proficiency.

No perception factors bear a significant relation with the TOEIC scores in Tables 5 and 6. In other words, the way jobseekers perceive the TOEIC did not influence the scores. They pursue relentless quest for proving the economic value of their unique human capital by submitting objective data on their English proficiency in hopes of getting material benefits in exchange (Bernstein *et al.*, 2015). The way they perceive the TOEIC may affect their study methods but bears no direct relation with scores. It is not a matter of how they perceive the test; it is how they can get a high enough score in the short term to maximize efficiency. This phenomenon is a manifestation of the commodification of English in the neoliberal era.

Validity of the TOEIC and the Oral Tests in the Neoliberal Job Markets

Compared to the TOEIC, the jobseekers perceived the oral tests as more of a job skill (M = 3.30). The results of Mann-Whitney U test in Table 4 supports this claim, showing that they perceive the job skill aspect of the TOEIC and the oral tests differently. It may mean that they perceive the oral tests are better suited to measuring job-related English skills. Nevertheless, the oral tests are still perceived as a scale of diligence the most (M = 3.81), followed by a spec (M = 3.56).

The result indicates that face validity of the oral tests is slightly higher than that of the TOEIC test. It is attributable to the historical trajectory of how the oral tests were utilized in Korean labor markets. Since the LG Group first started to use the TOEIC for employment in 1984 (YBMNET, n.d.), it has become indispensable to any jobseeker in Korea because more than 1,600 organizations, corporations, and institutes use the TOEIC for employment (Presidential Committee on the Younger Generation, 2014). Although its influence was on the rise, the TOEIC was criticized for not reflecting applicants' English proficiency around the mid-2000s, for its incapability of measuring communicative competence (Park, 2011). The TOEIC Speaking test and OPIc identified the niche, providing quantified data on applicants' oral communicate skills. The employers began adopting the tests in addition to the TOEIC. The oral tests do not necessarily encroach the market share of TOEIC; rather, the employers require the TOEIC as a default requirement to apply, and refer to the oral test scores as an additional criterion. What transpires in the case of oral tests is that they do not replace the TOEIC, but that they are an option for demonstrating applicants' skills. As a result, the jobseekers acknowledge that the speaking tests are more valid for assessing job-related skills than the TOEIC, and that they are an additional scale of diligence and a spec, rather than a gate to be passed through.

The social and historical construction of the tests is at the root of the three perception factors – the test as a gate, a spec, and a scale of diligence. They perceive English tests as a gate, especially in the case of the TOEIC, because they cannot even open the gate of employment unless they have a score ready. Since the TOEIC is a more widespread requirement than the oral tests, its role as a gate is perceived more prominently. Both types of tests serve as a spec used to distinguish themselves from competitors and compensate for other deficiencies in their resume. It is also linked with the competitive motivation (Kim, 2010; Kim & Kim, 2016) of winning the competition in the job market by achieving higher scores. Lastly, the tests become a scale of diligence, not necessarily an indicator of English proficiency. This is attributable to their test-oriented ways of preparing for the tests. They perceive that they can prove their diligence and perseverance in preparing for the position by presenting high scores. This perception is more salient in the speaking tests, because they require more effort than the multiple-choice TOEIC test.

Credentialism, Meritocracy, and the Entrepreneurial Self in Motivation to Learn English

In both types of tests, salient motivators are instrumental motivation (M = 3.68 for the TOEIC; M = 3.58 for the oral tests), competitive motivation (M = 3.26; M = 3.37), and ought-to L2 self (M = 3.29; M = 3.29), all of which imply the motive to accomplish goals extrinsic to English learning itself, whereas motivators contingent on learners' own fulfillment, integrative motivation (M = 1.78; M = 2.50), intrinsic motivation (M = 2.43; M = 2.67), and ideal L2 self (M = 2.23; M = 2.58), are less recognized. The jobseekers aspire to gain external benefits, win a competition, and meet the expectations of others by the TOEIC, while neglecting other motivators internal to the learning itself. Correlation and regression in Tables 5 and 6 indicate no motivators had positive relationship with the TOEIC score. Jobseekers with lower scores tend to feel obliged to meet external expectation (i.e., ought-to L2 self), get stressed about TOEIC (i.e., TOEIC stress), and become helpless because of its compulsory nature (i.e., compulsoriness of TOEIC). The more stress they feel, the more they are likely to achieve lower scores. In short, factors extrinsic to learning are more salient and no motivators bear any positive relation with the scores.

First of all, the saliency of the three factors reflects the widespread credentialism and meritocracy of English learning in Korea. Standardized English tests are meant to assess applicants' proficiency and provide valid and reliable data to employers. The original purposes notwithstanding, the jobseekers skirt the tedious procedure for developing genuine proficiency and opt for acquiring test-taking skills, being faced with pressure to defeat competitors in a zero-sum game; employers stratify the candidates based on quantified scores and spec in the recruitment process; the mean score of the tests among jobseekers gets inflated due to competition and easy access to the test-taking knacks (Jeong, 17 August 2015; MBC News, 4 May 2016); and finally, the distrust in the tests increase (Park, 2011).

Second, the fact that no motivators have positive significant relation with the score means that the quest of entrepreneurial self fails to spur jobseekers to achieve higher scores. The neoliberal regime is based on the premise that jobseekers are entrepreneurs of their own fate, thus relentless self-development and self-enterprising are deemed the essential virtues of neoliberal subjects (Holborow, 2015). The motivators in this study embody different types of the jobseekers' aspiration to invest in English to secure a job, ranging from intrinsic fulfillment to external rewards. None of the motivators, however, lead to higher scores, meaning that the entrepreneurial management of one's English skills under neoliberal regime is not a desirable way to stimulate higher achievement. Or, even worse, the pressure and stress from fierce competition may militate against higher achievement.

The credentialism, meritocracy, and the quest of entrepreneurial self sprout from neoliberal ideologies of English learning in job markets. English proficiency not directly measurable or profitable is meaningless in labor markets, as Holborow (2015) noted: "education, now an enabler of human capital development, becomes the crucial driver of the economy, a power-house of economic potentiality from which anything non-functional, not measurable in monetary terms or not immediately economically useful, is expelled" (p.16). It is incumbent upon the jobseekers to have their proficiency measured in scores. They turn into entrepreneurs of themselves, dutifully pass through the gate of English, demonstrate their diligence, and distinguish themselves from others in forms of scores. In other words, the product of their learning is the mighty score. Lave and McDermott (2002) wrote, "[t]he product of learning is *learning* which has been embodied in *a test score or promised credential*, which has become material: it is the *objectification* of *learning*" (p.35, italics in original). The jobseekers objectify their learning results, and submit them into the neoliberal job market, waiting to exchange it with material profit. In the course of such score-building competition, the English tests turn into the ends in themselves and spawn stress and demotivation unconducive to learning.

Conclusion

Overall, the study contributes to scholarship in the L2 motivation field by revealing the washback effect of the tests of English from the viewpoint of the L2 learners, questioning the validity of the tests in the social dimension, and finally providing detailed account on how L2 motivation intersects with credentialism and meritocracy. First, this study provides detailed information on the washback effect on the part of the jobseekers. The ideology of English is effective only if it is propagated and accepted by the public; otherwise, it remains an empty claim. The jobseekers' perception of the tests and efforts to achieve the scores reflect how the tests influenced the jobseekers' English learning activities. Second, the study raises questions on the validity of the tests in the social context. Even though they have high internal validity and reliability, the tests cannot be said to measure English proficiency adequately, if the tests are not used as intended. Lastly, the study shows that the jobseekers' motivation and demotivation are closely intertwined with credentialism, meritocracy, and the quest of entrepreneurial self in neoliberal job markets.

The findings call for a quantitative framework of L2 motivation which is more sensitive to ideology, power, and neoliberalism. Earlier psychometric frameworks, such as Gardner's (1985) socio-educational model, does not fully account for the social aspect of motivation because they attribute motivation to relatively stable individual characteristics. While studies grounded on sociocultural approaches examine power issues in individual learners' lives, they are relatively smaller in scale. Given that ideology and power not only influence individual L2 learners but shape the macro structure of the L2 learning context, it is important to examine L2 motivation in a large-scale study with a framework inclusive of ideology and power.

The scope of the study does not cover other factors that might influence the scores, such as their socio-economic status, study habits, and access to study materials. Another limitation lies in the fact the study only accounts for jobseekers' perspective on the standardized tests, leaving out how employers perceive and use the scores in actual recruitment processes.

The study has implications for various stakeholders of standardized English tests. For human resource managers or employers, the results indicate that they need to devise a more meaningful and genuine means of assessing applicants' English proficiency if their job responsibilities will include English. As for instructors and educators, the results demonstrate the need first to tailor teaching to the development of English proficiency so as to achieve high scores on the tests, rather than teaching to the tests. For test developers, the results indicate the need to complicate the test types to avoid test takers' using answer-choosing strategies and reform the test structure to better serve its original purpose.

Acknowledgement:

We would like to express our sincere gratitude to Dr. Jangho Lee for his insightful comments on the earlier versions of the study.

References

Bernstein, K. A., Hellmich, E. A., Katznelson, N., Shin, J., & Vinall, K. (2015). Introduction to special issue: Critical perspectives on neoliberalism in second/foreign language education. *L2 Journal*, *7*(3), 3–14.

Block, D., & Gray, J. (2015). “Just go away and do it and you get marks”: The degradation of language teaching in neoliberal times. *Journal of Multilingual and Multicultural Development*, *37*(5), 481–494.

Bourdieu, P. (1991). *Language and symbolic power*. Cambridge, MA: Harvard University Press.

Cho, H. (2015). The Spec generation who can’t say “No”: Overeducated and underemployed youth in contemporary South Korea. *Positions*, *23*(3), 437–462.

Coetzee-Van Rooy, S. (2006). Integrativeness: Untenable for world Englishes learners? *World Englishes, 25*(3/4), 437-450.

Dörnyei, Z. (2009). The L2 motivational self system. In Z. Dörnyei & E. Ushioda (Eds.), *Motivation, language identity and the L2 self* (pp. 9–42). Bristol, UK: Multilingual Matters.

Dörnyei, Z., & Ottó, I. (1998). Motivation in action: A process model of L2 motivation. *Working Papers in Applied Linguistics, 4*, 43–69.

Duchêne, A., & Heller, M. (Eds.). (2012). *Language in late capitalism: Pride and profit*. New York: Routledge.

Educational Testing Service. (2016). About the test. Retrieved from https://www.ets.org/toeic/listening_reading/about

Gardner, R. C. (1985). *Individual differences in second language achievement: Focus on attitudes and motivation*. London: Edward Arnold.

Gardner, R. C., & Lambert, W. E. (1959). Motivational variables in second-language acquisition. *Canadian Journal of Psychology*, *13*(4), 26–72.

Harvey, D. (2005). *A brief history of neoliberalism*. Oxford, UK: Oxford University Press.

Holborow, M. (2015). *Language and neoliberalism*. New York: Routledge.

Hyun, J.-E., & Kim, T.-Y. (2013). A study of Korean office workers’ motivation and demotivation in learning English. *Foreign Languages Education*, *20*(2), 163–188.

Jang, I. C. (2015). Language learning as a struggle for distinction in today’s corporate recruitment culture: An ethnographic study of English study abroad practices among South Korean undergraduates. *L2 Journal*, *7*(3), 57–77.

Jeong, K. (2015, August 17). “hetgallilttaen gwageoboonsa” ... jjikgi gareuchineun TOEIChakwon [“If you’re confused, choose past participle” ... private TOEIC institutes teach test-taking skills"]. *Chosun.com*. Retrieved from http://news.chosun.com/site/data/html_dir/2016/08/17/2016081700251.html

JTBC News. (March 9, 2016). *JTBC Newsroom*. Retrieved from http://news.jtbc.joins.com/article/article.aspx?news_id=NB11189537&pDate=20160309

Kachru, B. B. (1985). Standards, codification, and sociolinguistic realism: The English language in the Outer Circle. In R. Quirk & H. G. Widdowson (Eds.), *English in the world: Teaching and learning the language and literatures* (pp. 11–30). Cambridge, UK: Cambridge University Press.

Kim, M., Choi. D.-I., & Kim, T.-Y. (2018). A political economic analysis of commodified English in South Korean neoliberal labor markets. *Language Sciences*, 70, 82–91.

Kim, T.-Y. (2009). The sociocultural interface between ideal self and ought-to Self: A case study of two Korean students’ ESL motivation. In Z. Dörnyei & E. Ushioda (Eds.), *Motivation, language identity and the L2 self* (pp. 274–294). Bristol, UK: Multilingual Matters.

Kim, T.-Y. (2010). Socio-political influences on EFL motivation and attitudes: Comparative surveys of Korean high school students. *Asia Pacific Education Review*, *11*(2), 211–222.

Kim, T.-Y. (2017). The theretical interface between dynamic systems theory and sociocultural theory in L2 (de)mtovation research: A qualitative investigation. In M. Apple, D. Da Silva & T. Fellner (Eds.), *L2 selves and motivations in Asia contexts* (pp. 29-50). Bristol, UK: Multilingual Matters.

Kim, T.-Y., & Kim, Y.-K. (2016). A quasi-longitudinal study on English learning motivation and attitudes: The case of South Korean student. *The Journal of Asia TEFL*, *13*(2), 138-155.

Kubota, R. (2011). Questioning linguistic instrumentalism: English, neoliberalism, and language tests in Japan. *Linguistics and Education*, *22*(3), 248–260.

Lave, J., & McDermott, R. (2002). Estranged learning. *Outlines, 1*, 19-48.

Markus, H., & Nurius, P. (1986). Possible selves. *American Psychologist*, *41*(9), 954–969.

MBC News. (2016, May 4). Bakkwin TOEIC jeokeungharya sinspec gatchurya suheomsang ijunggo [Test-takers suffer to get familiarized with the new TOEIC and build up new specs] Retrieved from http://imnews.imbc.com/replay/2016/nwdesk/article/3962233_19842.html

Oxford, R., & Shearin, J. (1994). Language learning motivation: Expanding the theoretical framework. *The Modern Language Journal*, *78*(1), 12-28.

Park, J. S.-Y. (2009). *The local construction of a global language: Ideologies of English in South Korea*. New York: Mouton de Gruyter.

Park, J. S.-Y. (2011). The promise of English: linguistic capital and the neoliberal worker in the South Korean job market. *International Journal of Bilingual Education and Bilingualism*, *14*(4), 443–455.

Park, J. S.-Y., & Wee, L. (2015). *Markets of English: Linguistic capital and language policy in a globalizing world*. New York: Routledge.

Presidential Committee on the Younger Generation. (2014). *Urinara 100dae gieop ipsajiwonseoe banyeongdoen spec josa bogoseo [A report of the spec reflected in corporate applications required by the top 100 corporates in Korea]*. Retrieved from http://blog.president.go.kr/wp-content/uploads/2014/04/100_spec.pdf

Shim, H. (2015, August 25). Chojoonggosaeng jooleodeulja seongin sagyoyook sijang keojinda [The private market of adult education increases as the number of elementary and secondary students decrease], pp. 2014–2016. Retrieved from http://biz.chosun.com/site/data/html_dir/2014/08/24/2014082402709.html

Shin, H. (2015). Language "skills" and the neoliberal English education industry. *Journal of Multilingual and Multicultural Development*, *37*(5), 509–522.

Shin, H., & Park, J. S.-Y. (2015). Researching language and neoliberalism. *Journal of Multilingual and Multicultural Development*, *37*(5), 443–452.

Waninge, F., Dörnyei, Z., & de Bot, K. (2014). Motivational dynamics in language learning: Change, stability, and context. *The Modern Language Journal*, *98*(3), 704–723.

YBMNET. (n.d.). Yeonhyuk [History]. Retrieved from http://exam.ybmnet.co.kr/toeic/info/history.asp

YBMNET. (2015). Hwalyonghyeonhwang [The current state of how it's used]. Retrieved from http://exam.ybmnet.co.kr/toeic/status/grade.asp

Appendix

Questionnaire Section 2.

