



TESOL International Journal

Teaching English to Speakers of Other Languages

Volume 13 Issue 1 2018

ISSN 2094-3938



TESOL INTERNATIONAL JOURNAL

Volume 13 Issue 1, 2018

Chief Editor

Xinghua Liu

Published by the TESOL International Journal

<http://www.tesol-international-journal.com>

© English Language Education Publishing Brisbane
Australia

This book is in copyright. Subject to statutory exception no reproduction of any part may take place without the written permission of English Language Education Publishing.

No unauthorized photocopying

All rights reserved. No part of this book may be reproduced, stored in a retrieval system or transmitted in any form or by any means, electronic, mechanical, photocopying or otherwise, without the prior written permission of English Language Education Publishing.

Chief Editor: Dr. Xinghua Liu

ISSN. 2094-3938

TESOL International Journal

Chief Editor

Xinghua Liu

Shanghai Jiao Tong University, China

Associate Editors

Hanh thi Nguyen

Hawaii Pacific University, USA

Joseph P. Vitta

Queen's University Belfast, UK

Dean Jorgensen

Gachon University, South Korea

Khadijeh Jafari

Islamic Azad University of Gorgan, Iran

Editorial Board

Ai, Haiyang - *University of Cincinnati, USA*

Anderson, Tim - *University of British Columbia, Canada*

Arabmofrad, Ali - *Golestan University, Iran*

Batziakas, Bill - *Queen Mary University of London, UK*

Behfrouz, Behnam - *University of Buraimi, Oman*

Bigdeli, Rouhollah Askari - *Yasouj University, Iran*

Bretaña – Tan, Ma Joji - *University of the Philippines, Philippines*

Çakir, İsmail - *Erciyes University, Turkey*

Chang, Tzu-shan - *Wenzao Ursuline University of Languages, Taiwan*

Choi, Jayoung - *Georgia State University, USA*

Chuenchaichon, Yutthasak - *Naresuan University, Thailand*

Chung, Edsoulla - *University of Cambridge, UK*

Cutrone, Pino - *Nagasaki University, Japan*

Dang, Doan-Trang Thi - *Monash University, Australia*

Deng, Jun - *Central South University, China*

Derakhshan, Ali - *Golestan University, Iran*

Dodigovic, Marina - *American University of Armenia, Armenia*

Farsani, Mohammad Amini - *Kharazmi University, Iran*

Floris, Flora Debora - *Petra Christian University, Indonesia*

Hos, Rabia - *Zirve University, Turkey*

Ji, Xiaoling - *Shanghai Jiao Tong University, China*

Jiang, Xuan - *St. Thomas University, USA*

Kambara, Hitomi - *University of Oklahoma, USA*

Khajavi, Yaser - *Shiraz university, Iran*

Lee, Sook Hee - *Charles Sturt University, Australia*

Li, Chili - *Hubei University of Technology, China*

Li, Liang - *Jilin Normal University, China*

Li, Yiyang - *Wenzao Ursuline University, Taiwan*

Lo, Yu-Chih - *National Chin-Yi University of Technology, Taiwan*

Nguyen, Ha Thi - *Monash University, Australia*

Niu, Ruiying - *Guangdong University of Foreign Studies, China*

O'Brien, Lynda - *University of Nottingham Ningbo, China*

Rozells, Diane Judith - *Sookmyung Women's University, S. Korea*

Salem, Ashraf Atta Mohamed Safein - *Sadat Academy for Management Sciences, Egypt*

Sultana, Shahin - *B. S. Abdur Rahman University, India*

Ta, Thanh Binh - *Monash University, Australia*

Tran-Dang, Khanh-Linh - *Monash University, Australia*

Ulla, Mark B. - *Mindanao State University, Philippines*

Witte, Maria Martinez - *Auburn University, USA*

Wu, Chiu-hui - *Wenzao Ursuline University of Languages, Taiwan*

Yan, Yanxia - *Xinhua University, China*

Yu, Jiying - *Shanghai Jiao Tong University, China*

Zhang, Lidong - *Shanghai Jiao Tong University, China*

Zhang, Xinling - *Shanghai University, China*

Zhao, Peiling - *Central South University, China*

Zhao, Zhongbao - *Hunan Institute of Science and Technology, China*

Contents

Foreword from the Editor <i>Xinghua Liu</i>	i
Student Views on the Use of L1 in the Foreign Language Classroom <i>Glen Clancy</i>	1
A Longitudinal Analysis of Ability Grouping with College EFL Learners <i>Chiu-hui (Vivian) Wu</i> <i>Chia-jung Tsai</i> <i>Yi-Min Chiu</i>	20
Language Input Effects on L2 Composition Peer review Feedback <i>Clay Williams</i>	33
Learning to Spell in English by Chinese Students: A Cross-sectional Study <i>Donald R. Bear</i> <i>Sam von Gillern</i> <i>Wei Xu</i>	47
Exploring Young English Learners' Perceptions of the Picture Word Inductive Model in China <i>Xuan Jiang</i>	67
Critical Pedagogy in EFL Teacher Education in the United Arab Emirates: Possibilities and Challenges <i>Rana Raddawi</i> <i>Salah Troudi</i>	79
Teachers' Needs in the Advancement of Communicative Language Teaching (CLT) in Taiwan <i>Shu-Hsiu Huang</i> <i>Li-Chin Yang</i>	100

Foreword from the editor

Before introducing the first issue of the year, I am pleased to share some exciting news with our readers. From 2018 onwards, *TESOL International Journal* will be indexed in SCOPUS, in addition to already being indexed in the MLA (Modern Language Association) index, Asian Education Index, Social Science Research Network, Google Scholar, and Open J Gate & Ulrich's Web. This is a remarkable achievement and it is a testament to the success of the journal. This achievement, of course, would not have been possible without the constant support and help from readers, contributors, and our dedicated editorial team.

This issue of *TESOL International Journal* contains seven papers. In the first paper, **Glen Clancy** surveyed the views of students receiving tertiary English language education in Japan on the use of the L1 in English classes. The study reveals a negative correlation between desired L1 application in EFL classes and the L2 proficiencies of the students, and a discrepancy in the desired objective for L1 application for different student levels of L2 proficiency.

Chiu-hui (Vivian) Wu, Chia-jung Tsai and **Yi-Min Chiu** investigated the effect of organizing English classes in a general EFL program in Taiwan by students' language proficiency. They found that students gained in English proficiency over time and they suggest a more nuanced understanding about the value of a proficiency-based placement.

In the third paper, **Clay Williams** studied the impact of language choice (L1 or L2) in written peer review sessions on students' transmission and reception of productive commentary. The study indicated that L2 written peer reviews may be more beneficial at identifying/rectifying paragraph-level and structural issues, whereas L1 peer review was slightly better for correcting rhetorical and logical issues.

Donald R. Bear, Sam von Gillern and **Wei Xu** researched the developmental features in English orthography by Chinese EFL students from grades two to eight. They found students made predictable spelling errors that reflected a similar developmental sequence to native English speakers.

Xuan Jiang examined perceptions of Chinese fourth and seventh grade EFL learners regarding a new teaching method in English, namely the Picture-Word Inductive Model. The study found that the surveyed students had overall positive views of the method.

In the sixth paper, **Rana Raddawi** and **Salah Troudi** studied the successes and challenges of adopting a critical approach to English education in the United Arab Emirates. They suggest that to make this teaching approach successful in the Arab world, new action plans for EFL teacher pre-service and in-service training programs are needed.

In the last paper, **Shu-Hsiu Huang** and **Li-Chin Yang** examined teachers' needs for better communication-oriented classroom practices in Taiwan. They found that teachers expressed a need for in-service training and assistance from native English-speaking teachers. Further, students, parents, and school authorities need to value and support the development of communication proficiency.

Readers who are actively engaged in research or have done research related to English language education are encouraged to contact us about the possibility of publishing with *TESOL International Journal*. Apart from individual paper submissions, we also welcome proposals for special issues.

Xinghua (Kevin) Liu
School of Foreign Languages, Shanghai Jiao Tong University, China
Email: luxinghua@sjtu.edu.cn

Student Views on the Use of L1 in the Foreign Language Classroom

Glen Clancy*

La Trobe University, Australia

Abstract

The use of first languages (L1) in the English as a Foreign Language (EFL) classroom in Japanese Universities is often a source of robust debate. In recent years, there has been an increase in counterarguments against the L2 monolingual EFL classroom and a strengthening of support for L1 usage. This study examines the views of students receiving tertiary English language education on the use of the L1 in English classes. The analysis suggests that students studying English as a second language (L2) in Japanese universities overwhelmingly preferred the use of L1 to aid in the facilitation of learning in EFL classes. Moreover, there were notable trends in attitudes between differing levels of student L2 proficiencies. The results reveal a negative correlation between desired L1 application in EFL classes and the L2 proficiencies of the students, and a discrepancy in the desired objective for L1 application for different student levels of L2 proficiency. These findings suggest the approach adopted for EFL courses in Japanese universities should implement judicious use of L1 and recognize that different levels of L2 proficiency will affect preferences for L1 usage in the EFL classroom.

Keywords: L1 usage, L2, student views, EFL, teacher L1 fluency, purpose of L1

Introduction

Robust debate surrounds the use of the first language (L1) in the foreign language classroom. It is an emotive issue that can call into question the teaching skills of foreign language instructors (Burden, 2000a); teachers can be made to feel guilty or inadequate for using an L1 in the L2 (second language) classroom (Littlewood & Yu, 2011). There is no consensus for L1 usage in the foreign language classroom and there is variation in L1 use between countries and teaching institutions (Rolin-Ianziti & Varshney, 2008). There has, however, been a general shift in the debate in the past few decades from a strict emphasis on exclusive L2 usage towards a more balanced view recognizing the benefits of appropriate usage of an L1 (McMillan & Turnbull, 2009; Rolin-Ianziti & Varshney, 2006).

In Japan, many educational institutions have individual policies for the use of an L2 exclusively (Tsukamoto, 2011); however, the prohibition of L1 use in English classes is often advocated beginning in elementary school through to tertiary education. For example, some researchers argue that a monolingual approach is the most effective as it emulates the method in which a child acquires their L1 (Butzkamm, 2003; Cummins, 1998; Pennycook, 1994; Phillipson, 1992). L1 use is often regarded as a barrier to effective teaching in many secondary school classrooms (Tsukamoto, 2011). The overuse of L1 in Japanese high schools has been attributed to factors including "Japanese English teachers' own lack of communicative ability," a "lack of teacher training," and "the emphasis placed on university entrance examinations" (McMillan & Rivers, 2011, pp. 251-252). An English-only approach has recently been hailed as the solution to Japan's low international English rankings (McMillan & Rivers, 2011). The Japanese Ministry of Education, Culture, Sports, Science, and Technology (MEXT) announced in 2008 that English classes in high schools should be conducted in L2 (Tsukamoto, 2011).

* Email: gccquarterly@gmail.com. Tel.: +613 9479 1111 Address: Plenty Rd & Kingsbury Dr, Bundoora VIC 3086

Despite the Japanese government's push for greater L2 exclusivity in English classes, there has been an increase in counterarguments against the monolingual English as a Foreign Language (EFL) classroom and a strengthening of support for L1 usage. In recent years, there have been arguments made for the judicious and theoretically principled use of the L1 in the EFL classroom (Cook, 2001; Levine, 2003; Liebscher & Dailey-O'Cain, 2004; Turnbull, 2001).