The perception of the socio-economic importance of the TOEIC in the job market

#	Question items	Answer disagree ---------- agree 1	2	3	4	5
1	Applicants with low GPAs need to secure higher TOEIC scores.					
2	It is okay to have a TOEIC score just higher than the minimum score required for the initial screening process.					
3	The TOEIC score is an indicator that reflects how diligently one has prepared for a job.					
4	Studying for the TOEIC hard will help one use English after employment.					
5	To defeat competitors in the job market, the TOEIC score needs to be as high as possible.					
6	If an applicant achieves a safe score in TOEIC, it is okay to stop studying for it.					
7	The TOEIC score is a tool to disqualify those who are not diligent.					
8	The content of the TOEIC is helpful for getting work done in the workplace.					
9	The higher one’s TOEIC score is, the more competitive one becomes in the job market.					
10	It is okay to have a good enough score on the TOEIC because it is the basic requirement for employment.					
11	The TOEIC score reflects the effort put into job hunting rather than the English proficiency.					
12	Studying for the TOEIC will be useful for performing job responsibilities.					

Questionnaire Section 3.

Motivation and demotivation to study for the TOEIC for employment

#	Question items	Answer				
		disagree ---------- agree				
		1	2	3	4	5
1	It is fun to know more about knowledge in English by studying the TOEIC.					
2	I study the TOEIC because I want to easily befriend with speakers from Anglophone countries.					
3	The TOEIC is important because getting into a good company or school would be difficult if the score is low.					
4	I study the TOEIC because I want to have a head start than others.					
5	Since it is happy to imagine myself getting a good job with a high score, I study the TOEIC.					
6	I study the TOEIC because my family and close friends say it is important.					
7	I feel that others would laugh at my TOEIC score.					
8	The process of improving my TOEIC score is exciting.					
9	I study the TOEIC because I want to talk with foreigners and engage in cultural exchange activities.					
10	I study the TOEIC in order to achieve a particular goal, like getting into a good company or school.					
11	I study the TOEIC because I don’t want to yield to any of my competitors.					
12	My dreams will come true, if I study the TOEIC hard.					
13	I study the TOEIC because my close friends and acquaintances think it is natural to study for it.					
14	I feel like I won’t be able to get the score I want no matter how hard I study for the TOEIC.					
15	I study the TOEIC because it is interesting and enjoyable.					
16	I study the TOEIC because I want to travel to or live in a foreign country such as the US or Canada.					
17	I study the TOEIC for my future career.					
18	I study the TOEIC hard because I will lose the competition if I have a low score.					
19	It is pleasant to imagine my TOEIC score improve, so I keep studying it hard.					
20	The reason I study for the TOEIC is that people around me say it is important for jobseekers.					
21	I get nervous whenever I need to submit my TOEIC score.					
22	I lost interest in studying the TOEIC because it is mandatory.					
23	I lost interest in studying the TOEIC because it takes too much time and effort.					
24	I lost interest in studying the TOEIC itself.					

Young Children's L2 Vocabulary Learning through Cooking: The Case of Korean EFL Children

Jaeuk Park
Newcastle University, UK
Naya Choi
Seoul National University, South Korea
Jieun Kiaer
University of Oxford, UK
Paul Seedhouse
Newcastle University, UK

Bioprofile:

Jaeuk Park is a PhD candidate at Newcastle University, UK, and is currently writing his thesis on the Korean Digital Kitchen. His research interests include foreign language teaching and learning, CALL, Multimodality, interactional competence and intercultural competence. **Email: j.u.park@ncl.ac.uk**

Naya Choi is an associate professor of Department of Child Development and Family Studies in Seoul National University. She studies children's language, literacy, and cognitive development and education. **Email:** choinaya@snu.ac.kr **(corresponding author)**

Jieun Kiaer an associate professor in Korean Language and Linguistics in the Oriental Institute of the University of Oxford. Her research interests include the pragmatic aspects of syntactic architecture and the role of prosody in syntax, and also second language acquisition. **Email: jieun.kiaer@orinst.ox.ac.uk**

Paul Seedhouse is professor of Educational and Applied Linguistics and School Research Director in the School of Education, Communication and Language Sciences, Newcastle University, UK. His monograph *The Interactional Architecture of the Language Classroom* was published in 2004 and won the Modern Languages Association of America Mildenberger Prize. He has worked with colleagues in Computer Science at Newcastle to develop iLab: *Learn*, a centre to develop digital educational technology. He led the French Digital Kitchen project, which won the EU Language Label prize in 2011. **Email:** paul.seedhouse@ncl.ac.uk

Newcastle University, Newcastle upon Tyne, NE1 7RU, UK
Seoul National University, 222-408, 1 Gwanak-ro, Gwanak-gu, Seoul, South Korea 08826
University of Oxford, Oxford, OX1 2JD, UK

Abstract

This paper aims to combine cooking with language learning to create a sustainable language learning environment for young children. We investigate the pedagogical benefits of cooking as a task in Task-based Language Teaching (TBLT) for young learners of English in South Korea. We follow the definition of 'task' by Bygate *et al.* (2001). Our focus is on vocabulary acquisition. Using both quantitative and qualitative methods, we prove that tasks such as cooking produces not only better performance, but also creates a positive, enjoyable learning environment for young children. The benefits of cooking in language learning within a TBLT setting has been widely attested in the case of adult learners who are learning European languages (Seedhouse, 2017). There is, however, a lack of equivalent research which demonstrates the pedagogical impact of cooking in the case of young learners, although this is often generally assumed in early-years education.

Key words: Cooking, Young EFL learners, TBLT, vocabulary learning, English Fever in South Korea

Introduction

The current environment in South Korea is one in which Korean children are forced to learn English from a very young age, sometimes to the detriment of their academic, psychosocial, and physical wellbeing (Park, 2007). Under the umbrella of the South Korean educational sector, English education stands out as one of the most worrying social phenomena that young Koreans encounter. This zeal for English education has become so extreme that each day, on average, around 100 young Korean students go abroad for their education (Park, 2012). Learning a new language at an early age is not a problem in itself - the problem lies in strict, rigid teaching methods young learners experience, which are often based on total immersion and "One Language at a Time (OLAT)" or an English-only policy. Such methods been heavily criticised by many scholars (Auerbach, 1993; Cook, 1997; Firth & Wagner, 1997; Kachru, 1994; Park, 2007 among others). Kiaer and Morgan (2016) report various case studies where young children suffer as a result of these methods. In this study, we introduce an innovative way in which young learners of English can enjoy learning as well as perform better. We hope the outcome of this study will raise awareness to the problems with the current English teaching methods used South Korea and other Asian countries with English fever.

Literature Review

Young children's English language learning in South Korea

Due to the English Fever phenomenon in South Korea, many young children suffer from foreign language anxiety (FLA) towards English (Park, 2007; 2009). As a result of highly memorisation-based teaching and pressure for high performance, young children not only lose their motivation to learn a new language but further develop foreign language anxiety and related psychological problems. Kiaer and Morgan (2016) report a case study of four-year old children who developed different nervous habits. Similar cases have been reported in other Asian countries like Taiwan (Liu & Chen, 2013).

Cooking in Task-based language teaching (TBLT)

In this study we present cooking as an innovative language learning method for young EFL learners to help them learn L2 English without FLA. We explored the role of cooking as a task in Korean young children's English vocabulary learning. Task-Based Language Teaching (TBLT) is a well-established approach to language learning which prompts learners to achieve a goal or complete a task (Skehan, 1998; 2003). According to Bygate *et al*. cited by Ellis (2003), task is "an activity which requires learners to use language, with emphasis on meaning, to attain an objective" (p.5). Much like real-world tasks such as asking for directions, TBLT aims to develop students' language through providing a task and then using language to solve it. Some of the main principles of TBLT are as follows: meaning is primary (language use rather than form); there is a communication problem to solve; classroom tasks relate directly to real world activities; the assessment is done in terms of outcomes. Not only does Ellis (2003) insist that "there is a clear psycholinguistic rationale (and substantial empirical support) for choosing 'task' as the basis for language pedagogy" (p. 320), but Skehan (1998) suggests that tasks can provide learners with an opportunity to get involved in activities, which in turn sharpen their grasp of the language. Tasks thus serve as powerful mediators of language learning. Seedhouse *et al.* (2014) in particular shows that situated tasks like cooking make language learning effective yet also enjoyable**.**

Using the real-world task of cooking, researchers have demonstrated its pedagogical effects on foreign vocabulary learning. Pallotti *et al*. (2017) have shown two empirical studies relating to cooking: one in which 50 adult participants were engaged in a cooking activity to assess 16 English vocabulary items, and the other where 36 subjects got involved in the same task to learn 10 Italian vocabulary items. Both studies demonstrated statistically significant improvements in the acquisition of vocabulary items. In addition to this, Park and Seedhouse (2017) explored the effect of touching real objects on vocabulary acquisition by comparing the learning outcomes from a real-world learning environment with classroom environments using a quasi-experimental design. Working with 48 international participants, this paper has demonstrated a significant difference in Korean vocabulary learning. The study has contributed to building up one more dimension of psycholinguistic factors to vocabulary learning in the field of applied linguistics by showing that engaging in physical activities makes a difference. Central to their arguments was that cooking tasks are beneficial for foreign vocabulary learning, and these empirical studies as a whole offer a detailed perspective on vocabulary learning.

Cooking in young children's learning

Cooking in TBLT has been widely used for adults' L2 language learning (Seedhouse, 2017), yet to our knowledge it has not been adopted at all for young children's L2 language acquisition so far. Yet, cooking is one of the most popular activities adopted in young children's learning in general. Young children learn best when they are actively experiencing the world. They can explore cooking materials using all five senses and learn physical and chemical scientific concepts through mixing, cutting, and kneading, as well as basic concepts such as size, shape, color, taste, touch and smell (Dahl, 1998). Cooking is an interesting activity that can make a very effective and authentic experience in a curriculum that integrates all areas of language, mathematics, science, art, nutrition, and creativity, as well as enjoyment of food (Dahl, 1998; Jill, 1995). For young children, the multisensory learning experience that cooking provides can boost their eagerness and enthusiasm in foreign language learning, particularly encouraging their interactive, spontaneous participation during the session (Griva & Sivropoulou, 2009).

In spite of a range of theoretical claims on the effects of cooking for vocabulary learning, there has been few empirical study which examines the link between a real-world activity and its pedagogical influence. The majority of vocabulary studies on children's L2 vocabulary learning in context have focused on learning with texts presented in a single modality. In recent years, many studies have explored the effects of presenting information using multiple modalities such as text, audio and visual materials on L2 vocabulary learning (Al-Seghayer, 2001; Duquette & Painchaud, 1996; Mousavi *et al*., 1995). The results showed positive effects in retention of words learned from multimodality. However, researchers have paid insufficient attention to one more dimension for young learners' vocabulary learning – kinesthetic mode.

In summary, previous studies are limited in that they have applied a cooking-based approach only to adult learners in European contexts for vocabulary learning, not young learners. Taking these limitations into consideration and hoping to further research in this area, this study attempts to explore how a cooking activity engaging all five senses assists young Korean EFL learners' vocabulary learning. This paper is expected to broaden the research scope by exploring its impact on young learners in other contexts.

Methodology

Research Design

This study employed a mixed-methods approach which can bridge the gap between different ways of seeing, interpreting and knowing (Greene, 2007). Combining both quantitative (QUAN) and qualitative (QUAL) approaches has been known to offer 'a powerful third paradigm choice that often provide the most informative, complete, balanced, and useful research result' (Johnson *et al*., 2007, p.129). Of the various mixed-methods research models, we chose Creswell and Plano-Clark's (2011) *Explanatory Design*. This design starts with QUAN data, followed by QUAL data for detailed exploration of the findings. In this study, test results demonstrated the extent to which young students from two different groups could learn through a cooking activity, while interviews explain why learning differences occur. That is, QUAN offers 'learning outcomes' whereas QUAL data show 'learning processes' for the different outcomes.

The Experiment

This study has three features: pre-test -- post-test design, a treatment group and a control group, and random assignment of study participants (Shapley *et al.*, 2010). The classical experimental study involves comparing the impact of an intervention with what would have happened if there had been no intervention. What is essential for the validity of an experiment is the randomization of participants. The experiment 'engenders considerable confidence in the robustness and trustworthiness of causal findings' (Bryman, 2012, p.50). In other words, an experimental research design assures internal validity. This design was therefore fitting for this paper.

To explore the hypothesis of this study, participants were recruited and randomly assigned to one of two groups: the treatment group and the control group. To minimise confounding variables, the participants in the same year were chosen in the same school.

The treatment group carried out a cooking task using tangible objects, and the control group conducted the same cooking task, simply with photos of the objects. To be specific, the difference was that the former group experienced a real-world cooking, whereas the latter one simulated the activity. This study aimed to examine if a meaningful task using all five senses is more effective than an ordinary task involving fewer senses with regards to vocabulary acquisition. The intervention was thus whether young learners used all five senses (in particular, touch and taste) or fewer senses (sight) in performing the task of cooking. Two groups alike went through pre-, immediate post- and delayed post-testing two weeks later. To implement the experiment, we adopted Nation and Webb's (2011) experimental design as illustrated below.

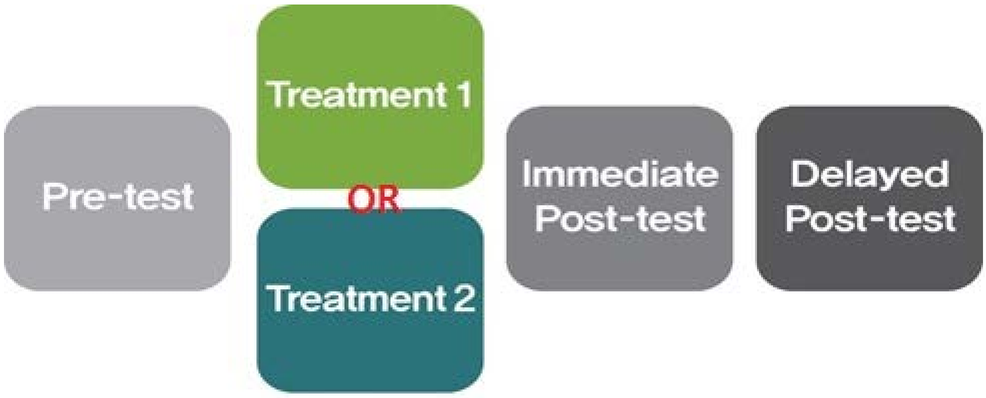


Figure 1 Experimental Design for Vocabulary Learning (Nation and Webb, 2011)

Participants

Participants were all from the same kindergarten in Kimpo, Gyeonggi Province in South Korea. The kindergarten is located in a relatively rural area and English is not the main focus of education unlike many kindergartens in Seoul, particularly in wealthy area like Gangnam. They had almost an ab-initio level of English background. For this study, 66 young learners were chosen and then randomly divided into two groups with 33 people in each group, on the assumption that 30 would be targeted in case a few students do not turn up. As participants were young learners whose physical and intellectual development vary month by month which might be one of the confounding variables, the specific age for each group is described. Overall average age of all participants was 66.92 months (*SD* = 2.90). The average age of OG students was 66.48 (*SD* = 2.65), while the age of their PG counterparts was 67.36 (*SD* = 3.12), showing no significant difference in age.

The number of children unable to complete the experiment was higher than our initial predictions. This was due to a high drop-out rate and young students' lack of concentration resulting in incomplete task implementation. A total of 50 students took part in two groups with 21 males (42% percent) and 29 females (58%).

Table 1 Participants

	Objects Group	Photo Groups	Total
Male	12 (48%)	9 (36%)	21 (42%)
Female	13 (52%)	16 (64%)	29 (58%)
Total	25	25	50

Target words selection

We selected 10 nouns as the target words. We chose nouns because they are the easiest word class to learn, and 'concrete nouns are learnt more easily than abstract ones, and because the concrete noun items are learnt more quickly and effectively if objects are nonverbally referred and used as stimuli' (Ellis & Beaton, 1993, cited by Seedhouse, 2017, p.210). Depending on the number of syllables of a target word, three of them (cranberry, cashew nut and persimmon) are very difficult words, four of them (bowl, raisin, strawberry and plate) are medium-level difficulty and three (banana, honey and yogurt) are easy words. English words are prevalent in contemporary spoken Korean. Young children may be more familiar with English words than Sino-Korean words (Kiaer, 2014). For example, the word *banana* does not have a native Korean counterpart, so we expect everyone may know this word. As the dish young students made was fruit salad, one of the children's favourites, 10 target vocabulary items included the name of ingredients and equipment necessary to make it.

Task and test procedures

The basic task and test procedures followed Park and Seedhouse's (2017) model below, with a slight amendment. In classroom A, a teacher used real materials physically demonstrating, how to make the dish, and in classroom B she simply employed photos simulating how she would make it. This was how a teacher guided two groups to be able to carry out the cooking task according to a three-phase framework of TBLT. The teacher was an American student majoring in Korean in Korea University. She was fluent both in Korean and English. She introduced herself in Korean in order to help children to feel relaxed. All of the instructions in the class were delivered both in English and in Korean, firstly in English then in Korean, in both groups.

Figure 2 Test and Task procedures

In *pre-task*, the teacher asked students to collect each item, such as 'cashew nut' and 'honey', one by one. Students were required in *during-task* to manipulate what they had collected. The control group learners took the same step, simulating the dish-making activity with photos. A teacher offered instructions such as 'put cashew nuts into a bowl'. During both stages, the teacher offered verbal repetition of the word when requested, but not more than twice for the same word.

The post-task involved eating the dish as well as sharing their ideas about which food item was the most delicious. The two photos below show the stark contrast between the two classrooms.

As the written form of a word constitutes part of vocabulary knowledge, the teacher used visual materials throughout the lesson in both classrooms, in the form of PowerPoint (PPT) slides on a screen, on which the written form of each target word was shown. There were 46 slides, which showed young learners individual items of the target word as well as images of the process of making the dish (Appendix A). The same slides were projected in both groups to avoid any confounding variables.

Figure 3 Different Learning experiences in the during-task step of mixing ingredients – Photo Group (PG) vs. Object Group (OG)

Figure 4 Using the PPT in the class

To gauge students' vocabulary knowledge after the task, there were three tests: pre-test, immediate post-test and delayed post-test. The pre-test was an oral production test. Each individual was shown ten real objects and then asked to produce them in English one by one, during which a research assistant held an audio-recorder by hand to record his or her performance. No receptive test was administered in the pre-test, as learners' random matching may compromise the actual vocabulary learning. The two post-tests consisted of a receptive and productive test. The productive test was intentionally always administered before the receptive test. This was because if learners first take the receptive test in which written words of the target vocabulary are provided, this may expose learners to the target word once more, potentially contributing to the increase in vocabulary scores. The receptive test involved matching written names of target words with the correct photos on a test sheet (Appendix B) within 2 minutes. They were instructed to keep what they did not know blank.

This was because otherwise they would take a guess which may compromise the assessment of their actual knowledge. The productive test was the same as in the pre-test. Two weeks later, delayed post-tests involved both receptive and productive tests.

In the receptive tests, each correct answer was counted for one point with the wrong counting for zero, and totals for each condition were then calculated. However, a different method was needed in the verbal production tests to quantify the ability in the L1-L2 & L2-L1 production as students might be able to show low levels of learning. Therefore, this study adapted the Lexical Production Scoring Protocol-Spoken (LPSP-Spoken) that (Barcroft, 2002) suggests as a way to quantify the scores. As learners acquire new words in bits and pieces, the measure that is sensitive to partial word learning is appropriate (Barcroft, 2002). The framework includes a five-way scale: 0, 0.25, 0.5, 0.75, and 1, depending on how a learner performs in production tests (Appendix C). It was based on this scoring scheme that this study could mark learner's partial knowledge of lexical items. The grade range is as below.

Table 2 Marking Range

	Minimum	Maximum
Productive test	0	10
Receptive test	0	10

Results

Learning outcomes

This section presents statistical evidence to determine whether the treatment group shows greater gains in vocabulary learning than the control group. This quantitative evidence demonstrates students' learning as a product of cooking as a TBLT activity.

This study employed two statistical approaches for analysis (Pallant, 2013). First, as the study saw a change in vocabulary scores over the three time periods (pre, immediate and delayed), it used the 'One-Way Repeated Measures ANOVA', which is a technique to determine whether there is a significant difference in the data in different columns. However, this tool cannot explain the difference of the mean scores between two groups or two intervals for comparison. Therefore a t-test was used. The next table presents the results of a one-way repeated measures ANOVA on Pre- (prior to the intervention), Immediate (following the intervention) and Delayed- (two-week follow-up) for productive and receptive tests in order.

Table 3 Mean Differences for Productive Tests through ANOVA

Group	Pre-test	Immediate-test	Delayed-test
Objects Group	1.53 (1.44)	3.20 (1.50)	3.30 (1.90)
Photos Group	1.09 (1.11)	2.25 (.93)	2.10 (.92)

Two groups experienced a gradual rise in vocabulary scores, with a little decline in Photos Group between Immediate- and Delayed-test. There was a significant effect for time, F (2) = 62.59, $p < .001$, multivariate partial eta squared = .69, which indicates a large effect size. Moreover, there were also significant interaction effects between Time and Group ($p < .05$).

Table 4 Mean Differences for Receptive Tests through ANOVA

Group	Pre-test	Immediate-test	Delayed-test
Objects Group	1.53 (1.44)	3.08 (2.97)	3.64 (3.37)
Photos Group	1.09 (1.11)	1.68 (1.68)	1.60 (1.56)

Table 4 shows similar results for the receptive and productive tests. Both groups saw an increase in vocabulary acquisition over time, with a little drop in Photos Group between Immediate- and Delayed-test. There was also a significant effect for time, F (2) = 10.68, $p < .001$, multivariate partial eta squared = .18, which shows a large effect size. Furthermore, there were also significant interaction effects between Time and Group ($p < .05$). These significant changes are seen in line graphs below.

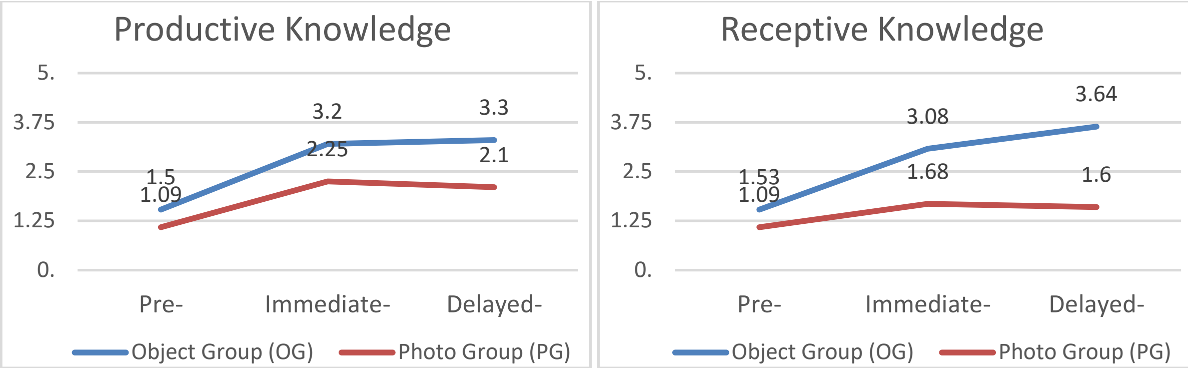


Figure 5 Productive and Receptive Test Results

There was significant difference in vocabulary points over three periods for both productive and receptive tests. The following table demonstrates the difference between the mean scores of each group.

Table 5 Mean scores of Vocabulary Knowledge

Tests	Group	OG (n=25)	PG (n=25)	*T*	*p*
Productive Knowledge	Pre	1.53 (1.44)	1.09 (1.11)	1.21	.232
	Immediate post	3.20 (1.50)	2.25 (.93)	2.69*	.010
	Delayed post	3.30 (1.90)	2.10 (.92)	2.84**	.007
Receptive Knowledge	Immediate post	3.08 (2.97)	1.68 (1.68)	2.05*	.047
	Delayed post	3.64 (3.37)	1.60 (1.56)	2.75**	.009

The table above shows the overall vocabulary gains over time between two groups. Although there was a slight difference between two groups (1.53 and 1.09) in vocabulary knowledge in the pre-test, the difference was not statistically significant. This shows that both PG and OG groups do not have any background English. Overall vocabulary gains over time are clearly shown in Table 4. In both productive and receptive knowledge, OG showed a greater increase in vocabulary acquisition than PG.

The mean differences between the pre-test and post-test should demonstrate change over time to prove that the OG benefitted from using real objects for vocabulary learning. Thus, we compared whether there is a statistically significant gap between immediate post-test with pre-test and delayed-post-test with pre-test. The results are shown in Table 6 below.

Table 6 Mean Differences between pre-test and Productive immediate post-test

Group	*M(SD)*		*T*	*p*
	Pre	Immediate post		
Object Group	1.53 (1.44)	3.20 (1.50)	-7.81***	.000
Photo Group	1.09 (1.11)	2.25 (.93)	-5.53***	.000

Table 7 Mean Differences between pre-test and Productive delayed post-test

Group	*M(SD)*		*T*	*p*
	Pre	Delayed post		
Object Group	1.53 (1.44)	3.30 (1.90)	-7.77***	.000
Photo Group	1.09 (1.11)	2.10 (.92)	-4.91***	.000

In tables 6 and 7, we can see that the two groups saw increases in vocabulary gains in the productive tests and this was statistically significant. The mean differences were, however, higher in OG than PG: for immediate test, OG (MD = 1.67, t = -7.81, p < .001) and PG (MD = 1.16, t = -5.53, p < .001); for delayed test OG (MD = 1.77, t = -7.77, p < .001) and PG (MD = 1.01, t = -4.91, p < .001). Similar results were seen for receptive tests as in Tables 8 and 9 below. While both groups registered a rise in learners' vocabulary gains, OG only showed statistically significant levels: immediate test (MD = 1.55, t = -3.11, p < .01); delayed test (MD = 2.11, t = -3.68, p < .01).

Table 8 Mean Differences between pre-test and Receptive immediate post-test

Group	*M(SD)*		*t*	*p*
	Pre	Delayed post		
Object Group	1.53 (1.44)	3.08 (2.97)	-3.105**	.005
Photo Group	1.09 (1.11)	1.68 (1.68)	-1.477	.153

Table 9 Mean Differences between pre-test and Receptive delayed post-test

Group	*M(SD)*		*t*	*p*
	Pre	Delayed post		
Object Group	1.53 (1.44)	3.64 (3.37)	-3.675**	.001
Photo Group	1.09 (1.11)	1.60 (1.56)	-1.786	.087

The mean differences in vocabulary scores were higher for OG than PG for both the post-test and delayed-post-test, and in both receptive and productive areas. The differences were statistically significant in all cases. Interestingly, in both productive and receptive tests, OG experienced gradual increase in their gains over time from 3.20 to 3.30 and 3.08 to 3.64, whereas PG registered a decline from 2.25 to 2.10, and 1.68 to 1.60.

These results clearly demonstrate that when carrying out a cooking activity, manipulating real objects is more effective for vocabulary acquisition than simply looking at photos. Young learners were able to learn English vocabulary items at a significantly higher level for both spoken and written media, and in both the immediate-post-test and delayed-post-test. Furthermore, even with regards to long term gains, students using real objects scored more than their counterparts.

The overall scores were higher in productive tests than in receptive tests. It is probably because of the nature of the lecture and tests. Students could be exposed to the phonological sounds of target words during a teacher's demonstration, which might help them take oral productive tests. On the other hand, in receptive tests where an ability to read the words and match with photos was required, a teacher rarely offered help to students. That might explain why there was a gap between two tests. What was really interesting was the score differences of stark contrast in all tests, even delayed post-tests between two groups. What could account for this?

Learning processes

The difference in the scores of the two groups can be understood by exploring the learning processes involved through analysis of the focus group interview transcripts. The interview approach is valuable because it "offers the opportunity of allowing people to probe each other's reasons for holding a certain view" (Bryman, 2012, p.503). This section presents learners' views about their own learning experiences. Six kids from each group (12 in total) were randomly selected to participate in focus group interviews. The interview followed the administration of the immediate post-test after the cooking session, and the interview question sheet is attached (Appendix D).

The interview was conducted in Korean and translated into English. Interviewee's answers were short but very clear. The first question was about whether or not, and why, they like cooking when they are at home with family members. All kids like cooking itself for reasons as shown in some illustrative verbatim quotes below:

Table 10 Answers to "Do you like cooking? If so, are there any reasons?"

Object Group (OG)	Photo Group (PG)
Yes. (Jae, Han, Ju, Cha, Tae, Seo) *Just a big fun. (Jun)* *Fun helping my mum. (Jae)* *It was great fun to bake bread together with my mum. (Cha)*	*Yes. (Da, Jun, Min, Eun, Hye, Son)* *I like just making the dish. (Min)* *Fun eating the dish. (Da)* *Because I can eat it and if it's delicious I can cook it myself again. (Eun)*

Clearly, every young learner enjoyed the activity of cooking and there were various reasons for this. They enjoyed eating and tasting the dish, doing the activity itself and doing something with their parents. This shows their experience is related to the physical senses. When performing activities, in particular cooking, all five senses are activated in relation to their learning and memory input (Trubek & Belliveau, 2009). Furthermore, young learners were fond of 'doing' itself, which helps learn a foreign language (Doughty & Long, 2003). This led us to wonder if 'learning by doing' had helped them gain the vocabulary knowledge. Asked if they had learned English by cooking, two groups had different and similar views as shown:

Table 11 Answers to "Does a cooking activity help you learn English words? If so, how did it help you?"

Object Group (OG)	Photo Group (PG)
Strawberry (Han) *Cranberry (Jae)* *Bowl (Han)* *I learned how to cook. (Jae, Ju, Cha)* *I learned English and how to cook. (Han, Tae)* *I also love learning English through singing. (Chae)*	*I learned 'strawberry'… that's it. (Son* *I love learning English through cooking and singing. (Son)*

Students in both groups had learnt English vocabulary items as demonstrated by the statistics discussed above, but to a different extent. Students from OG could make more sounds of the target words than PG learners, according to the interview scripts. Interestingly, the learners in both groups saw ‘cooking’ and ‘singing’ as a motivating vehicle by which they can learn English. Just as (Paterson & Willis, 2008) found that music can provide an excellent way for children to learn and memorise words and phrases, young learners were found to be more engaged in learning English by ‘doing’ - carrying out real world activities. In a similar way, ‘cooking’ opened a space in which kids are exposed to spoken and written English with a clear purpose for listening to English and trying to understand in a clear context for using English.

We can see specifically how children think about their own experience in the next verbatim quotes.

Table 12 Answers to “Which one would you prefer between making the dish using real objects and cooking just using photos of them? In either case, why do you think it is more fun?”

Objects Group (OG)	Photo Group (PG)
Real (Han) *Fun eating the dish. (Ju)* *Delicious (Han)* *Smell of yogurt and strawberry (Ju)* *I could feel the taste. (Han)* *Strawberry and plain yogurt changed its colour from white to pink when mixed. (Jae)*	*Not real (Eun)* *It was a fake! (Da)* *Boring (Min, Eun)* *Not being able to eat (Min)* *No fun because there is no food. (Jun)* *Angry because I couldn’t eat. (Da, Eun)*

OG students found their learning activity more real, enjoyable, hands-on with five senses, whereas PG learners were found the activity as not real and boring, due to not being able to eat the food. Two students even became angry, just because they had no end product. Thus, whether the activity children had performed was ‘real’ or ‘unreal’ determined their attitudes toward two difference experiences. We thought this might result in very different levels of motivation for students in two cases, which led to different level of learning processes.

Throughout the interviews, we noticed something distinctive which might account for the difference in their learning outcomes. It is illustrated in their comments below:

Table 13 Answers to "Was the dish you made any good? What did you like? (OG)" "How did you feel when cooking the dish using photos? What did you like? (PG)"

Object Group (OG)	Photo Group (PG)
I'd like to make it again at home tonight. (Han, Jae) *I remember the recipe. (Jae, Han, Cha, Ju)* *first strawberry, (Han, Ju-in chorus) then banana, (Ju) and persimmon and raisin (Han)*	*Strawberry, banana and cranberry (Da)* *Banana (Hae)* *Banana and strawberry (Da)*

The real activity of cooking with end product was so impressive that children wanted to do the same task again when they got home. Asked how they could do it again, they said that they would try to remember the procedures of making the dish. This helped them not only think about the recipe, but practice and repeat the target vocabulary one by one. Repetition has a tremendously powerful consequences for the learning process (Kurhila & Kotilainen, 2017). It is in this moment of reflection that they were internalizing and learning the linguistic information. On the other hand, there was no sign of this process in a PG case. They were able to recall a few vocabulary items, but no more meaningful action for language learning was shown.