In contrast with the vast majority of EFL literature concerning L1 usage in the foreign language classroom that focuses on the analysis of pedagogical methods and theories, this study examines the opinions of students who are the recipients of tertiary English language education. University students receiving L2 (English) education from a foreign EFL teacher were asked about their preferences for the foreign EFL teacher being fluent in the L1 (Japanese), using L1 when appropriate, the purpose of L1 usage, and whether a foreign EFL teacher should pretend to lack L1 proficiency. Quantitative and qualitative analysis of the data found that Japanese university students largely favored L1 usage to aid learning in EFL classes, but there were also significant differences in preferences across the various levels of student L2 proficiencies.

Literature Review

The Advantages of a Monolingual Approach in EFL Classes

The monolingual approach to English teaching is often revered for emulating the method in which a child acquires their L1 (Butzkamm, 2003; Cummins, 1998; Pennycook, 1994; Phillipson, 1992). L2 exclusivity enhances subconscious learning, and there is a direct relationship between comprehensible L2 input and proficiency (Krashen, 1982). The key attraction to a monolingual approach is student exposure to the target language—the greater the exposure to L2, the faster students will learn (Ellis R., 2005). Macaro (1997) argues that L1 has no pedagogical value for students and in fact is a barrier to L2 learning. L1 use in the EFL class may also decrease the motivation of students by dismissing the importance of L2 as a communicative tool (Littlewood, 1992). Ellis (1985), on the other hand, while promoting the monolingual approach, recognizes that an L1 may be necessary to explain and organize tasks and manage the behavior of students to facilitate the functioning of the EFL class, although they lament that this may be detrimental to language acquisition through the reduction of L2 input. There is also the risk that allowing L1 in the foreign language classroom will lead to excessive use (Turnbull, 2001). Turnbull (2001) argues that the main issue with L1 is formulating the appropriate parameters for “an optimal or acceptable amount of [L2] and L1 use” (p. 531).

The Advantages of a Bilingual Approach in EFL Classes

More recently, there has been growing support for the bilingual approach to foreign language classrooms, with greater recognition of judicious and theoretically principled L1 use (Cook, 2001; Levine, 2003; Liebscher & Dailey-O'Cain, 2004; Turnbull, 2001). Swain and Lapkin (2000) argue that judicious use of the L1 can indeed support L2 learning and use. To insist that no use be made of the L1 in carrying out tasks that are both linguistically and cognitively complex is to deny the use of an important cognitive tool. (p. 268)

The bilingual learner is now often presented as the best model for L2 acquisition, one who can use skills learned in one language to facilitate advancement in another (Butzkamm & Caldwell, 2009; Cook 2001). For example, bilingual learners consciously use mental strategies and use their L1 to forecast what works for L2 (Gass & Mackey, 2000). Cook (2001) argues that L1 use in the foreign language classroom may help develop “genuine L2 users” (p. 412) that are “mediators” (p. 407) between the L1 and L2 rather than “imitators” (p. 407) of native speakers.

There are three prevalent theories that provide evidence for L1 having a facilitating effect in the foreign language classroom. Firstly, the cognitive processing theory (Ellis N., 2005) demonstrates that L1 and L2 are not held in separate conceptual stores, and the mental lexicon is best explained as a series of connections, which are not language specific until activated (Ellis N., 2005; Kroll, 1993; Libben, 2000). Connections with L1 will be

much stronger; hence abandoning these connections will disregard an important tool for L2 learners (Ellis N., 2005; Kroll, 1993; Libben, 2000). The second, a socio-cultural theory promoting the use of L1 in foreign language learning proposes that inner voice and private speech, which are crucial devices in the way we think and act, are almost always carried out in L1 (Antón & DiCamilla, 1998; Brooks, Donato, & McGlone, 1997). Codeswitching in naturalistic environments makes up the third theory advocating the benefits of L1 in foreign language teaching, which compares switching from L2 to L1 in the classroom with naturalistic codeswitching outside the classroom (Håkansson & Lindberg, 1988). This theory identifies the advantages of using L1 for conveying message-oriented information in enhancing the process of foreign language lessons (Håkansson & Lindberg, 1988). Although the benefits of codeswitching are not conclusive, there is no evidence that teacher codeswitching is detrimental to lexical acquisition (Turnbull & Dailey-O’Cain, 2009) or that codeswitching by the teacher has a “negative impact on the quantity of students’ L2 production” (Macaro, 2005, p. 72).

Raschka, Sercombe, and Chi-Ling (2009) list four common functions of L1 in the EFL classroom: socialization (i.e., “when teachers turn to the students’ first language to signal friendship and solidarity”), topic switch (i.e., “when the teacher switches code according to which topic is under discussion”), classroom management (i.e., “where teachers negotiated progression of classroom activities in the students’ mother tongue”), and metalinguistic functioning (i.e., “where tasks were performed in the target language but comment, evaluation and talk about the task could take place in the first language”). However, the function of codeswitching between the L1 and L2 often fluctuates depending on the discourse in the classroom at a specific time (Raschka, Sercombe, & Chi-Ling, 2009). Eldridge (1996) argues that the functions of codeswitching interrelate in highly complex means making it difficult to identify the exact function of specific cases of L1 usage in the EFL classroom.

Tsukamoto (2011) breaks down the advantages of L1 in EFL learning into three categories: maintaining a comfortable class atmosphere, facilitating greater student comprehension, and class-time efficiency. The use of L1 can act as a tool to stimulate greater student participation by creating a relaxing atmosphere (Polio & Duff, 1994). Burden (2000b) found that an English-only approach isolates students and that the L1 is effective in providing a sense of security for learners in taking risks with the L2. Castellotti and Moore (1997) argue that L1 usage can create a low-anxiety classroom environment conducive to learning. Student comprehension may also be enhanced through L1 use. Krashen (1981) argues that bilingual learning provides knowledge and literacy in a student’s native language and indirectly enhances L2 proficiency. Students will often naturally equate the L2 with their native language; therefore, blocking this process may have negative effects (Harbord, 1992). Finally, L1 use can benefit L2 learners by increasing class-time efficiency and substituting time wasted on misunderstandings for more productive activities (Atkinson, 1987).

Teacher Views of L2 Usage

Researchers have also examined teacher and student views of L2 use in the foreign language classroom; however, in the past two decades, the bulk of studies have largely focused on teacher opinions (Macaro, 2001; Polio & Duff, 1994). Studies have found that teachers use L1 to explain new vocabulary and grammar, for instructions, and for student discipline (Kaneko, 1992; Macaro, 2001; Polio & Duff, 1994) and for creating a comfortable classroom atmosphere and teacher/student affinity (Kaneko, 1992; Polio & Duff, 1994; Rolin-Ianziti & Brownlie, 2002). The shift towards the recognition of the advantages of L1 in foreign language education has been reflected in studies on the opinions of EFL teachers. Makulloluwa’s (2013) study of EFL teacher opinions in Sri Lanka found that “a majority of the teachers demonstrated a positive attitude towards the use of L1 in the classroom” (p. 592). Similarly, Timor (2012) concluded in a study of EFL teaching in Israel that teachers demonstrated a “positive pedagogical stance” with “regard to Hebrew as the [L1] in EFL classes” (p. 13). McMillan and Rivers (2011) conducted an attitudinal study of 29 native-English speaker teachers at a Japanese university and found, contrary to the official university policy promoting exclusive L2 use, “many teachers believed that selective use of the students’ L1, by the teacher or by students, could enhance L2 learning in various ways within a communicative framework” (p. 251).

Student Views on L2 Usage

On the other hand, there has been limited research published on student views of L2 use in the foreign language classroom (Rolin-Ianziti & Varshney, 2008). Levine (2003) found through a study of 600 foreign language students and 163 foreign language instructors that, despite the prevailing “monolingual principle” in the US, the L1 reduces student anxiety and serves meaningful pedagogical functions. Rolin-Ianziti and Varshney published two studies in 2006 and 2008 exploring Australian university student views on L1 use in French, German, Japanese, and Spanish language classes (Rolin-Ianziti & Varshney, 2006; Rolin-Ianziti & Varshney, 2008). They found that students recognized the value of L1 use in the foreign language class; however, this indicated “necessity of a delicate balance between the L1 and the [L2] within the language classroom” (Rolin-Ianziti & Varshney, 2006, p. 78). L1 use was seen as a “double-edged sword” (Rolin-Ianziti & Varshney, 2006, p. 78). For example, “while students see grammatical items as easier to grasp in L1 due to their complexity, they also see the necessity for learning structures from language in use, in a natural context” (Rolin-Ianziti & Varshney, 2006, p. 78). Although there was no consensus, the majority of students viewed the role of L1 use as a facilitator of medium-oriented interactions, i.e., focusing on form rather than content, such as vocabulary and grammar explanations (Rolin-Ianziti & Varshney, 2008).

There has also been little research published on student views of L1 use in EFL classrooms in Japan. Burden conducted two studies examining Japanese university student views across varying levels of English proficiency, on student and teacher L1 use in English classes (Burden, 2000a) and another on changes in Japanese student views of L1 use in an English conversation class throughout a single university semester (Burden, 2004). Burden (2000a) found that the majority of students believed the teacher should have knowledge of the L1, and that the teacher and students themselves should use the L1 during class. This trend decreased as English proficiency levels increased, although postgraduate students bucked this trend and had similar views to the “pre-intermediate” level students (Burden, 2000a). Students were split into proficiency categories (Pre-intermediate, Intermediate, Advanced, and Postgraduate) based on their year level (Burden, 2000a). The most common reasons given for appropriate L1 use by the teacher was “relaxing the students,” “explaining the differences between [L1] and English grammar,” “explaining new words,” and “talking about tests” (Burden, 2000a, p. 144). Burden (2000a) concluded that students “recognize that communicative lessons with native speakers should be conducted in the [L2], while reserving the right to ask about usage through the [L1], thus creating a more relaxed, humanistic classroom where they can freely express themselves” (p. 139). Similarly, Tsukamoto (2011) conducted a “small study” on Japanese university students’ perceptions of L1 use in English classes. Tsukamoto surveyed 42 English major students asking, “Did you feel the instructor needed to use Japanese in class?” The majority of students (83%) believed the English instructor did not “need” to use L1 in the class (Tsukamoto, 2011, p. 150).

Research Questions

This study will contribute to current research (Burden, 2000a; Tsukamoto, 2011) by further exploring Japanese university student opinions (across varying levels of English proficiency) of the importance of a foreign English teacher’s L1 fluency, the need for L1 usage in English classes, and the specific purposes of L1 in English classes. This study will also present unique research regarding Japanese university students’ opinions of whether foreign English teachers should pretend to lack L1 proficiency in the EFL classroom. The research questions addressed by this study are: Do students prefer foreign EFL teachers to be fluent in the L1? When do students think it is appropriate to use the L1 during an EFL class? What purpose do students think the L1 serves, if any, in the EFL classroom? Do students think a foreign EFL teacher should pretend to lack L1 proficiency?

Methodology

This study employed both quantitative and qualitative methods of research to explore Japanese university

student views of L1 use in the EFL classroom. Questionnaires were administered to 175 first- and second-year undergraduate students from three universities in the Kanto region by the author and three other EFL instructors. The students undertook English classes taught by foreign native English teachers and were of varying English proficiencies. The questionnaire contained seven questions eliciting both quantitative and qualitative data on student views of the importance of a foreign English teacher's L1 fluency, the necessity for L1 use in EFL classes, the specific purposes of L1 in EFL classes, and whether a foreign English teacher should pretend to lack L1 proficiency. Students' L2 (English) proficiency levels were categorized through their Test of English for International Communication (TOEIC) Listening and Reading test scores—an exam that is designed to test students' everyday English skills working in an international environment.

Student respondents were asked if they preferred a foreign teacher that was fluent in the L1 (Japanese) or a foreign teacher that could not communicate in the L1. The students were also asked if they preferred the foreign teacher to speak no L1 or to speak L1 in certain situations. The questionnaire also contained a question asking students for what purpose, if any, did they want the foreign teacher to use L1. The options included “new vocabulary,” “grammar,” “instructions” (e.g., class instructions, homework, assignments), “administration matters,” “facilitating rapport/humour,” and “creating a comfortable atmosphere/enhancing communication.” Finally, students were asked if they preferred the foreign teacher to pretend to lack proficiency in the L1.

The data gathered from these questions were analyzed quantitatively both amongst the total number of students surveyed (175) and comparatively across each TOEIC proficiency level. Qualitative analysis was carried out through examining answers given to open questions regarding the respondents' reasoning for their preferences for a foreign teacher being fluent, or not, in the L1 and for a foreign teacher pretending to be fluent, or not, in the L1.

Results

This study explores four areas related to Japanese university students' opinions of L1 use in the EFL classroom across various levels of English proficiency. First, students were asked about the importance of foreign English teachers being fluent in L1 (Japanese); second, students were asked about whether the foreign English teacher should speak L1 (Japanese) at appropriate times during the EFL class; and third, students were asked about the specific purposes of L1 in EFL classes. Finally, students were questioned about whether the foreign English teacher should pretend to lack proficiency in the L1 (Japanese).

Desired Teacher Fluency

Overall, the majority of participants (66.29%) preferred that the foreign English teacher be fluent in L1 rather than having no knowledge of L1 (see Figure 1).

Unsurprisingly, there was a trend of decreased preference for teacher L1 fluency for participants with higher levels of English proficiency (see Figure 2).

Seventy-five percent of students with TOEIC scores ranging from zero to 300 preferred a teacher that was fluent in L1; sixty-seven percent of students with TOEIC scores ranging from 301 to 450 preferred a teacher that was fluent in L1; and seventy-three percent of students with TOEIC scores ranging from 451 to 600 preferred a teacher that was fluent in L1. On the other hand, only forty-seven and fifty percent of students with TOEIC scores ranging from 601 to 750 and 751 to 900, respectively, preferred a teacher that was fluent in L1. Common themes for preferring a foreign English teacher fluent in L1 across all levels included were that it is easier to learn, understand, and communicate during class. For example, one student with a TOEIC score of 528 wrote:

If the teacher speaks all English then often I won't understand so I would like the teacher to occasionally explain in Japanese.

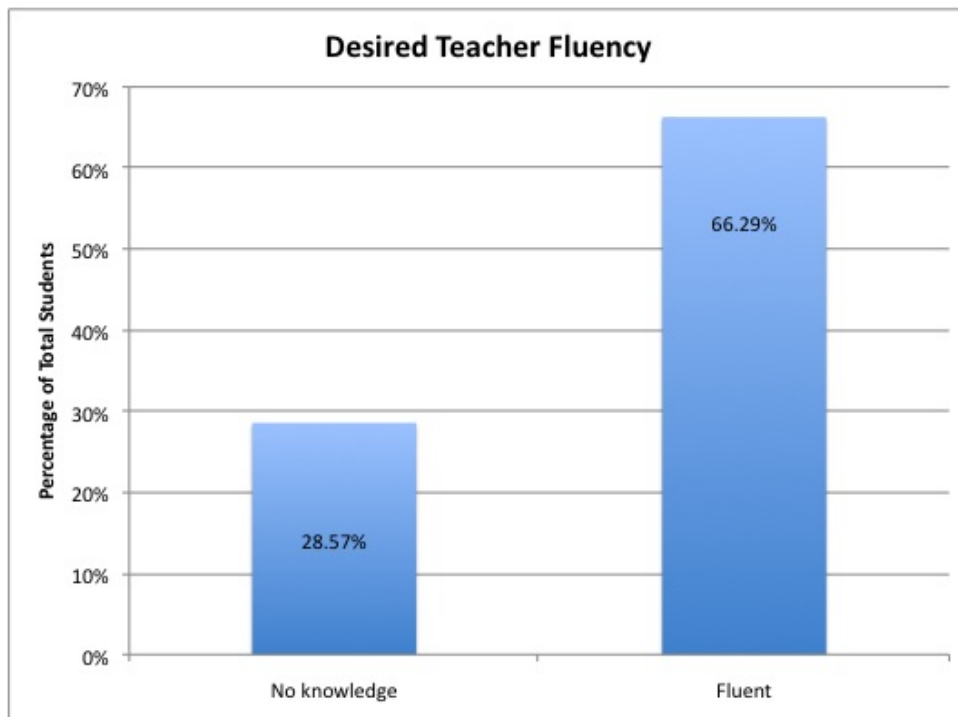


Figure 1. Students' preference of teacher L1 fluency (total students)

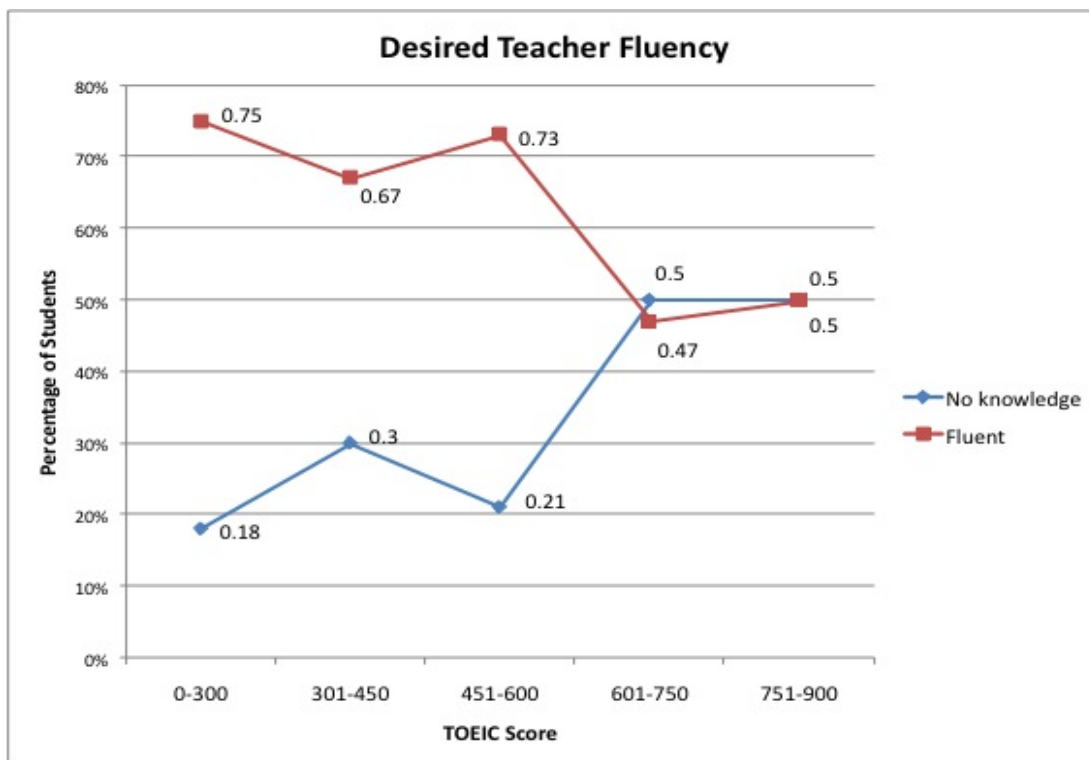


Figure 2. Students' preference of teacher fluency (per proficiency level).

Another student with a TOEIC score of 710, wrote:

If there is no way to communicate in English the teacher can understand. Also, the teacher can correct our English.

Other common reasons given for preferring a foreign English teacher fluent in the L1 were that the teacher could avoid “misunderstandings,” that it was useful “to communicate as a last resort” when L2 usage failed, and that students could “relate to the teacher.” For example, one student with a TOEIC score of 575, wrote:

It's better to relate to the teacher and easier to understand. If the teacher has gone through the same hardships learning a foreign language they will understand the student's position.

Common reasons given amongst lower level students were that a fluent foreign teacher enabled the class to “function smoothly” and encouraged student motivation. For example, a student with a TOEIC score of 155, wrote:

I can't understand what is being said from the beginning so I lose motivation.

Conversely, common themes amongst students that preferred the foreign English teacher had no knowledge of the L1 include not depending on the L1 and therefore being forced to try harder to use English. For example, one student with a TOEIC score of 450, wrote:

If there is no other way than to speak in English then the student will try their best.

Another student with a TOEIC score of 775, wrote:

If we know the teacher speaks Japanese we might depend on it.

Other common reasons given for preferring the foreign English teacher had no knowledge of the L1 was to become familiar with the L2 and increase the quantity of English practice during class. For example, one student with a TOEIC score of 160, wrote:

The class is a chance to speak English so we can get used to English conversation.

Another student with a TOEIC score of 603, wrote:

To make an English only environment and quickly get used to it.

Using L1 When Appropriate

The majority of all students (85.71%) preferred that the foreign English teacher use L1 when appropriate during class (see Figure 3).

Unsurprisingly, there was a negative correlation between English proficiency and preferring the teacher use L1 when appropriate (see Figure 4).

There was a gradual downward trend of desired L1 usage when appropriate with increasing English proficiency. Ninety-one percent of the students with the lowest-level English proficiency (TOEIC scores ranging from zero to 300) preferred that foreign English teachers use L1 when appropriate compared to just 67% of the students with the highest-level English proficiency (TOEIC scores ranging from 751 to 900).

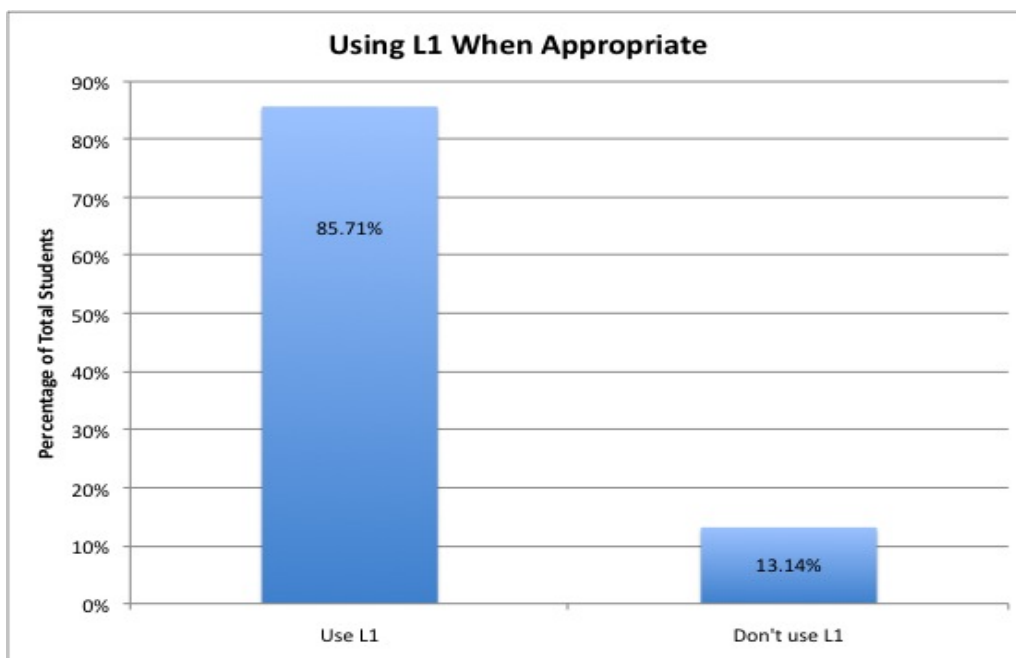


Figure 3. Students' preference for teacher using L1 when appropriate (total students).