On the whole, cooking the dish using real objects was found to be more enjoyable and effective, which increased their motivation, thereby contributing to students' vocabulary learning. In contrast, simple observation of photos was seen as boring and unreal, which resulted in relatively less successful outcomes. This is not to say that students did not have a sense of motivation in their experience, but rather suggests that multi-sensory materials could afford a relatively higher level of learning compared to the other case. Highly motivated learners achieved greater success in language learning than those who are not (Gu & Johnson, 1996; Sanaoui, 1995). It is for these reasons that two groups achieved different learning outcomes.

The reasons for different learning outcomes and processes also emerged from interviews of a teacher, who led the two classrooms. She pointed out clear differences in terms of motivation as shown below:

Q: What was your teaching experience like?

A: it was really neat to be able to do an experience like this to see the difference between kids learning when they actually do the experiment as opposed to just learning vocabulary by flashcards or pictures.

Q: Did you feel any difference between the two groups?

A: Yeah the first group was just the picture group, and I think they were kind of disappointed and less active in the experiment because it wasn't actually food and just pictures so it didn't mean as much to them, I think. And then the second group was more active and wanted to participate more I think because it was actually food and they knew they'd be able to eat it late […] just because they know they can eat the food because they have it and if they're only doing it with pictures, it's fun, but it doesn't give as much motivation. I think also because kids like to be able to do things by themselves and if it's actually real then they can touch it and feel it, and it's more sensory.

The teacher found that while PG students could have fun, but to a lesser degree due to only photos and lack of end product, OG learners showed more active participation because of real objects and foods they can touch and feel. The multi-sensory activity gave them an unrivalled level of motivation, thus enhancing vocabulary learning. This supports the findings of a previous study (Trubek & Belliveau, 2009). However, it turned out that she was wary of the feasibility of the task itself as shown below:

Q: Do you think this is feasible to use in institutions to teach?

A: I think they definitely could, maybe not as an everyday thing, that would be pretty hard because it takes a lot of time, money, resources, and other things, but I think every once in a while it's definitely a good idea.

When asked in terms of pedagogic practicality in a classroom, she said that available resources and commitments might make it hard for the task to be implemented. Given the current educational situation for Korean EFL young learners in terms of likely barriers, it is true that the implementation of real world tasks is hard to be put into practice. Nevertheless, we recommend applying real world tasks. This study indicates that a cooking activity might be able to be used as a resource to increase students' motivation. The findings suggest that real world tasks serve as a vehicle through which EFL young learners improve language skills in an enjoyable manner.

Discussion

This paper aims to combine cooking with language learning to create a sustainable language learning environment for young children. The so-called English Fever in South Korea has been the main cause of foreign language learning anxiety in young Korean EFL learners (Park, 2007). In order to create a fun and enjoyable language learning environment for young learners, in this paper we adopted cooking in Task-Based Language Teaching (TBLT). The results support the claims of TBLT in that language learning combined with real-life connection is much more effective than language learning without any real-life connection or task (Seedhouse, 2017). Focus group interviews show how children who have participated in real cooking activity enjoyed the lesson, whereas children who only learned vocabulary through manipulating non-real objects like photos soon lost their interest.

In South Korea, for EFL vocabulary learning, the repetition-based method without any real-world connection is extremely popular and widespread. Although this approach may be efficient in short term, the vocabulary learnt without any real-world connection may soon be lost in due course. Similarly, (Peterson & Mulligan, 2012) report a surprising negative repetition effect, in which participants who studied a list of cue-target pairs twice recalled fewer targets than a group who studied the pairs only once. TBLT-based vocabulary learning may look more costly and less effective compared to simple repetition-based vocabulary learning. Contrary to this belief, our study shows that real-life connection and the actual activity are the powerful tools for vocabulary learning. These findings have both pedagogical and practical implications. Educators might be able to come up with other tasks and activities for children (e.g., such as traditional games) and adapt them for the classroom as pedagogical resources. We can see many examples of schools putting into practice programs where, for example, students are sent on field trips to places like museums. We believe that this trend is related to the physical component of such learning experiences and the associated educational benefits. Given the sheer joy of experiencing a real-world activity, the curriculum will be able to be developed in a way that has a great deal to offer to language learners.

Conclusion

Due to English Fever, young EFL learners in South Korea not only lose motivation to learn a language but also develop foreign language anxiety and related psychological, physical and social problems. One of the key causes of the problems is a highly pressurised environment and demand for high performance. This paper shows an innovative way of learning; language learning through cooking. We show that using cooking as a task in the TBLT can help young EFL learners not only enjoy language learning but also help them to perform better. Despite theoretical claims that using all five senses would naturally contribute positively to young learners' foreign language learning, there have not been many studies to support such claims. In this sense, this paper builds up on previous literature by conducting empirical research showing that physical, real-life tasks significantly make a difference in young EFL learners' vocabulary learning, and also affects their level of their motivation. Considering a rising interest in cooking in many countries, we believe that the approach can be widely used.

Although this study attempted to support its argument, participants were too young for us to elicit strong opinions in interviews. It is therefore recommended that older groups of young learners be included in future research in order to examine their perspectives and attitudes in more detail towards learning through real-life tasks. Analysis through participants' detailed comments on such learning methods would reveal the potential of this real-world task-based approach.

References

Al-Seghayer, K. (2001). The Effect of Multmedia Annotation Modes on L2 Vocabulary Acquisition: A Comparative Study. *Language Learning & Technology*, 5(1), 202-232.

Barcroft, J. (2002). Semantic and Structural Elaboration in L2 Lexical Acquisition. *Language Learning*, 52(2), 323-363.

Bygate, M., Skehan, P. & Swain, M. (2001). *Researching Pedagogic Tasks: Second Language Learning, Teaching and Testing*. Longman.

Dahl, K. (1998). Why cooking in the curriculum?. *Young Children*, 53(1), 81-83.

Doughty, C. J. & Long, M. H. (2003). Optimal Pycholinguistic Environments for Distance Foreign Language Learning. *Language Learning & Technology*, 7(3), 50-80.

Duquette, L. & Painchaud, G. (1996). A comparison of vocabulary acquisition in audio and video contexts. *Canadian Modern Language Review*, 53, 143-172.

Ellis, R. (2003). *Task-based language learning and teaching*. Oxford: Clarendon Press.

Greene, J. C. (2007). *Mixed methods in social inquiry*. San Francisco: Jossey-Bass.

Griva, E. & Sivropoulou, R. (2009). Implementation and Evaluation of an Early Foreign Language Learning Project in Kindergarten. *Early Childhood Education Journal*, 37, 79-87.

Gu, Y. & Johnson, R. K. (1996). Vocabulary Learning Strategies and Language Learning Outcomes. *Language Learning*, 46(4), 643-679.

Jill, K. (1995). Cooking in the kindergarten. *Young Children*, 50(6), 32-33.

Kiaer, J. (2014). The history of English words in Korean. Lincom Europa.

Kurhila, S. & Kotilainen, L. (2017). Cooking, Interaction and Learning: the Finnish Digital Kitchen as a Language Learning Environment, in P. Seedhouse (Ed.) *Task-Based Language Learning In an Immersive Digital Environment: The European Digital Kitchen*, 157-179. London: Bloomsbury.

Liu, H. J. & Chen, T. H. (2013). Foreign Language Anxiety in Young Learners: How It Relates to Multiple Intelligences, Learner Attitudes, and Perceived Competence, *Journal of Language Teaching and Research*, 4(5), 932-938.

Mousavi, S., Low, R. & Sweller, J. (1995). Reducing cognitive load by mixing auditory and visual presentation modes. *Journal of Educational Psychology*, 87, 319-334.

Pallant, J. F. (2013) *SPSS survival manual : a step by step guide to data analysis using IBM SPSS / Julie Pallant*. Crows Nest, N.S.W: Allen & Unwin.

Park, J. K. (2007). *Korean Parents in "English Fever" and "early study-abroad" children in the U.S.: parental beliefs and practices concerning first-language peers*. Indiana University.

Park, J. K. (2009). "English fever" in South Korea: its history and symptoms. *English Today* 25(1), 50-57.

Paterson, A. & Willis, J. (2008). *English through music*. Oxford: UK: Oxford University Press.

Peterson, D. J. & Mulligan, N. W. (2012). A negative effect of repetition in episodic memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 38, 1786-91.

Sanaoui, R. (1995). Adult Learners' Approaches to Learning Vocabulary in Second Languages. *The Modern Language Journal*, 79(1), 15-28.

Seedhouse, P. (2017). *Task-Based Language Learning In an Immersive Digital Environment: The European Digital Kitchen*. London: Bloomsbury.

Seedhouse, P., Preston, A., Olivier, P., Jackson, D., Heslop, P., Balaam, M., Rafiev, A. & Kipling, M. (2014). The European Digital Kitchen Project. *Bellaterra Journal of Teaching & Learning Language & Literature*, 7(1), 1-16.

Shapley, K. S., Sheehan, D., Maloney, C. and Caranikas-Walker, F. (2010). Evaluating the implementation fidelity of technology immersion and its relationship with student achievement. *Journal of Teaching, Learning, and Assessment*, 19(4). 3-68.

Skehan, P. (1998). *A cognitive approach to language learning*. Oxford: Oxford University Press.

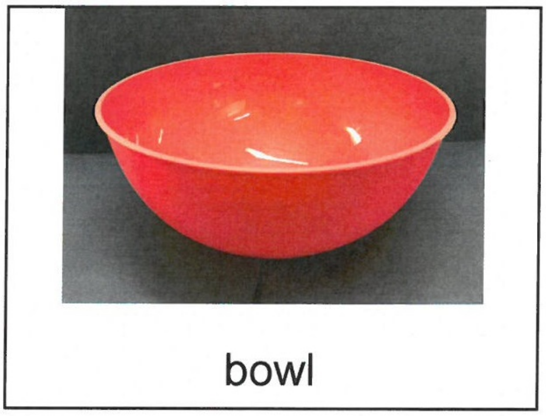
Skehan, P. (2003). Task-based instruction, *Language Teaching*, 36(1), 1-14.

Trubek, A. B. and Belliveau, C. (2009). Cooking as Pedagogy: Engaging the Senses through Experiential Learning. *Anthropology News*, 50(4), 16-16.

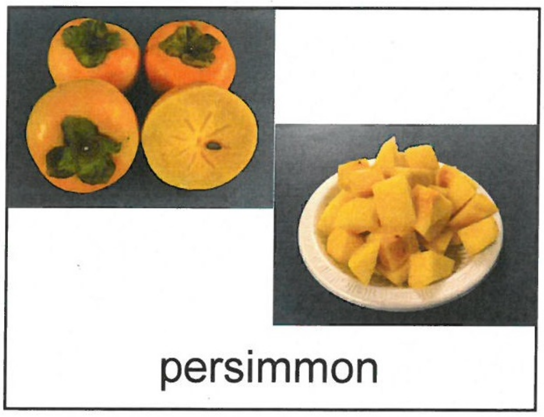
Appendices

Appendix A PPT slides (Shortened version)

bowl



persimmon



Now,
let's start cooking!

Get a bowl ready in front of you.

Put persimmon
into a bowl.

Put persimmon into a bowl.

Put cashew nuts
into a bowl.

Put cashew nuts into a bowl.

Pour yogurt into a bowl.

Pour yogurt into a bowl.

Using a spoon, mix all foods persimmon, strawberries, banana, cashew nuts, raisins, cranberries, yogurt and honey properly.

Using a spoon, mix them properly.

Now, enjoy the dish!

Question 1
Do you remember every ingredient and equipment you use to make this dish? Talk to your partner.

Question 2
Which one did you like most?

Appendix B Lexical Production Scoring Protocol (LPSP-Spoken)

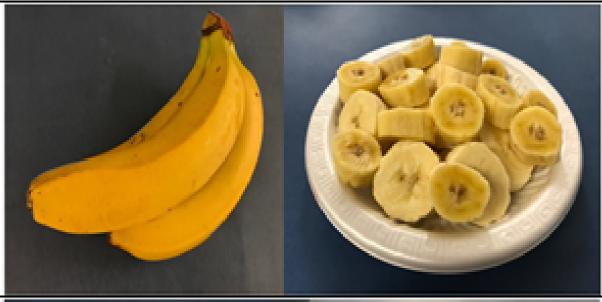
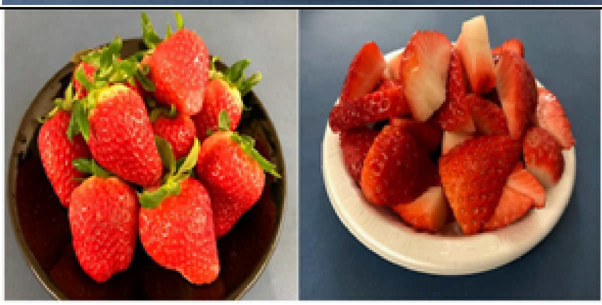
.00 points	.25 points	.50 points	.75 points	1 point
None of word is written;thisincludes: • nothingiswritten • theletterspresent do not meet any "for .25" criteria • Englishwordonly is written	1/4 of word is written; this includes: • any 1 letter is correct • 25-49.9% of the lettersarepresent • correct # of syllables	½ of word is written; this includes: • 25-49.9% of letters correct • 50-74.9% of letters present	3/4 of word is written; this includes: • 50-99.9% of letters correct • 75-99.9% of letters present • 100% letters correct but other letters added	Entire word is written; this includes: • 100% letters correct • 100% letters correctwithaccent addedoromitted

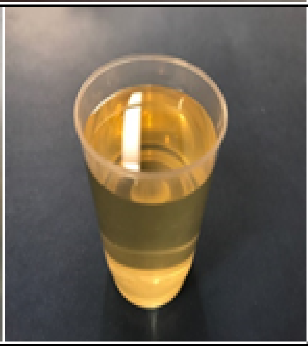
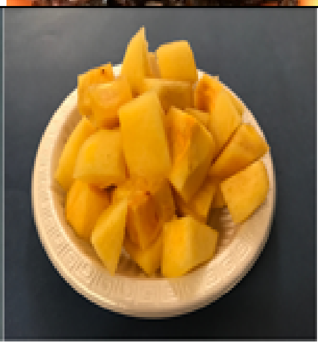
Instructions: (1) "Correct" refers to any letter written and placed in its correct position within a word; "present" refers to any letter written but not placed in its correct position. (2) Determine percentages by dividing letters correct and letters present by the number of letters in the target word. If more letters are written than are in the target word, divide by the large number. (3) If the same target word is written more than once, score it only once in the space where it should be written or, if it is not written in the correct space, score it in the first space where it is written based upon the target word for that space.

Appendix C Productive and Receptive Test Sheets

※ Please look at the image and read extra explanations to write the proper vocabulary item in Korean language.

Name: Age : Sex : Male ( ) Female ( ).

No.	**Photos**	**Extra Info**	**Answer**
1.		-Name	
2.		- Name	
3.		- Name	
4.		-Equipment Name	
5.		- Name	

6.	 	- Name - Not grape	
7.		-Equipment Name	
8.	 	- Name - Sweet - Sticky	
9.	 	- Dried fruit - Name	
10.	 	- Name	

※ *Please match the image with explanation with the proper vocabulary item in the box below.*

If you have no idea, please leave it empty OR fill 'NO' in the blank.

1.	2.	3.	4.	5.
cashew nut	raisin	yogurt	bowl	plate
6.	**7.**	**8.**	**9.**	**10.**
honey	persimmon	banana	strawberry	cranberry

		
		
		
		Good Job! ^ ^

Appendix D Interview Questions

Interviews on English learning through Cooking

1. 요리하는 것 좋아해요 (Do you like cooking?)

- 좋아한다면 왜 요리하는 것이 재미있어요? (? If so, are there any reasons?)

2. 요리하면서 어려운 점이 있었다면 무엇이에요? (What was difficulty you had while you were cooking?)

3. 영어 배우면서 사진만 보고 배우는 것이 재미있어요? 아니면 실제 재료를 가지고 요리하는 것이 재미있어요? (Which one would you prefer between making the dish using real objects and cooking just using photos of them?)

   - 어떤 경우이던 왜 재미있다고 생각해요? (In either case, why do you think it is more fun?)

4. 요리하면서 영어 단어를 배우는 것이 도움이 되었어요? 만약 그랬으면 어떻게 도움을 줬어요? (Does a cooking activity help you learn English words? If so, how did it help you?)