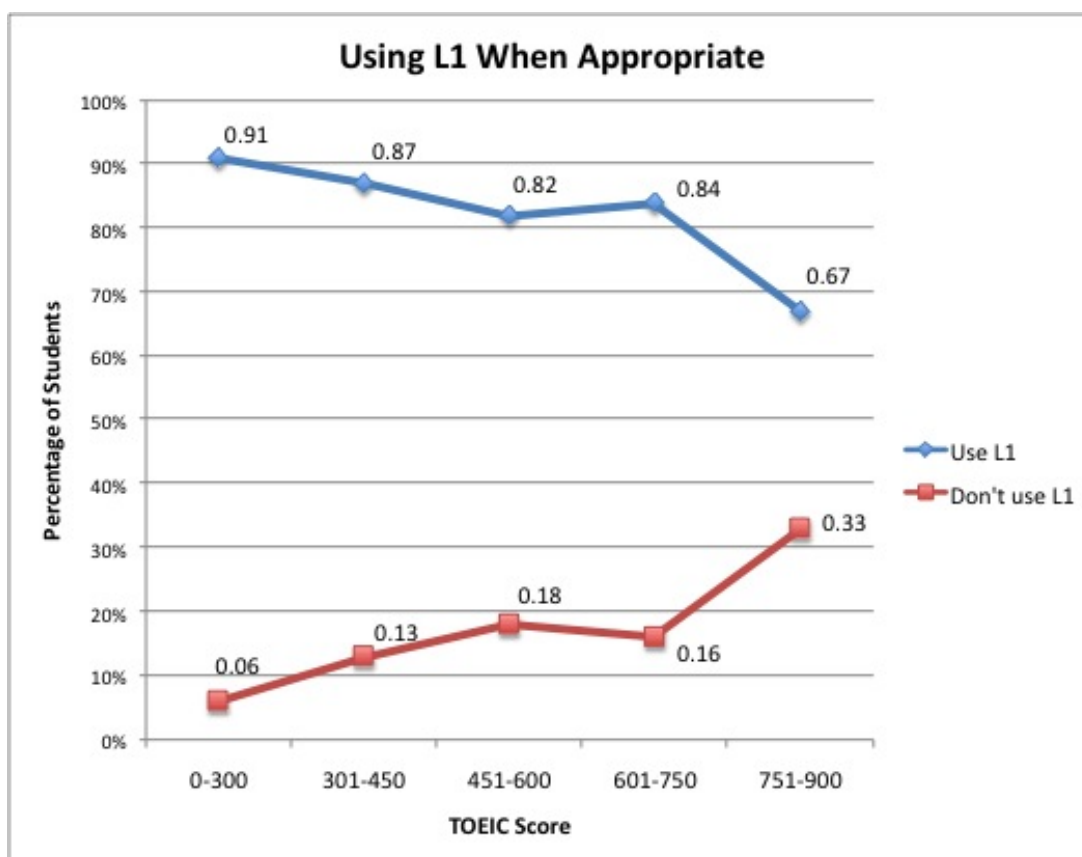


Figure 4. Students' preference for teacher using L1 when appropriate (per proficiency level).

Purpose of L1

New vocabulary was considered the most important area for the EFL teacher to use L1 (42% of all students identified new vocabulary) (see Figure 5). However, facilitating teacher/student rapport and humor (39% of all students), class instructions (35% of all students), class atmosphere/aiding communication (35% of all students), and grammar (34% of all students) were similarly highly regarded. Surprisingly, L1 for administrative purposes was only deemed necessary by 11% of all participants. Only 5% answered that L1 served “no purpose.”

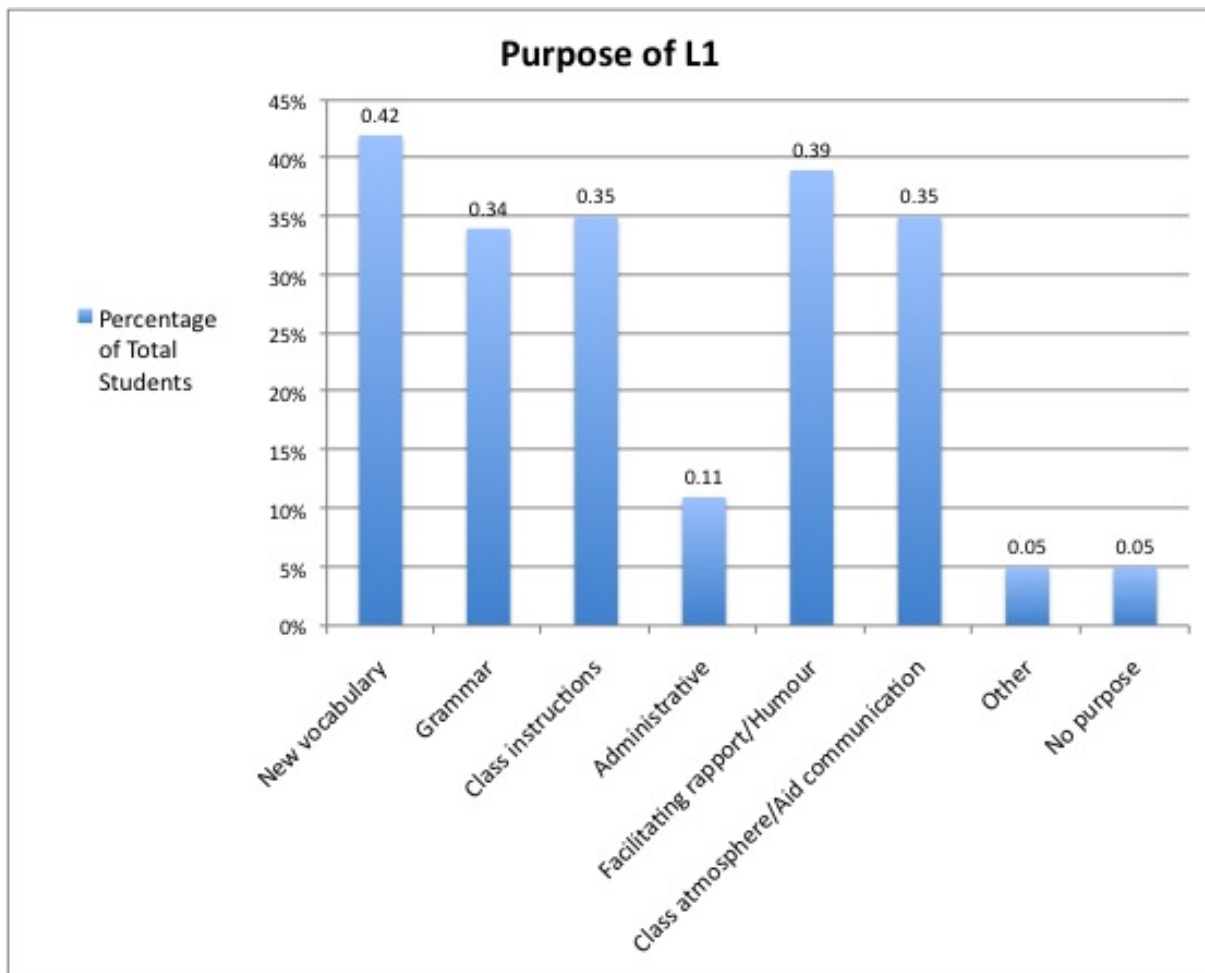


Figure 5. Desired purpose of L1 (total students).

The students with the highest TOEIC scores (751-900) identified new vocabulary (50% of the highest-level students) and class atmosphere (42% of the highest-level students) as the most desired areas for the EFL teacher using L1 (see Figure 6). Total percentages add to more than 100% due to students answering one or more desired purposes for L1 usage.

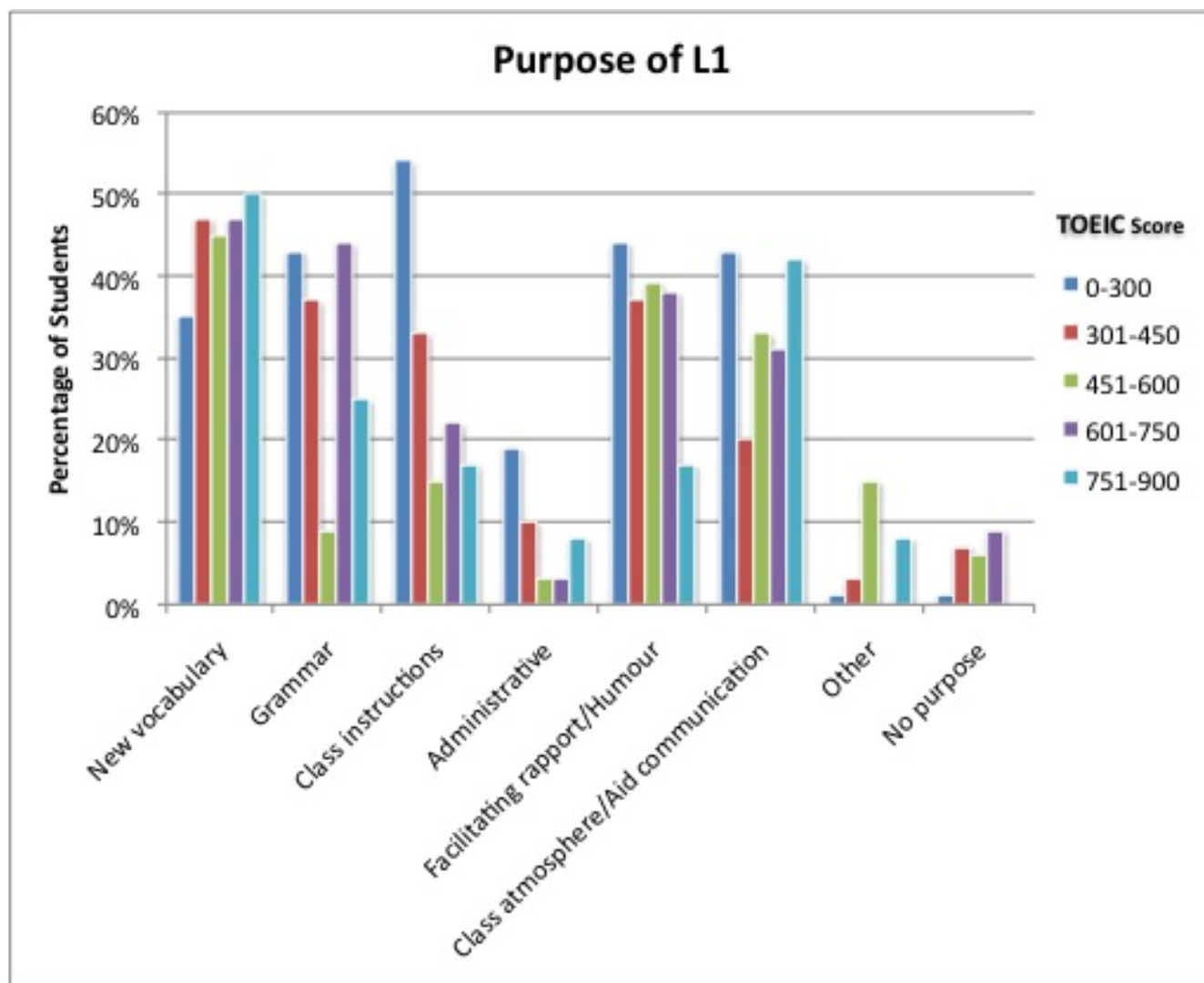


Figure 6. Desired purpose of L1 (per proficiency level)

New vocabulary had a slight trend of being more favourable the higher the TOEIC level of the students (see Figure 6). On the other hand, the three areas of facilitating teacher/student rapport and humor, class instructions, and administration all had a decreasing trend the higher the TOEIC level of the students.