5. 영어 단어 말고 다른 것도 배울 수 있다고 생각해요? (Do you think you can learn other than English words?)

   - 왜 그렇게 생각해요? (Why do you think so?)

Thank you very much for sharing your thoughts with me

A Meta-Analysis of Effectiveness of Teacher and Peer Feedback: Implications for Writing Instructions and Research

Suthathip Thirakunkovit
Mahidol University, Thailand
Bee Chamcharatsri
University of New Mexico, U.S.A.

Bioprofile:

Suthathip Thirakunkovit is currently a lecturer in the English and Applied Linguistics Programs at the Faculty of Liberal Arts, Mahidol University, Thailand. Her research interests broadly focus on language test development, test validation, corpus linguistics and second language writing. **Email:** suthathip.thi@mahidol.edu **(corresponding author)**

Bee Chamcharatsri is Associate Professor in the Department of Language, Literacy, and Sociocultural Studies at the University of New Mexico. He teaches composition courses to international students and TESOL preparation courses to graduate students. His research interests include emotions and second language writing, world Englishes, and writing center. **Email:** bee@unm.edu

Abstract

This meta-analysis investigates the effectiveness of teacher and peer feedback. Over 5,000 articles from 41 journals in the fields of composition studies and second language writing and four electronic databases were identified. After careful screening and applying inclusion and exclusion criteria, only 27 primary studies, yielding 52 effect sizes, were analyzed. The results indicate a Hedges' g effect size of .79 for the effectiveness of writing feedback in general. Feedback from teachers produces a larger effect size (Hedges' g = .90), compared with peer feedback (Hedges' g = .68). However, after examining the studies on peer feedback more closely, the researchers have found that there is a noticeable difference in effect sizes between peer feedback without training (Hedges' g = .60) and peer feedback with prior training (Hedges' g

Mahidol University, Faculty of Liberal Arts, 999 Phuttamonthon 4 Road, Salaya, Nakhon Pathom 73170 Thailand
University of New Mexico, 1 University of New Mexico, Albuquerque, NM, USA 87131

= .74). If specific instructions are given to students during peer training, it could result in higher improvement in the quality of student writing. Therefore, based on the results, the researchers would like to argue that not all peer feedback leads to significant performance improvement. The teachers have to provide specific training during the peer review process in order to enhance the quality and effectiveness of peer feedback. In this regard, peer feedback can have a strong positive effect on improving the quality of student writing.

Keywords: meta-analysis; writing feedback; teacher feedback; peer feedback

Introduction

For students who learn how to write, receiving feedback on their writing is an integral part of their learning. According to second language acquisition research, corrective feedback is helpful for students to notice their errors. One of many controversial conversations on feedback is the source of feedback, whether it should come from teachers or students or from both sources. Students usually believe that teachers provide better feedback because they are considered "experts." However, some students think that teachers sometimes take over or "appropriate" students' ideas and perspectives. Some students enjoy peer feedback because they are given opportunities to learn from one another, whereas others do not because their experiences suggest that grammatical issues will be the focus of the feedback, and they feel that their written texts are being judged by someone who is not a language expert, as opposed to their teacher (Ferris & Hedgcock, 2005; Hyland & Hyland, 2006; Leki & Carson, 1994; Matsuda, 2012). With such differences in the preferences of teacher and peer feedback, this meta-analysis aims to explore the effectiveness of teacher and peer feedback in different classroom contexts.

Literature Review

What is a Meta-Analysis?

Gene Glass (1976) first introduced the term "meta-analysis" into the field of social sciences. This method of study refers to a systematic statistical technique that combines the results from a large collection of different studies in order to derive at an estimate of the effectiveness of a treatment or technique, called *effect size* (Glass, 1976; Goldberg, Russell, & Cook, 2003). The term "*effect size*" refers to a measure that illustrates the average difference between two groups in terms of standard deviation units. According to Cohen's (1988) effect size benchmarks, a Cohen's *d* of .2 is considered small, .5 is considered medium, and .8 is considered large. To illustrate, an effect size of .2 means that the score of the average participants in the experimental group is only .2 standard deviations above the average participants in the non-experimental group. With such a small difference, we can expect plenty of overlap in the performances between participants in the two groups. This means that the new treatment only has a small effect on the participants in the experimental group.

Previous Meta-Analyses on Feedback

Meta-analysis has gained the attention of many scholars in the field of applied linguistics and second language (L2) writing (Biber, Nekrasova, & Horn, 2011; Kang & Han, 2015; Li, 2010; Plonsky & Brown, 2015; Russell & Spada, 2006). In particular, for over a period of 20 years, many meta-analyses have been conducted to provide insights on the issue of written feedback in different classroom contexts. Many of them have paid particular attention to the issue of written feedback on linguistic errors or so-called corrective feedback.

To begin with, Russell and Spada (2006) performed a meta-analysis of 15 studies published in academic journals between 1988 and 2003 to examine the effect of corrective feedback used by teachers in L2 contexts, and they reported a very large effect size of 1.16.

The more recent studies that investigate the effectiveness of corrective feedback include Li (2010), Biber *et al.* (2001), and Kang and Han (2015). Li (2010) examined 33 primary studies and investigated the effects of 11 moderator variables e.g. type of feedback, context of the study, treatment length, mode of delivery, and year of publication to determine whether these variables had significant impacts on writing improvement. His findings showed that there was an overall medium effect of .64 for corrective feedback; however, only research setting (lab, class, and group) showed significant differences in the effects of corrective feedback under both fixed-effects and random-effects models.

Biber *et al.* (2011) examined the effects of different types of corrective feedback on the quality of students' L1 and L2 writing with 23 studies published between 1982 and 2007. This meta-analysis delved into the analysis of 16 moderator variables, including teacher and peer feedback. A medium effect of .53 was found for the effect of feedback in the studies with a treatment/control design and a large effect of .98 in the studies with a pre- and post-test design. For the sources of feedback, the researchers only compared the results of teacher feedback with other sources of feedback i.e. peer, tutor, self, and computer combined. They stated that there was only little or no significant difference between feedback provided by teachers and feedback gained from other sources. However, the results of this meta-analysis must be interpreted with caution because the number of studies in some sub-group moderator analyses were very small i.e. less than five. Therefore, the effectiveness of L2 written corrective feedback could not be ascertained.

Kang and Han (2015) analyzed 21 primary articles and investigated eight moderator variables. Their study showed an overall medium effect size of .68 for the effect of written corrective feedback. Only three moderator variables i.e. research settings, proficiency levels, and genres of writing task showed significant differences between groups. Based on their results, they made a strong claim about written corrective feedback that it was indeed effective in helping L2 students improve grammatical accuracy even though its effectiveness could vary depending on a number of factors, including learners' proficiency, classroom setting, and the genre of writing tasks.

Although the findings of these meta-analyses indicate the effectiveness of corrective feedback in general, the debate about which source of feedback, teachers or peers, is more effective has not been explored in any previous meta-analyses.

Previous Studies on Teacher Corrective Feedback

In the field of L2 writing, there is still an ongoing debate about whether or not corrective feedback from teachers is helpful or a hindrance to student learning (Brown & Larson-Hall, 2012; Ferris, 2010; Hyland & Hyland, 2006). The controversial conversation of feedback began when Truscott (1996) published his article in *Language Learning* and made a strong claim about how corrective feedback was, in fact, ineffective and not useful to students. In some cases, it could even be harmful to student learning because the students might feel pressured when they are notified of the errors; therefore, teachers should stop providing corrective feedback to drafts of student writing. His argument has later sparked many subsequent studies.

While some scholars agree with his argument, Ferris and her co-author concluded that error correction is actually effective (Bitchener & Ferris, 2012; Ferris, 2010; 2011). Ferris (2011), in particular, made an argument that providing corrective feedback on student writing is helpful and effective if teachers know how to provide feedback to students. For example, she emphasized that teachers should focus only on two to three errors at a time, and they should know what types of written feedback students prefer in order to maximize their students' learning processes. Sheen (2007) also pointed out that written corrective feedback that is provided with metalinguistic explanation could be more valuable to students' writing improvement than explicit corrections without any explanation. Moreover, Lee (2011) who conducted a study of teacher perspectives on written feedback in Hong Kong concluded that writing teachers should consider their teaching contexts because, in her study, the context was the significant factor that determined what types of corrective feedback were actually helpful to students.

In conclusion, the literature indicates that there are controversial views on the effectiveness of corrective feedback; however, overall corrective feedback given by the teachers seems to have positive impact on students' writing development if the teachers are aware of the factors that can influence the effectiveness of corrective feedback.

Previous Studies on Peer Feedback

While the majority of students highly value teacher feedback because they believe that their teachers were more competent in the language, and their feedback can provide accurate information on the correctness and appropriateness of their writing (Lee, 2005), peer feedback is widely practiced in many writing classrooms in North American universities (Hyland & Hyland, 2006). A number of writing instructors in first year composition courses in North America use a process-oriented approach in their classrooms, meaning that students are asked to write and revise their papers multiple times. Therefore, in writing classrooms, it is common that students receive written feedback from both teachers and their peers during the writing process.

In the literature, many researchers have supported the use of peer feedback because they have found that peer feedback could lead to better revisions and overall improvement in writing quality (Ferris & Hedgcock, 2005; Hyland & Hyland, 2006; Rollison, 2005; Villamil & Guerrero, 1996; Yu & Lee, 2016). Ferris and Hedgcock (2005) argued that peer feedback provides students a sense of audience in composing a paper. From the reader’s perspective, students can provide feedback on their peer’s written drafts; as Ferris and Hedgcock (2005) have elaborated “mastering the interpretation of text … is a significant step toward achieving membership in the community of readers and experts for whom that text is meaningful” (p. 51). However, in order to make peer feedback helpful and effective, Ferris (2011) suggested that teachers need to train and model for students how to provide feedback to their peers. Instead of asking students to focus only on grammatical issues, teachers have to provide training sessions focusing on different types of written feedback such as explicit or implicit, direct or indirect, global or local feedback, and different types of questions to use on the drafts of their peers. Another value of peer feedback is the non-threatening communication taking place between peers (Rollison, 2005), i.e. the students do not feel that their writing is being judged by someone who is more superior. Students should not feel threatened or experience a loss of face because students are trained to receive and provide constructive feedback on their peer's drafts. On top of that, the peer feedback activity can become a collaborative effort between the writer and feedback provider because they need to negotiate what is written, how it is written, and how it needs to be revised (Villamil & Guerrero, 1996). In the EFL context, Yu and Lee (2016) conducted a peer feedback study with Chinese students. They reported that during the peer feedback activity, students had adopted different roles such as a reader, a writer, a listener, a speaker, which help students develop socio-cognitive development as a writer.

While the effectiveness of teacher and peer feedback have been documented by a number of studies, we have a less than clear idea about how much teacher and peer feedback can improve students' writing when compared with one another. Therefore, the researchers conducted this meta-analysis in order to examine the effectiveness of teacher feedback and peer review and explore to what extent peer feedback can help support teacher feedback. The purpose of this study is to gain more understanding of peer feedback in particular.

The research questions are as follows:

1. How much does feedback help improve students' writing in general?
2. Is the effectiveness of feedback affected by the sources of feedback?
3. Is the effectiveness of feedback affected by research context?
4. Is the effectiveness of feedback affected by the duration of writing course?
5. Are there potential effects of publication bias in the studies included in the analysis?

Methodology

Data Collection

In conducting a meta-analysis, four major analytical steps are generally required: 1) Identifying relevant studies, 2) Determining inclusion or exclusion criteria, 3) Developing a coding scheme; and 4) Selecting a model for computing effect sizes. The following section is an elaboration on each step in conducting the current study.

Identifying Relevant Studies

Forty-one journals in the fields of first and second language writing and four electronic databases (Appendix A.) were used to identify studies that explored the effectiveness of teacher and peer feedback. In order to minimize the effects of publication bias, unpublished studies such as master's theses, Ph.D. dissertations, and conference proceedings were included in the study. The researchers used different keywords and their combinations to identify relevant studies. The keywords include *writing feedback*, *effectiveness of writing feedback*, *the effect of writing feedback*, *teacher feedback*, and *peer feedback*. Using these keywords, more than 5,000 articles were initially identified to be relevant to the effectiveness of feedback. These search results also included more than a hundred of unpublished sources, mainly Master's theses and Ph.D dissertations.

Determining Inclusion or Exclusion Criteria

The inclusion criteria for the meta-analysis are stringent. Only the studies that have the following characteristics are included in the analysis.

The studies that:

1) used a quantitative analysis as a method of investigation;

2) primarily examined the effectiveness of writing feedback;

3) had a valid and reliable method of determining writing quality e.g. writing scores;

4) provided enough information for effect size calculation i.e. the sample size, mean scores, standard deviations, and p-values; and

5) reported the score differences between drafts or between a pre- and post-test are used.

The researchers decided to use the data that came only from matched samples i.e. the participants were measured more than once because the researchers wanted the results of this meta-analysis to reflect the common practice of the majority of writing instructors, in which the students were asked to write multiple drafts.

Developing a Coding Scheme

Once all potential studies were retrieved, a coding scheme was developed. The coding scheme for this study was developed inductively i.e. by reading through more than 30 research studies to identify potential variables on which their research variables could vary. The coding scheme developed for this study includes 27 variables and are categorized into three broad groups: descriptions of the study, statistical procedure, and writing outcome. (Appendix B.)

The researchers worked independently to code the data. One of the researchers coded all the articles, and then rechecked three times to ensure the maximum accuracy. Since most of the variables can be coded in a very straightforward manner i.e. the information of many variables was explicitly stated in the studies, only 20% of the studies were double-coded by the other researcher. All the coding disagreements were resolved during discussion.

Selecting a Model for Computing Effect Sizes

Most meta-analysis studies are based on two statistical models: fixed-effects and random-effects models. With the fixed-effects model, it is assumed that there is one true effect size underlying all studies and that the participants in the studies were randomly sampled from the same population. With the random-effects model, the true effect size for each primary study is assumed to come from different groups of populations (Borenstein, Hedges, Higgins, & Rothstein, 2009). Based on the different assumptions of these two models, the researchers decided to use the random-effects model because when considering the information given in each study, the random-effects model can account for the variations in the data such as different classroom contexts, different class levels, and different language levels.

Data of the Study

From all databases, more than 5,000 relevant articles were identified with 283 studies remaining after applying the inclusion and exclusion criteria. Since the researchers wanted to focus only on studies that allowed the students to do multiple drafts, a total of 27 studies, yielding 52 effect sizes, were finally included in this meta-analysis. Note that many studies contributed more than one effect size because they focused on more than one source of feedback or they had different groups of students in the studies. Any later publication that re-used the same data from previous publication was excluded from the analysis.

Tables 1 and 2 summarize the characteristics of all 27 studies. All studies were published between 1975 and 2014. One study was published between 1971 and 1980, three studies were between 1981 and 1990, four studies were between 1991 and 2000, 14 studies were between 2001 and 2010, and five studies were between 2011 and 2014. Among the 27 studies, 18 studies were published journal articles, seven studies were master's theses or dissertations, one study was a conference proceeding, and one was a book chapter.

Most of the studies recruited adult learners who learned English as a second or foreign language, and eight studies investigated L1 English speakers. Approximately two thirds of the L2 studies were conducted with university students who had either intermediate or advanced English proficiency. In terms of duration of the writing course, most of the studies focused on the courses that lasted one semester or less. Thirteen studies investigated only teacher feedback, 11 studies examined only peer feedback, and three studies compared teacher feedback with peer feedback. In the case of the studies that included both teacher and peer feedback, the researchers will examine each source of feedback separately if it has been found that the participants between two groups were independent.

In terms of the focus of the feedback, more than 70% of the studies involved more than one aspect of writing. Only four studies specifically focused on grammar. Finally, with respect to the mode of feedback, more than half of the 27 studies focused on written feedback. Thirteen studies investigated the effectiveness of both written and oral feedback, and only two studies investigated only oral feedback.

Table 1 Characteristics of Included Studies by Research Context

Study	Document Type	Research Context	School Level	Proficiency Level	Duration of the Writing Course
Berg (1999)	Journal	L2	Language program	Intermediate	1 semester
Berman (1992)	Thesis	L2	College	Advanced	6 weeks
Bitchener (2008)	Journal	L2	Language program	Low/ Intermediate	8 weeks
Bitchener *et al*. (2009)	Journal	L2	College	Intermediate	2 semesters
Bitchener *et al*. (2010)	Journal	L2	College	Advanced	10 weeks
Butcher (2006)	Dissertation	L2	College	Advanced	1 semester
Cho *et al*. (2007)	Journal	L2	College	Not reported	1 semester
Ciftci *et al*. (2012)	Journal	FL	College	Advanced	1 semester
Davis *et al*. (1997)	ERIC	L1	College	Not reported	1 semester
Diab (2011)	Journal	FL	College	Intermediate	1 semester
Hillocks Jr. (1982)	Journal	L1	Middle	Not reported	4 weeks
Huang *et al*. (2009)	Conference proceeding	FL	College	Not reported	1 semester
Kamimura (2006)	Journal	FL	College	Mixed	1 semester
Lange (2011)	Thesis	L1	Middle	Not reported	7 weeks
Liu (2008)	Journal	L2	College	Intermediate	1 semester
Matsumura *et al*. (2002)	Journal	L1	Primary	Not reported	Not reported
McGroarty *et al*. (1997)	Journal	L1+L2	College	Intermediate	1 semester
Patchan (2011)	Dissertation	L1	College	Mixed	1 semester
Riazantseva (2012)	Journal	L2	College	Intermediate	1 semester
Ryoo (2013)	Dissertation	L2	College	Not reported	1 semester
Sayed (2010)	Journal	FL	College	Not reported	8 weeks
Suzuki (2008)	Journal	FL	College	Intermediate	1 semester
Truscott *et al*. (2008)	Journal	FL	College	Intermediate	2 semesters
Villamil *et al*. (1998)	Journal	L2	College	Intermediate	1 semester
Wade (1988)	Dissertation	L1	Primary	Mixed	2 semesters
Wolter (1975)	Dissertation	L1	Primary	Not reported	3 weeks
Yang *et al*. (2006)	Journal	FL	College	Not reported	1 semester

Table 2 Characteristics of Included Studies by Feedback Features

Study	Source of Feedback	Focus of Feedback	Mode of Feedback	Number of Effect Sizes (k)
Berg (1999)*,**	Peer	Global	Written/Oral	4
Berman (1992)	Teacher	Holistic	Not reported	2
Bitchener (2008)	Teacher	Grammar	Written/Oral	3
Bitchener *et al*. (2009)	Teacher	Grammar	Written/Oral	3
Bitchener *et al*. (2010)	Teacher	Grammar	Written/Oral	3
Butcher (2006)*	Teacher/Peer	Holistic	Written	2
Cho *et al*. (2007)*	Teacher	Content	Written	2
Ciftci *et al*. (2012)**	Peer	Organization	Written/Oral	2
Davis *et al*. (1997)	Teacher	Holistic	Written/Oral	2
Diab (2011)**	Peer	Global	Written/Oral	1
Hillocks Jr. (1982)	Teacher	Holistic	Written	1
Huang *et al*. (2009)*	Peer	Holistic	Written/Oral	2
Kamimura (2006)**	Peer	Holistic	Written/Oral	2
Lange (2011)**	Teacher/Peer	Content	Written	2
Liu (2008)	Teacher	Local	Written	2
Matsumura *et al*. (2002)	Teacher	Global	Written	2
McGroarty *et al*. (1997)*	Peer	Holistic	Oral	2
Patchan (2011)**	Peer	Global	Written	4
Riazantseva (2012)	Teacher	Local	Written	1
Ryoo (2013)	Teacher	Holistic	Written/Oral	1
Sayed (2010)*	Peer	Holistic	Written/Oral	2
Suzuki (2008)*	Peer	Holistic	Oral	1
Truscott *et al*. (2008)	Teacher	Grammar	Written	1
Villamil *et al*. (1998)**	Peer	Holistic	Written/Oral	1
Wade (1988)	Teacher	Local	Written	1
Wolter (1975)	Teacher	Global	Written	1
Yang *et al*. (2006)*	Teacher/Peer	Holistic	Written/Oral	2

*Without peer training

**With peer training

The dependent variable of the study is the effect size derived from each primary study, which indicates the magnitude of the effect of teacher or peer feedback in improving students' drafts. In this study, the researchers decided to use Hedge's g instead of Cohen's d for the effect sizes because many studies included in this study contain small sample sizes. According to Borenstein *et al*. (2009), when sample sizes are small (usually below 20), the effect size estimate can be overstated (so-called biased). To correct for this bias, they recommended the researchers to use Hedge's g.

The independent variable i.e. the moderator for the study is the source of feedback. Drawing on the findings from previous studies, the researchers additionally investigated three other moderator variables: research context, duration of the writing course, and publication type. The researchers decided to investigate only these three variables because they are theoretically associated with the effect of feedback, and each variable can be represented by an adequate number of studies that allow the researchers to make comparisons within each variable group. The rationales for investigating these three moderators are discussed below.

Research Context

The contexts of research on writing feedback can generally be divided into three categories: L1, L2, and foreign language (FL). According to Li (2010), the dynamics of learning environments can be different in these contexts, so it is possible that the effects of feedback may differ. Therefore, the researchers believe that it is worthwhile to also investigate the context of research. Even though this meta-analysis put a lot of emphasis on feedback in L2 and FL contexts due to a large number of studies conducted in L2 and FL contexts, the researchers decided to also include studies in L1 contexts in order to see whether there are any differences between L1-English and non-L1-English contexts.

Duration of the Writing Course

According to Norris and Ortega (2000), length of instruction is one of the variables that had significant impact on the magnitude of the effectiveness of L2 instructions. Their meta-analysis found that longer lengths of instructions could lead to larger effects. Therefore, it is noteworthy to explore whether the duration of writing training has any significant influence on the effectiveness of feedback.

Publication Type

It is generally believed that academic research with significant findings is more likely to be published. This phenomenon can be called publication bias. Therefore, this meta-analysis also includes data from unpublished sources in order to examine whether there is a bias against studies without significant findings.

Data Analyses

The data were analyzed using professional meta-analysis software called CMA, Version 3. The differences of writing scores of 52 matched groups were analyzed under the random-effects model. According to Borenstein *et al.* (2009), when a researcher works with the data that come from independent groups, the natural unit of an effect size is the regular standard deviation. However, if a researcher works with matched groups or dependent samples i.e. when the same participants were measured more than once, the resulting effect sizes are dependent and the natural unit of an effect size then becomes the standard deviation of the score differences between drafts or between pre-test and post-test. In this case, the correlations between matched groups are required to compute the standard deviation within groups. However, none of the studies included in the analysis reported correlation coefficients between drafts or between the pre-test and post-test. In this study, the researchers, thus, decided to use the correlation of .5 in the analysis because, in actual practices, we could expect a moderate non-bias level between two drafts of writing.

Results of the Study

In order to investigate how much feedback helps improve students' writing in general, the overall effect size of writing feedback was estimated. Additionally, the heterogeneity across all effect sizes was checked by calculating the Q statistics, which determines whether variation across effect sizes is greater than expected from sampling error. A significant value of Q statistics would indicate that the effect sizes are not homogeneous. If the heterogeneity test is shown to be significant, then the moderator analyses are required.

How Much does Feedback Help Improve Students' Writing in General?

Based on a sample size of 27 studies, yielding 52 effect sizes, writing feedback appears to have a moderate to large effect on students' writing improvement in general (Hedge's = .79; SE = .08; C.I. between .64 and .95). Moreover, there is a significant heterogeneity across effect sizes (Q = 282.36; $p < .000$). The I^2 statistics[1] is 81.94%, and the Tau Squared (τ^2) is .25. These suggest that the observed heterogeneity is due to actual differences among studies rather than random error. Table 3 reports the mean effect size that each study contributes.

Table 3 Effects of Overall Writing Feedback

Study	Mean Hedges' g	SE	95% Confidence Interval	
			Lower Limit	Upper Limit
Berg (1999)	0.17	0.29	-0.40	0.74
	0.62	0.35	-0.05	1.30
	1.63	0.44	0.76	2.49
	1.33	0.40	0.56	2.11
Berman (1992)	0.78	0.26	0.26	1.29
	0.07	0.24	-0.41	0.54
Bitchener (2008)	0.86	0.28	0.31	1.42
	1.25	0.31	0.63	1.87
	1.38	0.31	0.73	1.94
Bitchener *et al*. (2009)	2.52	0.57	1.41	3.63
	1.64	0.42	0.81	2.47
	1.07	0.35	0.39	1.75
Bitchener *et al*. (2010)	0.77	0.33	0.12	2.34
	0.34	0.20	-0.05	0.73
	1.02	0.36	0.32	1.72
Butcher (2006)	0.71	0.32	0.07	1.34
	0.75	0.26	0.24	1.26
Cho *et al*. (2007)	0.60	0.36	-0.11	1.31
	0.12	0.32	-0.50	0.74
Ciftci *et al*. (2012)	0.46	0.27	-0.07	1.00
	1.51	0.38	0.77	2.25
Davis *et al*. (1997)	1.86	0.37	1.14	2.58
	2.47	0.45	1.59	3.35

[1] The I^2 statistics indicates the percentage of variation across studies that occurs due to heterogeneity rather than chance.

Diab (2011)	0.70	0.24	0.24	1.17
Hillocks Jr. (1982)	0.85	0.14	0.57	1.13
Huang *et al*.(2009)	0.01	0.14	-0.26	0.28
	-1.18	0.18	-1.54	-0.83
Kamimura (2006)	0.79	0.33	0.14	1.43
	1.35	0.40	0.57	2.14
Lange (2011)	0.45	0.24	-0.03	0.92
	1.57	0.34	0.90	2.24
Liu (2008)	2.34	0.79	0.79	3.88
	1.60	0.62	0.39	2.81
Matsumura *et al*. (2002)	0.85	0.18	0.50	1.20
	0.31	0.15	0.01	0.61
McGroarty *et al*. (1997)	0.30	0.13	0.05	0.54
	0.33	0.12	0.09	0.57
Patchan (2011)	0.69	0.16	0.37	1.01
	0.46	0.16	0.16	0.76
	0.46	0.15	0.16	0.75
	0.46	0.15	0.16	0.76
Riazantseva (2012)	0.59	0.19	0.22	0.97
Ryoo (2013)	1.00	0.22	0.56	1.44
Sayed (2010)	2.10	0.57	0.99	3.21
	3.93	0.93	2.10	5.76
Suzuki (2008)	1.73	0.46	0.84	2.63
Truscott *et al*. (2008)	1.00	0.27	0.48	1.53
Villamil *et al*. (1998)	0.42	0.39	-0.35	1.19
Wade (1988)	0.02	0.26	-0.49	0.52
Wolter (1975)	0.98	0.43	0.14	1.82
Yang *et al*. (2006)	0.41	0.16	0.10	0.73
	0.75	0.18	0.39	1.11
Random-effects model	0.79	0.08	0.64	0.95

Moderator Analyses

When the heterogeneity test or Q statistics is shown to be significant, the researchers are required to perform moderator analyses. In this study, moderator analyses were performed on the effect sizes associated with four independent variables: the source of feedback, research context, duration of writing course, and publication type.

Is the Effectiveness Affected by the Sources of Feedback: Teacher and Peer Feedback?

The first moderator variable analyzed was the source of feedback. As evident in Table 4, feedback from teachers produces a large effect size (Hedges' g = .90; SE = .10; 95% C.I. between .70 and 1.09). Feedback from peer review produces a smaller effect size (Hedges' g = .68; SE = .12, 95% C.I. between .45 and .92). Based on the results, it is apparent that teacher feedback could improve students' writing quality more than peer feedback. However, the reviews of peer feedback studies specifically showed that providing specific peer training to students resulted in higher improvement in their writing (Hedges' g = .74) versus without peer training (Hedges' g = .60). A 95% confidence interval is often interpreted as a range within which we can be 95% certain that the true effect lies within the specified range. If we look at the confidence interval of the effect of peer feedback in which the students did not receive peer review training before the peer review process, it has a much wider range than the range of the studies with prior peer training. It can be interpreted that the distribution of the effectiveness of peer feedback without training is less certain when compared with peer feedback with training. Moreover, with an attempt to investigate the effect of peer feedback without peer training further, the researchers have found that there is one study that seems to inflate the true picture of effect sizes. From Table 3, it can be seen that an effect size from Sayed's study (Hedges' g = 3.93) is much larger than others'. If this study is removed from the analysis, the effect size of peer feedback without prior training comes down to .50 (SE = .22; 95% C.I. between .02 and .90). Since the sample size in each variable category is relatively small, an outlier or extreme value can affect the validity of meta-analysis results. For this case, the researchers decided that it might be best to delete the effect size that is considered an outlier and re-calculate the average effect size for the rest of the studies.

Is the Effectiveness of Feedback Affected by Research Context?

Significant differences in effect sizes were not found among the studies conducted in the L1, L2, and FL contexts, meaning that the contexts where the students learn how to write do not have a significant impact on the effectiveness of feedback ($F = .33$; $p = .72$).

Table 4 Mean Effect Sizes for the Source of Feedback

Study	k	Mean Hedges' g	SE	95% Confidence Interval	
				Lower Limit	Upper Limit
Teacher Feedback	26	.90	.10	.70	1.09
Peer Feedback	26	.68	.12	.45	.92
- With prior training	14	.74	.11	.53	.94
- Without prior training	12	.60	.23	.15	1.06

Table 5 Mean Effect Sizes for Research Context

Study	k	Mean Hedges' g	SE	95% Confidence Interval	
				Lower Limit	Upper Limit
L1	15	.68	.10	.47	.88
L2	24	.85	.10	.65	.06
FL	13	.86	.25	.38	.34

Is the Effectiveness of Feedback Affected by the Duration of Writing Course?

The results revealed that studies with two semesters provided the largest mean effect (Hedges' g = 1.15) when compared with those with shorter time periods. Therefore, the overall results of this study can support the arguments made by Norris and Ortega (2000) in that duration of instruction can have significant impact on the magnitude of the effectiveness of language improvement.

However, finding from this moderator analysis indicated that intensive writing courses could result in higher effectiveness of writing feedback than traditional 15-week semester length courses. By looking at the effect size of the writing courses that were shorter than a semester, we can see that it was significantly higher than the one-semester courses ($F = 84.50$; $p < .00$). Overall, the researchers have found that many universities and institutions now offer intensive writing courses that are between four and eight weeks. In particular, these courses are becoming more and more common for international students who come to study in English-speaking countries. However, it is still unclear how duration of a writing course could have an impact on the effectiveness of writing feedback. Because of limited information was provided in the published studies, the researchers could not determine whether or not the students in those studies had more class time per week than in a traditional format. Therefore, further study is necessary to examine the characteristics of intensive writing courses in general.

Table 6 Effects for the Duration of Writing Course

Study	k	Mean Hedges' g	SE	95% Confidence Interval	
				Lower Limit	Upper Limit
Less than a semester	14	.94	.02	.65	1.22
One semester	33	.69	.10	.50	.88
Two semesters	5	1.15	.37	.43	1.90

Are there Potential Effects of Publication Bias in the Studies Included in the Analysis?

Finally, the effect of publication types was examined as a potential moderator variable. With regard to how publication type can vary in effect, the researchers have observed differences in effect sizes from the studies that were published and those that were not. According to Table 7, studies that were published showed a larger effect size (Hedges' g = .86) than those that were not published (Hedges' g = .62), and the 95% C.I. of the studies that were published was also narrower.

Table 7 Effects for Publication Type

Study	k	Mean Hedges' g	SE	95% Confidence Interval	
				Lower limit	Upper limit
Published	35	.86	.08	.70	1.03
Unpublished	17	.62	.17	.30	.94

Funnel plots are used as a visual tool to detect bias in publication. According to Sterne and Harbord (2004), in the absence of publication bias, the distribution of the effect sizes should scatter widely and evenly at the bottom of the plot and spread narrowly at the top. However, publication bias in a meta-analysis can be detected when the funnel plot looks asymmetrical.