The most desired purpose for L1 by students with the lowest TOEIC scores (0-300) was class instructions, with 54% of the lowest-level students (see Figure 6). Facilitating teacher/student rapport and humor (44%), class atmosphere (43%), and new vocabulary (35%) were also highly valued purposes for L1 use in the EFL classroom by the lowest-level students.

Pretending to Lack L1 Proficiency

Overall, only a small minority of students (16.57%) believed that foreign English teachers should pretend to lack proficiency in the L1 (Japanese) (see Figure 7).

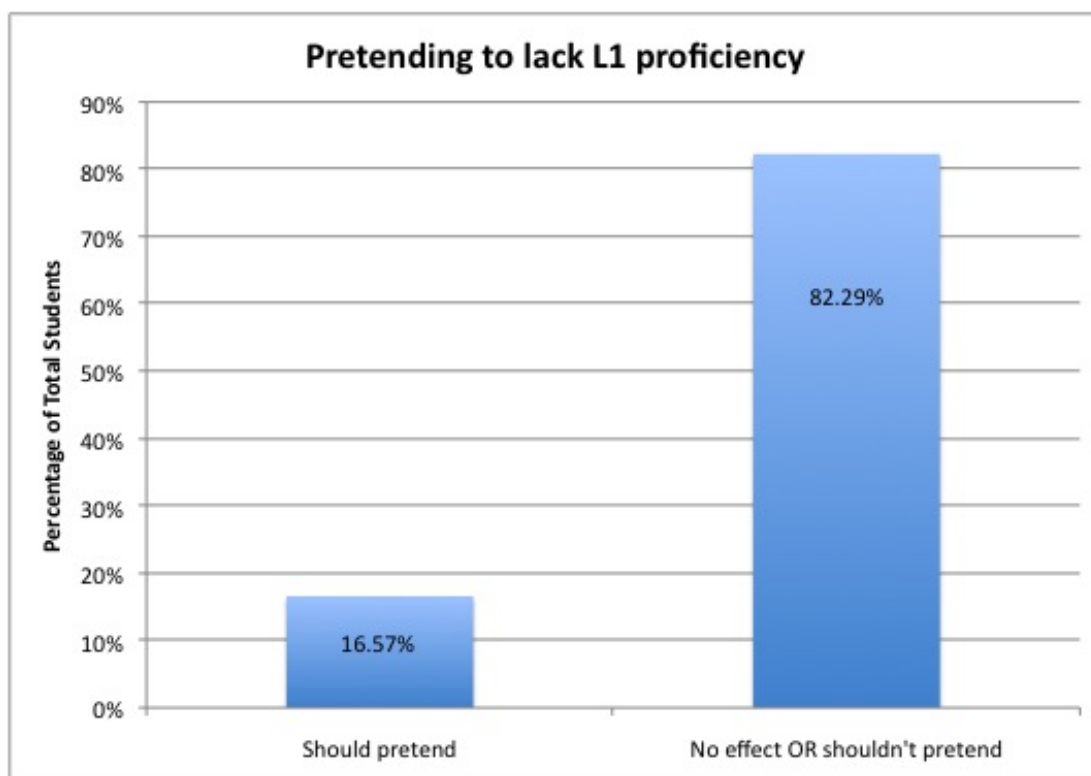


Figure 7. Students' preference of teacher pretending to lack L1 proficiency (total students)

As levels of English proficiency increased, more students believed the teacher should pretend to lack L1 proficiency (see Figure 8). However, surprisingly, the views of students with the highest TOEIC scores (751-900) bucked this trend and were almost identical to the students with the lowest TOEIC scores (0-300), with only eight and seven percent, respectively, indicating they preferred the foreign English teacher didn't pretend to lack L1 proficiency.

There were similar reasons given for rejecting the need for the foreign English teacher to pretend to lack L1 proficiency throughout all English levels of proficiency. One common reason given was to enable another means of communication with the teacher. For example, one student with a TOEIC score of 821, wrote:

Japanese (the L1) may be used as one method of communication.

Another intermediate level student with a TOEIC score of 515, supported this opinion, commenting:

It's easier to communicate if the teacher speaks Japanese.

Similarly, another student with a TOEIC score of 180 wrote:

It's better for class atmosphere and communicating more easily.

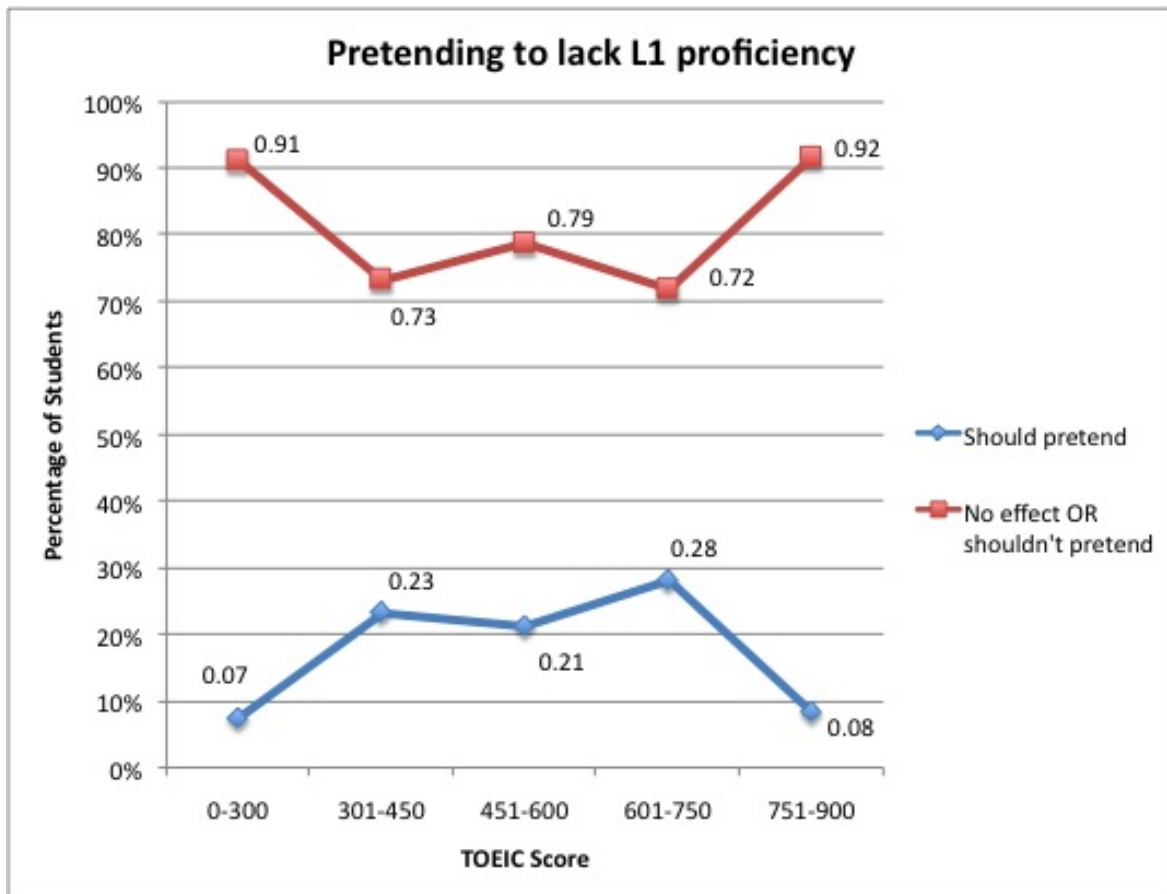


Figure 8. Students' preference of teacher pretending to lack L1 proficiency (per proficiency level)

Another common reason given for rejecting the need for the foreign English teacher to pretend to lack L1 proficiency was to facilitate a more effective lesson where misunderstandings could be explained in the L1. For example, one student with a TOEIC score of 825 wrote:

There are times when I want things explained in Japanese.

Similarly, another lower level student with a TOEIC score of 290 wrote:

I want things explained in Japanese if possible.

Another common reason given across most proficiency levels (excluding the highest-level students with TOEIC scores of 751-900) for rejecting the need for the foreign English teacher to pretend to lack L1 proficiency was to relate to the teacher. For example, one student with a TOEIC score of 405 wrote:

If the teacher speaks Japanese there is a greater sense of affinity and I will try harder to understand English.

Another student with a TOEIC score of 575, also commented on the importance of developing a relationship with the teacher:

I think it's better to understand the teacher and build a trusting relationship rather than pretending not to speak Japanese to increase English usage.

Similarly, another student with a TOEIC score of 600 wrote that knowing the teacher is learning the L1 can be stimulating for the class:

We can feel like we are learning a language together.

Comments by students that believed it was constructive for the teacher to pretend to lack L1 proficiency revealed common themes such as being encouraged to use L2 (English), L1 (Japanese) usage obstructing English practice, and an L2 exclusive class improving language acquisition. For example, one student with a TOEIC score of 210 wrote:

If Japanese is used our English won't improve.

Another student with a TOEIC score of 360 wrote:

If Japanese is spoken it will obstruct English practice.

One student with a TOEIC score of 703 was concerned about becoming too reliant on the L1 during the EFL class:

Because the students will tend to depend on communicating in Japanese.

Discussion

The opinions gathered from Japanese university students in this study provide insight into the views of L1 use in the EFL classroom by those receiving an English language education. The results of the student surveys largely mirror other studies examining student and teacher views on L1 usage in EFL classes (Kaneko, 1992; Polio & Duff, 1994; Rolin-Ianziti & Brownlie, 2002; Rolin-Ianziti & Varshney 2008). A significant percentage of student views advocating for the use of L1 by a foreign EFL teacher supports the shift in recent years of publications highlighting the advantages of L1 usage (Cook, 2001; Levine, 2003; Liebscher & Dailey-O'Cain, 2004; Turnbull, 2001).

Teacher L1 Fluency

Overall, a minority of students surveyed (28.57%) responded that the foreign EFL teacher should have no knowledge of the L1. Students that preferred that their foreign EFL teacher was not fluent in the L1—the majority of these having higher levels of English (L2) proficiency—believed this forced them to try harder because they could not rely on the L1, it assisted them in getting used to using L2, and also increased the quantity of L2 practice in the classroom.

These responses support R. Ellis' (2005) thesis that the main benefit of a monolingual approach in the EFL class is student exposure to the target language, thereby enhancing L2 acquisition. For example, one student with a TOEIC score of 810 wrote, "We can't talk a lot of English in Japan, so it's better not to speak Japanese in class." Another student with a TOEIC score of 450 wrote, "If there is no other way than to speak in English then the student will try their best." Moreover, Turnbull (2001) argues that allowing L1 usage in the EFL classroom may result in excessive use and asserts that the main problem with L1 is creating appropriate boundaries for "an optimal or acceptable amount of [L2] and L1 use" (p. 531).

Nevertheless, a majority (66.29%) of the students surveyed preferred their foreign EFL teacher to be fluent in their native language (Japanese). This point of view had a negative correlation with the English proficiency of the students, i.e., the lower the level of English proficiency, the more likely a student is to prefer that their foreign EFL teacher is fluent in their native language. Conversely, the higher a student's English proficiency, the less likely they are to prefer their foreign EFL teacher to be fluent in their native language. This is a significant finding that is rarely addressed in EFL literature. This study highlights that L2 proficiency plays a significant role in determining learners' desire for L1 usage in the EFL classroom.