Based on Figure 1, it can be interpreted that studies with statistically significant results are more likely to be published. The interpretation of the funnel plot is as follows:

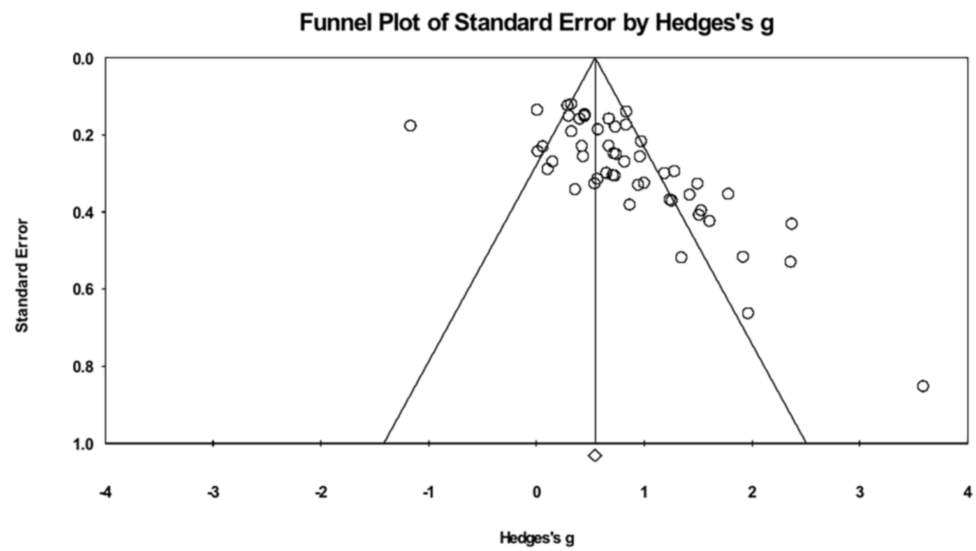


Figure 1 Funnel Plot

As shown in Figure 1, the funnel plot showed clear asymmetry with missing studies at the bottom of the left side of the plot. Most studies were clustered around the means of the upper part of the plot. This means that there were more studies with positive effect sizes on the right side of the funnel plot than on the left side. With regard to whether there is any evidence of publication bias, as the plot showed, there are more studies with significant positive effects. This means that studies with significant feedback effects were more likely to be published. Moreover, the researchers observed publications with very large effect sizes on the right side. It is important to note that a number of studies on the left side of the plot can also be seen. Therefore, the amount of bias in publication was not substantial. Considering the missing data at the bottom of the plot, the researchers can interpret that this meta-analysis consisted of several studies with large or sufficient sample sizes in the data.

Discussion and Conclusions

Summary of the Study

The present meta-analysis has demonstrated that overall both teacher and peer feedback are valuable aids in improving the quality of student writing. Even though the results of this meta-analysis comes as a confirmation to most educators and researchers that feedback from teachers could produce a larger effect size (Hedges' g = .90) than feedback from peers (Hedges' g = .68), there are some points about peer feedback that should be discussed further. Because students know that the teachers will grade their papers, they will simply revise their papers based on teachers' comments. This will result in the undermining of peer feedback. In fact, students can provide critical feedback to others' writing, and its effect can be as nearly as effective as teacher feedback if the students are trained before giving feedback to their peers. As reported in this study, peer feedback can be effective if students are appropriately trained to provide feedback on their peers' drafts. To illustrate this point, the groups of students who received pre training had greater improvement in revising their subsequent drafts (Hedges' g = .74). Based on the findings, the researchers ultimately want to provide some pedagogical implications for giving effective peer review training to the students and giving recommendations for future research directions.

Pedagogical Implications

There is no doubt that providing feedback is a time-consuming activity, and oftentimes teachers have too much paperwork to edit and evaluate. However, the results of this meta-analysis showed that with training, guidance and practice, students can be critical readers and provide useful feedback to their peers. Strong arguments for the use of peer response can mostly be found in these two stances: writing process and collaborative learning.

In the process of writing, peer review should be planned and guided carefully, so that it can help students become better writers and readers. Before the peer review starts, as suggested by Hyland (2001), the teachers need to emphasize the role of the peer-reviewer that he/she is a critical reader, not an evaluator. In this regard, the students will feel more comfortable, and peer review will be seen as an opportunity for them to learn how to write for an audience.

Before students are required to review their classmates' writing, they should be taught how to develop necessary skills for peer review. Given the time that is required for peer-review training sessions, the teachers should make sure that students can do their first practice of peer review during class, along with the guidance and suggestions of the instructor (Ferris & Hedgcock, 2005). It may take more than one class session to train students on how to provide feedback; instructors need to be patient in this process. Importantly, during peer training, teachers should provide the students with checklists or peer evaluation forms focusing on different aspects of writing (Diab, 2011; Kamimura, 2006). The checklists can be created by teachers or can be collaboratively created with the students. The students and teachers can use a few examples of student writing to create the lists. These lists can be a powerful tool for students to have during the peer review activities because they can provide a "roadmap" to successful assignment completion and help clarify the features of good writing.

During training, the students should also be taught about what types of feedback on content, organization, or sentence level structures are helpful and useful for their peers. Moreover, instead of providing only grammatical corrections in the drafts, teachers should encourage students to provide questions and comments by modeling it to the whole class. Teachers need to remind their students that their comments should be constructive and complete. For example, a peer-reviewer might say, "I found these explanations very clear, but I do not understand how they relate to your topic sentence. Could you make the connection by providing examples?" Teachers should also evaluate the feedback students give to their peers to make sure that constructive feedback is provided. For example, the peer-reviewer can indicate which sentences or paragraphs that seem out of order and in need of revision.

Moreover, based on the suggestions by Butcher (2006), Hyland and Hyland (2006), Rollinson, (2005), and Villamil and Gurrero (1998), peer review can promote the development of the collaborative learning process because peer review sessions allow the students to discuss and explain what the feedback means and how to effectively revise their papers. In other words, they learn how to approach a variety of feedback with an open mind. As many studies about peer feedback have suggested (Berg, 1999; Bitchener & Ferris, 2012; Butcher, 2006), teachers may organize peer feedback as a classroom collaborative activity. Teachers may pair up students who have stronger writing skills or students with more experience in writing with less skilled writers to facilitate the revision process. Regardless of language proficiency, students should have access to learning development opportunities through peer feedback activities as 'stronger' students may still require feedback directed at idea progression or logical organization. Beyond these aspects, peer questioning can further aid 'stronger' students in developing their drafts. In a similar vein, 'weaker' students can receive guidance with the more fundamental skills of writing, such as idea planning and outlining from their 'stronger' peers. Even though peer assistance can be a good technique, the teachers have to be able to find a balance between using strong students as helpers and enriching their learning and, at the same time, still allowing them to sustain the quality of their writing. Above all, through working collaboratively in evaluating each other's draft, the students had opportunities to interact and assist one another by making feedback comprehensible and providing social and affective support (Hyland & Hyland, 2006).

Through a peer training activity, the students can also develop their critical perspectives as readers in reading and providing feedback to written texts. Let the students know that feedback is not just about editing someone's text; it is about providing feedback as readers. By providing feedback from a reader's perspective, the comments can be inquisitive, insightful, and honest. At the same time, the students also learn about audience in writing that they do not simply write the papers for their instructors. In this regard, peer feedback can help the students to develop their autonomy and self-confidence as writers.

Recommendations for Future Research

While this research study reported and confirmed the results of previous teacher and peer feedback studies, the researchers believe that there is still room for more research on feedback. To be more specific, it has been found that there is a lack in quantitative studies that focus on the use of peer feedback in L1 writing. Moreover, when conducting this study, the researchers were unable to identify or locate all necessary descriptive statistical data in the studies that were being reviewed. By omitting such data in their reports, other readers or researchers may not be able to confirm or recalculate the statistical analyses. Therefore, the researchers would like to encourage future researchers to include all necessary descriptive statistics including means, standard deviations, and the number of students in the study, as well as the correlations between students' drafts. In doing so, the results regarding the effectiveness of writing feedback could be made more accurately and generalized to wider contexts. As Liu and Brown (2015) indicated, research methodology used in conducting research on written feedback was not carefully designed or reported in areas of individual differences of learners in our classrooms, written tasks objectives, and reporting of statistical analyses in research studies.

Acknowledgments

We would like to express our deep gratitude to Professor Antony Silva, Professor Luke Plonsky, and Dr. Todd Ruecker for their valuable input throughout the whole process of conducting this project.

References

*Berg, E. C. (1999). The effects of trained peer response on ESL students' revision types and writing quality. *Journal of Second Language Writing*, *8*(3), 215-241.

*Berman, R. (1992). *Effects of teachers' written comments on the learning by advanced ESL students of discourse in second language writing*. M.A. thesis, University of Toronto.

Biber, D., Nekrasova, T., & Horn, B. (2011). The effectiveness of feedback for L1-English and L2-writing development: A metaanalysis. *ETS Research Report Series*, *2011*(1), i-99.

*Bitchener, J. (2008). Evidence in support of written corrective feedback. *Journal of Second Language Writing*, *17*(2), 102-118.

*Bitchener, J., & Knoch, U. (2009). The contribution of written corrective feedback to language development: A ten month investigation. *Applied Linguistics*, *31*(2), 193-214.

*Bitchener, J., & Knoch, U. (2010). Raising the linguistic accuracy level of advanced L2 writers with written corrective feedback. *Journal of Second Language Writing, 19*(4), 207-217.

Bitchener, J., & Ferris, D. R. (2012). *Written Corrective Feedback in Second Language Acquisition and Writing.* New York: Routledge.

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). *Introduction to Meta-Analysis*. Chichester: Wiley.

Brown, S., & Larson-Hall, J. (2012). *Second Language Acquisition Myths: Applying Second Language Research to Classroom Teaching*. Ann Arbor: University of Michigan Press.

*Butcher, K. F. (2006). *The efficacy of peer review in improving ESL students' online writing* (Unpublished doctoral dissertation). University of New Orleans, New Orleans.

Chen, T. (2014, August). *An exploratory quantitative meta-analysis on peer feedback research in the ESL/EFL writing classrooms*. Paper presented at the International Association for Applied Linguistics (AILA), Brisbane, Australia.

Chen, T., & Lin, H. (2012, March). Effects of peer feedback on EFL/ESL writing improvement: A meta-analysis. Paper presented at the Georgetown University Roundtable on Linguistics and Languages (GURT), Georgetown University, Washington DC.

*Cho, K., & Schunn, C. D. (2007). Scaffolded writing and rewriting in the discipline: A web-based reciprocal peer review system. *Computers & Education*, *48*(3), 409-426.

*Ciftci, H., & Kocoglu, Z. (2012). Effects of peer e-feedback on Turkish EFL students' writing performance. *Journal of Educational Computing Research*, *46*(1), 61-84.

Cohen, J. (1988). Statistical Power Analysis for the Behavioural Sciences. New Jersey: Erlbaum.

*Davis, W., & Fulton, J. (1997). *The effects of professors' feedback on the growth of students' overall writing quality in two college freshman English courses*. Retrieved from ERIC database. (ED414570)

*Diab, N. M. (2011). Assessing the relationship between different types of student feedback and the quality of revised writing. *Assessing Writing*, *16*(4), 274-292.

Evans, N. W., Hartshorn, K. J., McCollum, R. M., & Wolfersberger, M. (2010). Contextualizing corrective feedback in second language writing pedagogy. *Language Teaching Research*, *14*(4), 445-463.

Ferris, D., & Hedgcock, J. S. (2005). *Teaching ESL Composition: Purpose, Process, and Practice* (2nd ed.). New York: Routledge.

Ferris, D., & Roberts, B. (2001). Error feedback in L2 writing classes: How explicit does it need to be?. *Journal of Second Language Writing*, *10*(3), 161-184.

Ferris, D. (2010). Second language writing research and written corrective feedback in SLA: Intersections and practical applications. *Studies in Second Language Acquisition*, *32*(2), 181-201.

Ferris, D. (2011). *Treatment of Error in Second Language Student Writing*. Ann Arbor: University of Michigan Press.

Glass, G. V. (1976). Primary, secondary, and meta-analysis of research. *Educational Researcher*, *5*(10), 3-8.

Goldberg, A., Russell, M., & Cook, A. (2003). The effect of computers on student writing: A meta-analysis of studies from 1992 to 2002. *The Journal of Technology, Learning and Assessment*, *2*(1), 1-52.

Goldstein, L. M. (2005). *Teacher Written Commentary in Second Language Writing Classrooms*. Ann Arbor: University of Michigan Press.

Hajimohammadi, R., & Mukundan, J. (2011). Impact of self-correction on extrovert and introvert students in EFL writing progress. *English Language Teaching*, *4*(2), 161-168.

*Hillocks Jr, G. (1982). The interaction of instruction, teacher comment, and revision in teaching the composing process. *Research in the Teaching of English, 16*(3), 261-278.

*Huang, S. Y., & Chang, Y. (2009). A mediation effect of computer-mediated communication on EFL writing. Proceedings from *World Conference on Educational Media and Technology*. 2009(1), 2588-2597.

Hyland, F. (2001). Providing effective support: Investigating feedback to distance language learners. *Open Learning*, *16*(3), 233-247.

Hyland, K., & Hyland, F. (2006). Feedback on second language students' writing. *Language Teaching*, *39*(2), 83-101.

Jacobs, G., & Zhang, S. (1989). Peer feedback in second language writing instruction: boon or bane?. Retrieved from ERIC database. (ED306766)

*Kamimura, T. (2006). Effects of peer feedback on EFL student writers at different levels of English proficiency: A Japanese context. *TESL Canada Journal*, *23*(2), 12-39.

Kang, E., & Han, Z. (2015). The efficacy of written corrective feedback in improving L2 written accuracy: A meta-analysis. *The Modern Language Journal*, *99*(1), 1-18.

*Lange, K. (2011). *Scientific explanations peer feedback or teacher feedback* (Unpublished doctoral dissertation. Arizona State University, Tempe.

Lee, I. (2005). Error correction in the L2 writing classroom: What do students think? *TESOL Canada Journal, 22*(2), 1-16.

Lee, I. (2011). Feedback revolution: What gets in the way?. *ELT Journal*, *65*(1), 1-12.

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

Li, S. (2010). The effectiveness of corrective feedback in SLA: A meta-analysis. *Language Learning*, *60*(2), 309-365.

*Liu, Y. (2008). The effects of error feedback in second language writing. *Arizona Working Papers in SLA & Teaching*, *15*, 65-79.

Liu, Q., & Brown, D. (2015). Methodological synthesis of research on the effectiveness of corrective feedback in L2 writing. *Journal of Second Language Writing*, *30*, 66-81.

Matsuda, P. K. (2003). Second language writing in the twentieth century: A situated historical perspective. *Exploring the Dynamics of Second Language Writing*, *1*, 15-34.

Matsuda, P. K., Cox, M., Jordan, J., & Ortmeier-Hooper, C. (Eds.). (2006). *Second-Language Writing in the Composition Classroom: A critical Sourcebook*. New York: Macmillan.

Matsuda, P. K. (2012). Let's face it: Language issues and the writing program administrator. *WPA: Writing Program Administration, 36*(1), 141-163.

Matsumura, S., & Hann, G. (2004). Computer anxiety and students' preferred feedback methods in EFL writing. *The Modern Language Journal*, *88*(3), 403-415.

*Matsumura, L. C., Patthey-Chavez, G. G., Valdés, R., & Garnier, H. (2002). Teacher feedback, writing assignment quality, and third-grade students' revision in lower-and-higher-achieving urban schools. *The Elementary School Journal*, *103*(1), 3-25.

*McGroarty, M. E., & Zhu, W. (1997). Triangulation in classroom research: A study of peer revision. *Language Learning*, *47*(1), 1-43.

Mustafa, R. F. (2012). Feedback on the feedback: Sociocultural interpretation of Saudi ESL learners' opinions about writing feedback. *English Language Teaching, 5*(3), 3-15.

Norris, J. M., & Ortega, L. (2000). Effectiveness of L2 instruction: A research synthesis and quantitative meta-analysis. *Language Learning*, *50*(3), 417-528.

*Patchan, M. M. (2011). *Peer review of writing: Learning from revision using peer feedback and reviewing peers' texts* (Unpublished doctoral dissertation). University of Pittsburgh, Pittsburgh.

Plonsky, L., & Brown, D. (2015). Domain definition and search techniques in meta-analyses of L2 research (Or why 18 meta-analyses of feedback have different results). *Second Language Research*, *31*(2), 267-278.

Rahimi, M. (2009). The role of teacher's corrective feedback in improving Iranian EFL learners' writing accuracy over time: Is learner's mother tongue relevant?. *Reading and Writing*, *22*(2), 219-243.

*Riazantseva, A. (2012). Outcome measure of L2 writing as a mediator of the effects of corrective feedback on students' ability to write accurately. *System*, *40*(3), 421-430.