Students who preferred an L1-fluent foreign EFL teacher—the majority having a lower English proficiency—believed it allowed some usage of L1 and benefitted learning when they “didn't understand.” For example, one student with a TOEIC score of 200 wrote, “Because the teacher can explain in more depth.” Another student with a TOEIC score of 300 wrote, “I want to ask the teacher not to skip something if I don't understand.” This illustrates that many students—especially those with a lower English proficiency—recognized L1 as an important tool to develop their foreign language skills. Cook (2005) supports this position, arguing that the bilingual learner is now often acknowledged as the best model for L2 learning, whereby a student can use their skills already learned in the L1 to facilitate advancement in the L2. The survey results also support Burden's (2000a) findings that the majority of Japanese university students learning English as a foreign language believe that the teacher should have knowledge of the L1.

Using L1 When Appropriate

The majority of all students surveyed (85.71%) believed that L1 should be used in the EFL class when appropriate. This outcome supports recent studies recognizing the advantages of the judicious and theoretically principled use of L1 in the EFL classroom (Cook, 2001; Levine, 2003; Liebscher & Dailey-O'Cain, 2004; Turnbull, 2001). This study also found a negative correlation between preferring L1 usage when appropriate and the L2 proficiency of the students, i.e., there is a decreasing trend of desired L1 usage when appropriate with increasing levels of L2 proficiency. These results demonstrate the importance of the L2 proficiency of the students in determining the parameters of L1 usage in the EFL class.

The overwhelming majority (91%) of students with the lowest-level L2 proficiency (TOEIC scores ranging from zero to 300) believed L1 usage should be allowed when appropriate, while only 67% of the students with the highest-level L2 proficiency believed that the L1 should be used when appropriate. Consequently, according to the recipients of tertiary English language education in Japan, the application of L1 when appropriate in the EFL class is more important for students with lower levels of English proficiency.

Purpose of L1

The survey produced a relatively complex mix of answers regarding the preferred purpose of L1 usage in the EFL classroom. Overall, new vocabulary (42% of all students) was considered the most significant area of importance for L1 usage, closely followed by the facilitation of teacher/student rapport and humor (39% of all students), class instructions (35% of all students), class atmosphere/aiding communication (35% of all students), and grammar (34% of all students). Significantly, only 5% of all students answered that L1 served no purpose in the EFL classroom, further highlighting student beliefs of the advantages of L2 usage in EFL lessons.

These results illustrate the similar findings of previous studies carried out on the application of L1 by teachers in EFL classrooms. Researchers found that teachers use L1 to explain new vocabulary and grammar (Polio & Duff, 1994) to give instructions (Kaneko, 1992; Macaro, 2001; Polio & Duff, 1994), and to create a comfortable classroom atmosphere and teacher/student affinity (Kaneko, 1992; Polio & Duff, 1994; Rolin-Ianziti & Brownlie, 2002). The results of this study also support the few published studies examining student views of L1 use in the EFL classroom. For example, Rolin-Ianziti and Varshney (2008) argue that the majority of Australian university students they surveyed viewed the role of L1 as a facilitator of medium-orientated interactions. In other words, L1 is an important tool for teaching aspects of language focusing on form rather than content such as new vocabulary or grammar. Similarly, this study found that students believed that L1

usage was most important for learning new vocabulary.

The students with the highest TOEIC scores (751-900) also identified new vocabulary (50% of the highest-level students) and class atmosphere (42% of the highest-level students) as the most desired areas for L1 usage. On the other hand, students with the lowest TOEIC scores (0-300) identified class instructions (54% of the lowest-level students), the facilitation of teacher/student rapport and humor (44% of the lowest-level students), class atmosphere (43% of the lowest-level students), and new vocabulary (35% of the lowest-level students) as the most important areas for L1 usage.

Respondents who believed that L1 usage was important for class atmosphere were relatively evenly spread amongst all levels of L2 proficiency; previous research supports this widely-held view amongst respondents. For example, Tsukamoto (2011) argues that maintaining a comfortable class atmosphere is one of the key advantages of using L1 in the EFL classroom. Polio and Duff (1994) found that the L1 is important for providing a sense of security for learners and can act as a tool to stimulate greater student participation. Moreover, Castellotti and Moore (1997) argue that the L1 can create a low-anxiety environment and enhance L2 learning. The results of this study suggest that the students surveyed were aware of these benefits, borne out of L1 usage, for creating a class atmosphere more conducive to learning.

These results also support Burden's (2000a) study in which students were split into proficiency categories. One of the most common reasons given for appropriate L1 use was "relaxing students." Burden argues that allowing students to ask about L2 usage through the L1 creates a more "relaxed, humanistic classroom where they can freely express themselves" (p. 139).

In contrast, the desire to use the L1 for class instructions was most prevalent amongst students with lower English proficiency levels. This again indicates that there is a discrepancy in the views of the role of L2 for differing levels of L2 proficiency. For example, one student with a TOEIC score of 300 wrote:

There are times when I can't understand anything or when I can't understand instructions so I want the teacher to be able to speak some Japanese even if the teacher isn't fluent.

The results illustrate that students who are less proficient in the L2 believe that the L1 is most effectively used as a tool for increasing class efficiency.

The data also revealed an overall trend of a negative correlation between the preference for L1 usage for class instructions and the level of L2 proficiency. In other words, the more proficient the students, the less likely they were to prefer class instructions communicated in the L2. This pattern is the same for student views of L2 usage for the facilitation of teacher/student rapport and humor and administrative tasks, once again illustrating the different preferences for L2 usage across differing levels of L2 proficiency. These results demonstrate that students with lower L2 proficiency levels feel they need the L1 to create a comfortable environment for learning and to aid in the functioning of the class. The findings in this study illustrate that the arguments presented by Tsukamoto (2011), Polio and Duff (1994), and Castellotti and Moore (1997), i.e., the use of L1 to create a comfortable atmosphere in the EFL classroom to initiate more effective learning, are perhaps even more relevant to students with lower levels of L2 proficiency.

Foreign EFL Teachers Pretending to Lack L1 Proficiency

Finally, students were questioned about whether foreign EFL teachers should pretend to lack L1 proficiency. The overwhelming majority of students wrote that the foreign EFL teacher should not pretend to lack L1 proficiency. Only 16.57% of students wrote the teacher should pretend, which is slightly higher than the ratio of students advocating for the L1 not to be used in the EFL classroom even when appropriate. These results illustrate that almost all students who prefer the foreign EFL teacher not to use the L1 believe that the teacher should pretend to lack proficiency in the L1. The most common reasons given were that a monolingual EFL classroom encourages L2 usage and improves learning. For example, one student with a TOEIC score of 210 wrote, "If Japanese is used our English won't improve." Another student with a TOIEC score of 703 was

concerned with becoming too reliant on the L1 during the EFL class saying, “Because the students will tend to depend on communicating in Japanese.”

The results also showed an increase in student preference for teachers pretending to lack proficiency in L1 as L2 proficiency amongst the students increased. However, the students with the highest-level L2 proficiency (TOEIC scores ranging from 751-900) bucked this trend and have a similar ratio of preference for teachers pretending to lack L1 proficiency as the lowest level L2 proficiency students (TOEIC scores ranging from zero to 300). This perhaps suggests that the highest-level L2 proficiency students are at a stage in their English language development where they are confident enough that L2 usage will not impinge on their learning.

The overwhelming majority of all students surveyed (and the majority of students from each language proficiency level) believe the foreign EFL teacher should not pretend to lack L1 proficiency. Ultimately, many students believe the foreign EFL teacher can use the L1 to enable communication when there is a misunderstanding and to enhance the class atmosphere. Moreover, it may be difficult for teachers and program coordinators to satisfy the minority of students that prefer monolingual lessons through the implementation of L2 usage parameters given the diversity of students in one EFL class. In practical terms, it is extremely difficult for a foreign EFL teacher to pretend to have a lack of understanding of the L1 for one student and not another. This decision will affect all students in the class.

The majority of student opinions with regards to foreign EFL teachers pretending to have a lack of L1 proficiency support recent arguments for a greater acceptance of L1 usage in the EFL class in creating a comfortable atmosphere conducive to learning (Castellotti & Moore, 1997) and its use as an important cognitive tool (Swain & Lapkin, 2000).

Conclusion

Approaches taken towards L1 usage in EFL classes in Japanese universities can vary from a strict monolingual system to flexible bilingual methods. Until recently there have been strong arguments made for the advantages of monolingual EFL lessons. For example, some researchers argue that a monolingual approach is the most effective as it emulates the method in which a child acquires their L1 (Butzkamm, 2003; Cummins, 1998; Pennycook, 1994; Phillipson, 1992). However, in recent years, there have been arguments made for the judicious and theoretically principled use of the L1 in the EFL classroom (Cook, 2001; Levine, 2003; Liebscher & Dailey-O’Cain, 2004; Turnbull, 2001).

This study examined the views of university students receiving L2 (English) education from a foreign EFL teacher regarding L1 usage in the EFL classroom. The students in this study were asked about their preferences for the foreign EFL teacher being fluent in the L1 (Japanese), using L1 when appropriate, the purpose of L1 usage, and whether a foreign EFL teacher should pretend to lack L1 proficiency.

There are several limitations to this study: first, the students surveyed carried out the TOEIC English proficiency tests at different times, which creates a degree of inconsistency in the comparison of L2 proficiency levels between students; second, the practical implementations of the results will be restricted by the diversity of the language proficiencies and learning preferences of each student within an EFL tertiary class; and third, there may be some debate concerning the significance of student views when it comes to the methodology of English language teaching in Japanese universities. Finally, students with lower levels of L2 (English) proficiency may also have lower levels of motivation given their lack of progression in the language since secondary school.

Most of the 175 university students surveyed (66.29%) preferred that the foreign EFL teacher was fluent in the L1 (Japanese). A common theme throughout the responses was that the L1 was a useful tool when students could not understand the L2 and it helped to facilitate more effective class communication. These results strengthen the theses supporting L1 usage in the EFL classroom. This study also found that a negative correlation exists between student preference for teacher L1 fluency and the L2 (English) proficiency of these students. The determination of a monolingual or bilingual approach to tertiary EFL courses in Japan would subsequently benefit from the recognition of the influences of English proficiency levels on the desired levels of

teacher L1 usage.

Students were also surveyed on their preferences of the usage of L1 when appropriate. The overwhelming majority (85.71%) believed that the L1 should be used in the EFL classroom when appropriate, further adding weight to arguments supporting bilingual approaches. Moreover, there is a negative correlation between the preference for L1 usage when appropriate and student L2 proficiency, highlighting the fact that the L2 proficiency of students should be taken into account when devising an approach to L1 application in the EFL classroom. According to these findings, lower-level L2 proficiency classes should accept a greater quantity of prudent L2 application.

Students were also questioned on their views on the specific purpose of L1 in the EFL classroom. This study revealed discrepancies in themes from students in different levels of L2 proficiency. Students with the highest level of L2 proficiency mostly identified new vocabulary as the most important area for L1 application while students with the lowest level of L2 proficiency mostly identified class instructions as the most important area for L1 application. These findings also suggest that the L2 proficiency of students in the EFL class should be considered when determining the objective of L1 usage.

Finally, students were asked about their opinions towards foreign EFL teachers pretending to lack L1 proficiency. The overwhelming majority of students (82.29%) believed that the foreign EFL teacher should not pretend or that pretending will have no effect on learning. These results are significant for English courses with policies prohibiting foreign EFL teachers from using L1 (Japanese) in classrooms. According to the overwhelming majority of the recipients of tertiary English (L2) education, a foreign EFL teacher should not pretend to lack L1 proficiency.