Rollinson, P. (2005). Using peer feedback in the ESL writing class. *ELT Journal*, *59*(1), 23-30.

Russell, J., & Spada, N. (2006). The effectiveness of corrective feedback for second language acquisition: A meta-analysis of the research. In J. Norris & L. Ortega (Eds.), Synthesizing Research on Language Learning and Teaching (pp. 131–164). Amsterdam: Benjamins.

*Ryoo, S. M. (2013). *Responding to non-native writers of English: The relationship between a teacher's written comments and improvement in second language writing* (Unpublished doctoral dissertation). Temple University, Philadelphia.

*Sayed, O. H. (2010). Developing business management students' persuasive writing through blog-based peer-feedback. *English Language Teaching*, *3*(3), 54-66.

Sheen, Y. (2007). The effect of focused written corrective feedback and language aptitude on ESL learners' acquisition of articles. *TESOL Quarterly*, *41*(2), 255-283.

Sterne, J. A., & Harbord, R. M. (2004). Funnel plots in meta-analysis. *Stata Journal*, *4*, 127-141.

*Suzuki, M. (2008). Japanese learners' self revisions and peer revisions of their written compositions in English. *TESOL Quarterly*, *42*(2), 209-233.

Truscott, J. (2004). The effectiveness of grammar instruction: Analysis of a meta-analysis. *English Teaching & Learning*, *28*(3), 17-29.

*Truscott, J., & Hsu, A. Y. P. (2008). Error correction, revision, and learning. *Journal of Second Language Writing*, *17*(4), 292-305.

*Villamil, O. S., & De Guerrero, M. C. (1998). Assessing the impact of peer revision on L2 writing. *Applied Linguistics*, *19*(4), 491-514.

*Wade, L. K. (1988). *An analysis of the effects of a peer feedback procedure on the writing behavior of sixth grade students*. Unpublished doctoral dissertation, University of Kansas.

*Wolter, D. R. (1975). *Effect of feedback on performance on a creative writing task* (Unpublished doctoral dissertation). University of Michigan, Ann Arbor.

*Yang, M., Badger, R., & Yu, Z. (2006). A comparative study of peer and teacher feedback in a Chinese EFL writing class. *Journal of Second Language Writing*, *15*(3), 179-200.

Yu, S., & Lee, I. (2016). Exploring Chinese students' strategy use in a cooperative peer feedback writing group. *System*, 58, 1-11.

The articles marked with asterisks (*) are those included in the analysis.

Appendix I: The List of the Journals and Databases

Applied Linguistics
Assessing Writing
Asian Journal of English Language Teaching
British Journal of Educational Technology
CALICO Journal
Computer Assisted Language Learning
College English
College Composition and Communication (NCTE)
Computers and Composition
ELT Journal
Electronic Journal of Foreign Language Teaching
English Language Teaching
Foreign Language Annals
IRAL: International Review of Applied Linguistics in Language Teaching
Issues in Writing
Journal of Basic Writing
Journal of Computers in Education
Journal of English for Academic Purposes
Journal of Educational Computing Research
Journal of Educational Psychology
Journal of Educational Research
Journal of Second Language Writing
Language Assessment Quarterly
Language Learning Language Teaching
Language Teaching Research
ReCALL
Research in the Teaching of English
Second Language Research
Spaan Fellow Working Papers
Studies in Second Language Acquisition
System
TESL Canada Journal
TESL-EJ Teaching English as a Second or Foreign Language
TESOL Quarterly
TESOL Journal
The Elementary School Journal
The Canadian Modern Language Review
The English Journal
The INTESOL Journal
The Modern Language Journal
Written Communication

Appendix II. Coding Scheme

Variables	Descriptors
Descriptions of the Study	
Document type	Journal article/ Book chapter/ Institutional report/ Conference proceeding/ Dissertation or master's thesis/ Unpublished studies
Period of data collection	Known/ Unknown
Research context	L1/ L2/ FL/ Mixed between contexts
School level	Primary school/ Middle school/ High school/ College/ Language program/ Other
Writing courses	First-year composition/ Business English/ Writing for Specific Purposes/ Creative writing/ Other
Proficiency level	Beginner/ Intermediate/ Advanced/ Not reported
Length of writing task	Paragraph/ Essay/ Portfolio/ Not reported
Genre of writing task	Narrative/ Argumentative/ email/ Academic research/ Other
Research method	Quantitative method/ Mixed method
Length of the study	Less than one semester/ One semester/ Two semesters/ More than a year/ not reported
Statistical Analysis Procedure	
Research design	Pre- and post-test method/ Draft/re-draft design/ Experimental/ Quasi-experimental
Statistical procedure	Descriptive statistics/ Paired t-test/ ANOVA/ Correlation
Sources of feedback	Teacher feedback/ Peer feedback/ Tutor/ Self-evaluation/ Computer-generated program/ Not reported

Feedback focus	Holistic/ Analytic
Feedback mode	Written/ Oral/ Both written and oral feedback
Writing Outcome	
Outcome measure of writing development	The number of errors reduced/ Writing scores

Improving Foreign Language Teaching: Towards a Research-Based Curriculum and Pedagogy

E. Macaro, S. Graham, and R. Woore, London: Routledge, 2015, 216 pp., $46.95, ISBN 978 1138779501

Pelin Irgin

University of Reading, UK

Bioprofile:

Pelin Irgin received her doctorate degree in English Language Teaching Program from Hacettepe University, Ankara, Turkey. As an academic visitor, she followed her PhD research on exploring the roles of strategy instruction in a blended learning environment at Institute of Education, University of Reading, Reading, UK. She has taught undergraduate courses in teacher education, applied linguistics and foreign language teaching methodology. Her current research areas are TEYL, EFL, listening in L1 and L2, and strategy instruction. **Email:** pelinirgin@yahoo.com

The book under review aims to discuss new trends on language skills and communication competences and present a system for language assessment and its reflection in language learning process. The book is divided into nine different chapters on debates. Chapter 1, “The problem with improving language learning”, introduced three main dilemmas: first, to what extent L2 learning situation should be reproduced out of the classroom environment; second, to what extent foreign language teachers should motivate their students to be more responsible for their own learning; and lastly, how teachers should decide on the target foreign languages to teach in the school setting formally and eight principles dealing with the research and education setting rooted in both UK educational concerns and international agenda. This chapter describes the reasons of the PDCinMFL (The Professional Development Consortium in Modern Foreign Languages) project such as reforming the principles of professional development by relating to L2 language skills and highlighting key features of pedagogical approach embodied in the principles, and how the following chapters are linked to the framework presented step by step throughout the book.

Chapter 2, “Oral interaction in the second language classroom”, considers ways of developing communicative competence of the students in the classroom environment and

University of Reading, Whiteknights, PO Box 217, Reading, Berkshire, RG6 6AH, UK

aspects of designing practice-focused interaction in pedagogical setting rather than waiting for the progress in students' oral fluency as a natural-endowed skill. This chapter is well illustrated with a number of cases in classroom interaction including teacher talks, student comments on the use of communication strategies and fluency development. Chapter 3, "Developing reading and listening" reviews how to develop learners' linguistics and strategic knowledge in both receptive skills to improve their proficiency. The chapter offers convincing evidence supporting a balanced approach to make learners become strategic readers and listeners.

Unlike the case presented in Chapter 3, the argument proposed in Chapter 4, "Developing the learner through writing", is that there is not an agreement about the purpose of writing in a foreign language classroom. This lack of consensus is described in detail which might motivate researchers study on L2 writing in their own contexts. This chapter presents a model for L2 writing and strategic writing, and emphasizes the uncertainty on what the reasonable level of creativity in L2 writing is. Moreover, the chapter provides a case study of an L2 writer which is outlined clearly.

Chapter 5, "A pedagogical assessment system", discusses foreign language assessment system. The chapter aims to present Main Development Strands (MDS) urging teachers to start with a holistic judgement based on weeks ending up with an assessment reference point during the period. In essence, some problematic sides of L2 assessment framework used in pedagogical setting are discussed by comparing the previous National Curriculum of the UK to the alternative assessment system presented by the authors. This presents a curriculum and an alternative L2 assessment framework which proposes five levels for all skills: beginner level, experimenter level, developing level, competent level, skilled level, and five levels are split into two designated as "just meeting the level" (e.g. Beginner Level 1) and "comfortably meeting the level" (e.g. Beginner Level 2). Chapter 6, "Preparing to become and continuing to be a language teacher", deals with barriers and handicaps teachers and student teachers should overcome, and the possible ways they should develop teaching practices. This chapter defines teachers' beliefs and explains to what extent they resist the change in professional setting. Also, the chapter discusses developing the teachers' professional identity and overcoming barriers which are drawbacks in their professional development.

Chapter 7, "The PDCinMFL project" (The Professional Development Consortium in Modern Foreign Languages), presents an actual nationwide research involving a cluster of local teachers. The PDCinMFL project provides a pedagogical understanding and practices clearly by explaining principles on putting foreign language teaching theories into practice, and relating the features of language skills to the principles of the project. It seems that there is a significant impact on practitioners' beliefs and practices when teaching have been conducted with a research-based curriculum. Chapter 8 "Putting the principles into practice" offers concrete principles embodied in L1 and L2 listening, reading, oral interaction and writing, and how they stand in a classroom practice. This chapter suggests a number of pedagogical tasks, which can be accessed openly on the PDCinMFL website. Chapter 9, "By way of conclusion" summarizes the aims of the book and the pedagogical and practical ideas of the different chapters in the book.

The book can serve as a standard reference for teachers, researchers, student teachers, curriculum designers, and even education policy makers. The integration of theory and practice contextualized on argumentative points works well. Generally, the book provides readers with clear, up-to-date and accessible account of information in Foreign Language Teaching by connecting theory and practice. As a useful resource, the book is recommended to the readers, especially to those interested in improving foreign language teaching and research-based curriculum and pedagogy.

Author notes

I acknowledge that this book review is original and is not being considered for publication elsewhere.

ASIAN EFL JOURNAL
The EFL Professional's Written Forum

Language, Identity and Education on the Arabian Peninsula- Bilingual Policies in a Multilingual Context

L. Buckingham, UK: Multilingual Matters, 2016. pp. 256. Hb.$149.95, ISBN: 9781783096596

Raheb Zohrehfard
University of British Columbia, Canada

Bioprofile:

Raheb Zohrehfard is now a Master of Education graduate in TESL from the Language and Literacy Education Department at the University of British Columbia. His focus is on identity issues of those who immigrate to English speaking countries. He tries to help non-native English speaking students to negotiate and construct their identity and legitimacy in their new community. **Email:** raheb.zohrehfard@alumni.ubc.ca

Focusing on both public and private spheres, this collection considers the importance of English and immigrants' languages in the context of Arabian Peninsula. Extending beyond English–Arabic societal bilingualism, the writers contend that the language practices of the Peninsula's citizens and residents serve multiple purposes in their daily lives. Chapters on home and heritage languages, identity, ELT, commercial signage and academic publishing help delve into the ubiquitous linguistic diversity in these dynamic settings.

In chapter 1, O'Neill discusses the multilingual landscape in the UAE and how the dominance of English has posed threat to the language, culture, and traditions of the Emiratis, as evidenced by the emergence of Arabizi (p.13). Language mixing is happening in the UAE and this is due to the rapid globalization, which has cornered the tradition and language leading to the indigenous inhabitants being at minority. There is a

Department of Language and Literacy Education, University of British Columbia, Vancouver, Canada

wide range of predispositions and expectations, which have polarized children and families (P.16).

Van den Hoven and Carroll in Chapter 2 investigate the sociolinguistic dimensions of language use in Abu Dhabi perceived by 16 female Emirati pre-service teachers. The study suggests that Indian, Farsi, Korean, and Filipino, though to differing extents, are seen as speech communities in Abu Dhabi. Although these languages function as peripheral languages on the margins of social life, they possess what it takes to consider Abu Dhabi a multilingual context. Descriptions of local understanding of the linguistic complexity in this context can serve as curricular resources for cultivating intercultural and inter-linguistic awareness.

Chapter 3 explores the use of English and the inclusion of cultural references on commercial signs on private enterprises in the capital cities of Yemen and Oman. By examining the textual features of bilingual shop signs in Sana'a and Muscat, Buckingham and Al-Athwary illustrate how English use and cultural references on these signs reflect different sociocultural realities in these cities. The language choice and selection of cultural references on commercial signs in Muscat and Sana'a reflect the different social function of English in these cities, while the manner in which English is used sheds light on the relative position of the language in comparison with Arabic and migrant languages in the broader national context.

Chapter 4, by Hopkyns, highlights identity construction for the nationals due to the internationalization strategy of the education sector and the high ratio of resident migrants to nationals. It investigates the effects of global English on Emirati cultural identity and reveals that Arabic is connected to home life and religion, with English associated with education and the wider world. With a large number of English speaking expatriates working in the UAE, English has become a very practical tool as a lingua franca at all levels of society.

Chapter 5 deals with the language situation as well as with the educational consequences in Qatar. In this chapter, Abu-EL-Kheir examines the complexities of Qatar's unique situation by investigating the interconnections between identity, language use and education through a series of interviews with one Qatari participant. The participant in this study sees her educational background and experiences related to her identity. It is shown that conceptualizing identity, as negotiated belonging is a useful analytic tool.

In chapter 6, Gallagher examines issues with regards to the language of teaching in the United Arab Emirates. The introduction of very early English immersion education for Arabic speaking children presents challenges, which puts the teaching and learning of Arabic in a somewhat parlous state. Children's self-identification as confident users of the two languages is at odds with their professors' estimation of their English proficiency. The medium of instruction policies and practices in the education system in the UAE need much detailed and nuanced research to be enhanced.

Chapter 7 investigates the International English Language Testing System (IELTS) and the suitability of the examination as a university entrance gatekeeper at national universities in the UAE. Having provided a brief history of the examination, Freimuth questions the appropriateness of high stakes international English language tests, arguing that IELTS, which originated from a different cultural context, cannot prove to be recognized as standard, offering an attempt to create a localized version of the exam.

In chapter 8, Solloway reports a study that attempts to investigate the way female Emirati higher education students regard English in today's world, how they perceive EMI policy in tertiary education in the UAE, and what penetrating effect English has had on the UAE society, culture and religion. The most striking finding of this study is that more than half of those surveyed believe that English and its symbiotic relationship with globalization constitutes a threat to the Islamic values and customs of the UAE.

The primary aims of Nuske in chapter 9 were firstly to investigate how TESOL practitioners characterized the dynamics of their teaching contexts based on their lived histories, and secondly, to examine how feasible the teaching of post-method and critical pedagogy is for Saudi EFL classrooms. While novice EFL teachers from the peninsula grapple with integrating knowledge and practices into their localized home contexts, they expressed positive attitudes toward post method pedagogy.

Chapter 10 by Buckingham and Ramachandran concerns publishing practices of scientific journals and how they can exercise an important function in developing local research capacity. They investigate editors' perspectives on the use of English in scientific research as a tool to build and maintain a professional community of practice, as well as the process of 'shaping' a manuscript to prepare it for publication. This chapter could have dug deeper into notions of textual borrowing, plagiarism, and textual ownership (Pecorari & Petrić, 2014) to examine writers' perceptions about academic writing.

This collection is a high-quality, worthwhile read that (1) examines the urban multilingual dynamics of the Arabian Peninsula, (2) considers the importance of both English and immigrants' languages in a context of rapid socioeconomic development, and (3) deepens insight into the inherent linguistic diversity in these dynamic societies. Nevertheless, the book could have expanded the discussion on (1) identity as a struggle of competing ideologies governed by varying degrees of socio-cultural capital (Darvin & Norton, 2015), (2) an appropriate linguistic repertoire required to write about discipline-specific topics (Sowden,2005), and (3) the cultural influences of multilingual researchers on their writing as a social practice.

References

Darvin, R., & Norton, B. (2015). Identity and a model of investment in applied linguistics. *Annual Review of Applied Linguistics, 35*, 36-56.

Pecorari, D., & Petrić, B. (2014). Plagiarism in second-language writing. *Language Teaching, 47*(03), 269-302.

Sowden, C. (2005). Plagiarism and the culture of multilingual students in higher education abroad. *ELT Journal, 59*(3), 226-233.

Author's Notes

Hereby I avow that the submission has not been previously published and is not being considered for publication elsewhere.