This study found that students of English (L2) education in Japanese universities overwhelmingly preferred the application of the L1 to aid in the facilitation of learning in EFL classes. Moreover, there were notable trends in attitudes between differing L2 proficiencies, with results revealing a negative correlation between desired L1 application in EFL classes and student L2 proficiencies. The findings also showed a discrepancy in the desired objective for L1 application for different student levels of L2 proficiency. The opinions analyzed in this study suggest that the approach adopted for EFL courses in universities (in Japan) should implement the judicious and theoretically principled application of L1 and recognize that different levels of L2 proficiency will affect preferences for L1 usage in the EFL classroom.

This paper recommends further research on how student preferences of L1 usage in EFL classes across different levels of English language proficiency vary for different English subjects; for example, subjects focused on English listening, reading, writing, or speaking. Further research into Japanese university student preferences for L1 usage in English classes using a larger sample size of students is also recommended to complement the findings of this study. The addition of more specific research regarding student preferences of L1 usage in EFL classes will provide important data in challenging the commonly held position within Japanese educational institutions of a strong emphasis on L2 exclusivity.

References

- Antón, M., & DiCamilla, F. J. (1998). Socio-cognitive functions of L1 collaborative interaction in the L2 classroom. *Canadian Modern Language Review*, 54(3), 314–342.
- Atkinson, D. (1987). The mother tongue in the classroom: a neglected resource? *ELT Journal*, 41(4), 241–247.
- Brooks, F. B., Donato, R., & McGlone, J. V. (1997). When are they going to say “it” right? Understanding learner talk during pair-work activity. *Foreign Language Annals*, 30(4), 524–541.
- Burden, P. (2000a). The Use of 'Only English' in a Learner-Centred University Classroom in Japan. *RELC Journal*, 31(1), 139–149.
- Burden, P. (2000b). The use of the Students' mother Tongue in Monolingual English “Conversation” classes at Japanese Universities. *The Language Teacher*, 24(6), 5–10.
- Butzkamm, W. (2003). We only learn language once, The role of the mother tongue in FL classrooms: Death of

- a dogma. *Language Learning Journal*, 28, 29–39.
- Butzkamm, W., & Caldwell, J. A. W. (2009). *The Bilingual Reform: a Paradigm Shift in Foreign Language Teaching*, Narr, Tübingen.
- Castellotti, V., & Moore, D. (1997). Alternances des langues et apprentissages: Revue de dialectologie des langues-cultures. Paris: Didier Erudition. In J. Rolin-Ianziti & R. Varshney (2006). Student perceptions of L1 use in the foreign language classroom: Help of hindrance. *Journal of Australasian Universities Modern Language Association*, 105, pp. 55–83, 148.
- Cook, V. (2001). Using the first language in the classroom, *Canadian Modern Language Review*, 57, 402–423
- Cook, V. (2005). Basing teaching on the L2 user. In: E. Llurda (Ed.), *Non-native Language Teachers: Perceptions. Challenges and Contributions to the Profession* (pp. 47–61), New York, Springer.
- Cummins, J. (1998). Immersion for the millennium: what have we learned from 30 years of research on second language immersion? In: M.R. Childs & R.M. Bostwick (Eds.), *Learning through Two Languages: Research and Practice*, pp. 34–47. Katoh Gakuen, Japan.
- Ellis, N. (2005). At the interface: Dynamic interactions of explicit and implicit language knowledge. *Studies in Second Language Acquisition*, 27(2), 305–352.
- Ellis, R. (1985). *Understanding Second Language Acquisition*. Oxford: Oxford University Press.
- Ellis, R. (2005). Principles of instructed language learning. *Asian EFL Journal*, 7(3), 9–24
- Gass, S. M., & Mackey, A. (2000). *Stimulated Recall Methodology in Second Language Research*. Mahwah, NJ: Erlbaum.
- Håkansson, G., & Lindberg, I. (1988). What's the question? Investigating questions in second language classrooms. *AILA Reviews*, 5, 101–116.
- Harbord, J. (1992). The use of the mother tongue in the classroom. *ELT Journal*, 46(4), 350–355.
- Kaneko, T. (1992). *The role of the first language in the foreign language classroom*. Temple University.
- Krashen, S. (1981). *Second language acquisition and second language learning*. Pergamon, Oxford.
- Krashen, S. (1982). *Principles and Practice in Second Language Acquisition*. Pergamon, Oxford.
- Kroll, J. (1993). Assessing conceptual representations. In R. Schreuder & B. Weltens (eds), *The Bilingual Lexicon* (pp. 249–277). Amsterdam/Philadelphia: John Benjamins.
- Levine, G. S. (2003). Student and instructor beliefs and attitudes about target language use, first language use, and anxiety: Report of a questionnaire study. *Modern Language Journal*, 87(3), 343–364.
- Libben, G. (2000). Representation and processing in the second language lexicon: The homogeneity hypothesis. In J. Archibald (ed.) *Second Language Acquisition and Linguistic Theory* (pp. 228–248), Oxford: Blackwell.
- Liebscher, G., & Dailey-O'Cain, J. (2004). Learner code-switching in the content-based foreign language classroom. *Canadian Modern Language Review*, 60(4), 501–525.
- Littlewood, W. (1992). *Communicative language teaching*. Cambridge: Cambridge University Press.
- Littlewood, W., & Yu, B. (2011). First language and target language in the foreign language classroom. *Language Teaching*, 44(1), 64–77.
- Macaro, E. (1997). *Target Language, Collaborative Learning and Autonomy*. Clevedon: Multilingual Matters.
- Macaro, E. (2001). Analysing student teachers' codeswitching in foreign language classrooms: Theories and decision making. *Modern Language Journal*, 85, 531–548.
- Macaro, E. (2005). Codeswitching in the L2 Classroom: a Communication and Learning Strategy. In: E. Llurda (Ed.), *Non-native Language Teachers: Perceptions, Challenges and Contributions to the Profession* (pp. 63–84). New York, Springer.
- Makulloluwa, E. (2013). Code switching by teachers in the second language classroom. *International Journal of Arts & Sciences*, 6(3), 581–598.
- McMillan, B., & Rivers, D. (2011). The Practice of Policy: Teacher Attitudes toward "English Only". *System: An International Journal of Educational Technology and Applied Linguistics*, 39(2), 251–263.
- Pennycook, A. (1994). *The Cultural Politics of English as an International Language*. Longman, London.
- Phillipson, R. (1992). *Linguistic Imperialism*. Oxford University Press, Oxford.
- Polio, C. G., & Duff, P. A. (1994). Teachers' language use in university foreign language classrooms: A

- qualitative analysis of English and target language alternation. *Modern Language Journal*, 78, 313–326.
- Rolin-Ianziti, J., & Brownlie, S. (2002). Teacher use of learners' native language in the foreign language classroom. *The Canadian Modern Language Review*, 58(3), 402–426.
- Rolin-Ianziti, J., & Varshney, R. (2006). Student perceptions of L1 use in the foreign language classroom: Help of hindrance? *Journal of the Australasian Universities Modern Language Association*, May, 105, 55–83.
- Rolin-Ianziti, J., & Varshney, R. (2008). Students' Views Regarding the Use of the First Language: An Exploratory Study in a Tertiary Context Maximizing Target Language Use. *The Canadian Modern Language Review*, 65(2), 249–273.
- Swain, M., & Lapkin, S. (2000). Task-based Second Language Learning: the Uses of the First Language. *Language Teaching Research*, 4(3), 251–274.
- Timor, T. (2012). Use of the mother tongue in teaching a foreign language. *Language Education in Asia*, 3(1), 7–17.
- Tsukamoto, M. (2011). Students' perception of teachers' language use in an EFL classroom, *Osaka Jogakuin Daigaku Kiyō*, 8, 143–154.
- Turnbull, M. (2001). There is a role for the L1 in second and foreign language teaching but... *The Canadian Modern Language Review*, 57, 531–540.
- Turnbull, M., & Dailey-O'Cain, J. (2009). *First language use in second and foreign language learning*. Clevedon: Multilingual Matters.

A Longitudinal Analysis of Ability Grouping with College EFL Learners

Chiu-hui (Vivian) Wu

Wenzao Ursuline University of Languages, Kaohsiung, Taiwan

Chia-jung Tsai

Wenzao Ursuline University of Languages, Kaohsiung, Taiwan

Yi-Min Chiu

Wenzao Ursuline University of Languages, National Cheng Kung University, Taiwan

Abstract

Ability grouping, organizing classes homogeneously by L2 proficiency, has been commonly used in Taiwanese English as a Foreign Language (EFL) classes. This quasi-experimental (within-subjects design) study examined proficiency gains of 785 Taiwanese university students over three years enrolled in a general English (GE) program that employed ability grouping. The standardized test used for this study was the *College Students English Proficiency Test (CSEPT)*. The results indicated students gained in English proficiency over time, from entry into the program to their last year of English instruction. Further post hoc analysis of the long-term proficiency changes showed that students with an observed A2 (CEFR) proficiency, upon entry, had more pronounced gains, over the three years, than their A1 and B1 counterparts. The study concluded that a leveled English curriculum maximized the learning experience for A2 level students and allowed them continuous proficiency gains. However, the fact that B1 level students did not show consistent progress is perhaps due to plateau effect when their test scores hit the graduation benchmark. As for the A1 students, their lack of achievement may be due to their low self-esteem. The pedagogical implication suggests the need to revisit the leveled (ability grouping) English curriculum for A1 and B1 level learners.

Keywords: ability grouping, general English curriculum, graduation benchmark, college English, language program design

Introduction

Over the past two decades, Taiwan has attempted to increase its international participation in the global market by prioritizing English language education through a national development plan (Chen & Hsieh, 2011). It is believed that increased English proficiency of Taiwanese citizens would give them greater opportunity to participate in international affairs (Chen, 2011). Thus, the status of English in Taiwan has shifted from being a foreign language to being a quasi-official language, which is illustrated by the fact that signs in English are used in many public places (Chen, 2011; Feng, 2012). In addition, to better prepare citizens and students for English proficiency—and hence internationalization—the Ministry of Education (MOE) in Taiwan has initiated and implemented policies for English curriculum reform. For instance, MOE recently made EFL courses compulsory starting at primary school (grades 3 to 6) rather than secondary school (grades 7-12), (Chen & Hsieh, 2011; Chern, 2002, 2010).

*E-mail: joycecjsai@hotmail.com, Tel.: 886-922452332, Address: Center for English Language Teaching, Wenzao Ursuline University of Languages, 900 Mintsu 1st R. Kaohsiung 807, Taiwan R.O.C.

Regarding tertiary or higher education, the MOE has set expectations for English curricula and language policies for some time (Hua & Beverton, 2013; Pan & Newfields, 2012). In the 1990s, the MOE began to move universities away from exclusively reading in English curricula to more comprehensive learning plans (e.g., emphasis on production). In 2003, the MOE further urged universities and colleges to set English graduation benchmarks and left the choice of how the benchmarks would be measured (e.g., standardized testing) to the universities (Pan & Newfields, 2012). In today's universities, GE programs are how universities prepare their students to reach the English proficiency benchmark prior to their graduation.

To better understand the GE programs implemented across Taiwan, Chern (2010) studied the programs at 60 universities, including public and private. Findings from the study revealed that 32 universities required students to take a one-year English course during their freshmen year, approximately four credit hours, and 20 universities required students to take English courses for two years with a total of four to six credit hours. Based on the findings from diverse GE programs implemented in Taiwan, Chern (2010) concluded that a systematic examination was needed to determine if the curriculum prepared students to meet English language proficiency benchmarks.

One widely utilized approach believed to be effective at the tertiary level of EFL education was ability grouping. When this strategy is implemented, students are placed in different levels of GE groups based on their English proficiency. Research in EFL contexts on the effects of ability grouping found positive results for college freshmen (e.g., Khazaenezhad, Barati, & Jafarzade, 2012; Kulik, 1992; Liu, 2008). Kulik (1992) contended that it would be a mistake if schools abolished ability grouping. Yet, the controversy and debate regarding its effectiveness has continued.

This current study was conducted to investigate how students of different English proficiency upon entry would progress over three years of ability grouping instruction. Because of university policies, a control group and/or other groups receiving non-ability grouping instruction was not possible. This lack of an experimental design meant that the findings of the current study could not facilitate claim a direct casualty between ability grouping and proficiency gains. The rationale for this current study, instead, was grounded in a desire to understand how ability grouping could have influenced the proficiency gains of students of different proficiency levels upon entry. In other words, this study contributes to an ongoing dialogue while also being aimed at inspiring future research that could address its unavoidable design limitations.

Literature Review

Ability grouping in language education

Ability grouping refers to the practice of placing students in a classroom or small groups based on ability or achievement. This is usually done by assessment of ability with standardized tests (Kim, 2012). This teaching and program design strategy has been used in education, especially in primary and secondary schools, since the 20th Century (Slavin, 1987). The earliest reviews regarding ability grouping were found in the 1920s and early 1930s (Kulik, 1992).

Subsumed under ability grouping are two types: (1) within-class and, (2) between-class grouping (Ireson & Hallam, 2001). Within-class grouping, or mastery learning, is usually practiced in a class, and students of different perceived levels are assigned to groups for specific or adaptive instruction to accommodate their learning needs (Ireson & Hallam, 2001). Between-class grouping, by comparison, is a school-level practice that places students in different ability groups or tracks by class (Ireson & Hallam, 2001). The current study specifically looked at the between-class grouping model for ability grouping with respect to EFL students.

Between-class grouping model for ability grouping

Ability grouping has been widely adopted in pre-secondary and secondary English language education in several countries such as the UK (Hallam & Ireson, 2003; Ireson, Hallam, Hack, Clark, & Plewis, 2002), and the USA (Slavin, 1990), and Korea (Jung, 2000; Kim, 2012). Previous research, nevertheless, has yielded divergent results regarding the effect of ability grouping on English proficiency (L2) attainment.

It has been evident in some survey-based studies with teachers that ability grouping facilitated English language teaching and learning, yet it catered more to the needs of higher-level learners than to those of their lower-level counterparts (Hallam & Ireson, 2003; Ireson et al., 2002). Hallam and Ireson (2003), for example, researched secondary school teachers' attitudes toward and beliefs about ability grouping. Their sample was comprised of more than 1,500 teachers from 45 schools in the UK. They found overall agreement among the teachers that ability grouping ensured maximum learning outcomes for the most advanced students.

Additionally, Hallam and Ireson (2003) found strong agreement that ability grouping was beneficial, especially for teachers, because it was advantageous for student learning, while also making class management easier. Furthermore, when using ability grouping, it was possible for teachers and programs to better design curriculum to meet the needs of a variety of students. Regarding subject matter, ability grouping was seen as specifically beneficial when it came to the disciplines of mathematics and foreign languages.

Other studies (e.g., Kim, 2012; Slavin, 1990), in contrast, reported ability grouping to be only slightly beneficial or ineffective. According to Slavin's (1990) review on 29 (experimental, correlational, or case) studies of between-class ability grouping for junior high and high school students, no positive effects on student achievement were observed. This review of ability grouping, in the US, included students in various courses over a period of five years.

Kim (2012) found that there was not a positive attitude among students toward between-class ability grouping. A survey was administered as part of the study to 754 students from six different Korean middle schools (grades 7-9). Due to a variation in ability grouping practices among the schools participating in the study, Kim focused on only three comparable schools. Findings revealed that in two schools with three group levels (high, intermediate, low), higher-level students' responses to between-class ability grouping were slightly positive or neutral, while lower-level students were neutral or negative about its effectiveness. Moreover, in one school with two group levels, both high and low-level students reported a negative attitude toward between-class ability grouping.

While the findings of ability grouping research in secondary contexts across various disciplines have been divergent, studies done in post-secondary EFL settings tended to observe positive results (Wen, 2011). Khazaenezhad, Barati, and Jafarzade (2012) conducted an experimental design inquiry using test and control groups to examine the effectiveness of ability grouping on college-level English language learners in Iran. The study investigated ability grouping (less-able, intermediate, and advanced groups) and various amounts of exposure to English (two, three, and four hours) in relation to academic gains in one semester. The study recruited 320 non-English major undergraduates and divided them into *different ability* groups and *non-ability* groups. Findings indicated that the students in the ability groups significantly outperformed their counterparts in the different ability groups as exemplified by their test scores. This clearly revealed the positive effects of ability grouping on the subjects' academic gains in GE training.

Ability grouping implemented in GE training in Taiwan

Specific to the Taiwanese EFL context, ability grouping has been a popular policy in secondary education and widely advocated and practiced by many universities and colleges (Chern, 2010; Feng & Chang, 2010; Lee & Su, 2009; Wen, 2011). Some studies have reported a positive impact on learning and a positive attitude from students regarding the effectiveness of ability grouping in the GE courses (Lee & Su, 2009; Liu, 2008; Wen, 2011).

For example, ability grouping was positively perceived by university instructors, as well as by students (Liu, 2008). In a survey, Liu investigated the perceptions of 582 freshmen and sophomores and 34 English teachers at university in central Taiwan. The focus of the survey was to measure the participants' attitudes toward ability grouping. The participants were divided into the following groups based on their scores from the *General English Proficiency Test* (GEPT) when they enrolled at the university: (a) basic, (b) intermediate, and (c) advanced. The GEPT is an English proficiency assessment designed by the *Language Training and Testing Center* (LTTC) of Taiwan to measure citizens in four skills of English proficiency. The four skills assessed were: (1) listening; (2) speaking; (3) reading; and (4) writing. From the learners' perspective, the results demonstrated that freshmen held positive attitudes toward ability grouping, particularly those with basic English proficiency. These students reported that

working with students of similar ability reduced the pressure and anxiety of learning and enhanced their motivation. However, positive attitudes toward ability grouping weakened by the end of sophomore year.

In a semester-long project, Lee and Su (2009) studied 2,230 non-English-major students from a technical university in Taiwan with respect to ability grouping. The participants were leveled into three ability groups based on English proficiency: (1) beginning, (2) intermediate, and (3) higher intermediate. The aim of the project was to compare achievement scores before and after taking the one-semester, freshman English course. Results from the students' achievement tests indicated a significant difference between pre- and post-test scores, with the intermediate level students making the greatest progress. Yet, the short intervention time (one semester) and lack of detailed description on leveled instruction lessened the validity of the findings.

Likewise, the purpose of Wen's (2011) study was to examine the effects of ability grouping on technical university students' general English learning achievement in a year-long, two-semester, program. The subjects in this study consisted of 792 freshmen from three colleges at one university (business, engineering, and electronics and information) who were divided into three ability groups: (1) high achievers, (2) medium achievers, and (3) low achievers. Ability grouping was based on students' English scores from the *Joint College Entrance Exam* (JCEE), a regular (non-technical) university entrance exam. All the participants took one pretest (listening and reading) before the year began and two posttests (listening and reading) at the end of the GE course. The findings indicated that low-achieving students did not benefit from ability grouping, but students in the medium and high groups showed significant progress on listening and reading scores.

Some studies related to the two previously discussed research lines—English as a native language (ENL) (Slavin, 1990) studies and EFL studies (Kim, 2012; Trautwein, Koller, & Kammerer, 2002)—voiced concerns about the potential negative consequences caused by the implementation of between-class ability grouping. These concerns were raised because students who were less proficient in EFL were deprived of what could be better instruction because the teachers had lower expectations of them compared to their more proficient counterparts (Kim, 2012). Kim (2012) also reported that between-class ability grouping often had adverse effects because it widened the gap between high- and low-level learners. Kim (2012) concluded that the effectiveness of ability grouping was determined by how it was implemented (e.g., the number of group levels) and if it was supported by other school policies. In studying the effects of ability grouping on students grades 6 to 9 in EFL and math classes, Trautwein, Koller and Kammerer (2002) found that the between-class ability grouping enhanced lower ability students' academic involvement in class.

These concerns about ability grouping, however, have been addressed by researcher such as Khazaenezhad et al. (2012) and Wen (2011). They suggested that the negative possible effects of ability grouping could be mitigated through careful planning and decision making where all the different agents in the teaching and learning process collaborated. Policy makers are tasked with the responsibility for how ability grouping is implemented, keeping in mind the different levels with respect to curriculum design, materials development. Teachers should be trained in how to deliver the instruction at the specific level to which they are assigned.

Taken together, two significant gaps exist in the above-discussed research. First, little empirical research to date has explored ability grouping and proficiency gains over time. Among the limited EFL studies done (e.g., Khazaenezhad et al., 2012; Wen, 2011), most were done during a period of one semester or one academic year, thereby contributing limited information about the short-term effects and not exploring the long-term effects. Second, relevant studies have used different proficiency measurements where subjects varied by absolute proficiency level. Lee and Su's (2009) and Wen's (2011), for example, were different in relation to their tests and subject proficiency level. Comparing their findings would therefore be problematic. Perhaps a more widely used reference framework could provide researchers common reference points for students' proficiency levels. These issues indicate the need for (a) long-term studies with a systematic examination of how universities' GE curriculum prepares students to meet language proficiency requirements and (b) the need for using a common reference framework to ensure a consistent interpretation of students' proficiency levels.

To address these research gaps in the literature, the current study conducted a longitudinal study in Taiwan within the context of EFL programs. Specifically, it adopted a quasi-experimental within-subject design and

interpreted students' proficiency levels based on the Common European Framework of Reference for Languages (CEFR – Council of Europe, 2011).

Research Questions

Having noted the aforementioned gaps in the research, this quasi-experimental study examined whether or not students enrolled in a longitudinal, ability-grouping GE curriculum improved over time at a university in Southern Taiwan (hereafter SU). The study proposed the following research questions:

1. How do the observed English (L2) proficiency scores of a group of Taiwanese EFL university students change over time where their GE curriculum was designed around ability-grouping principles?
2. How do the observed English (L2) proficiency scores of a group of Taiwanese EFL university students who were observed to have an A1, A2, or B1 CEFR level upon entry to the university change over time where their GE curriculum was designed around ability-grouping principles?

RQ2 was posed as consisting of three separate hypotheses and as a post hoc of RQ1.

Methodology

Subjects

This study used the three-year CSEPT test records from 785 students at SU. The subjects were first enrolled in the 2012 academic year and received a three-year-long intervention of leveled GE (ability grouping) instruction from Fall 2012 to Spring 2015. Because the study was designed as a longitudinal study, these 785 subjects took the pretest, the one-year posttest, the two-year posttest, and the three-year posttest. In other words, any subjects who did take the CSEPT these four times were excluded. The informed consent for each student was obtained prior to taking the pretest.

At the beginning of the subjects' freshman year at SU, they took the CSEPT pretest for placement purposes. Based on the CSEPT scores, all freshmen were grouped within Levels One to Eight. Table 1 illustrates the range of scores from Levels One (lowest) to Eight (highest), their equivalence on CEFR, and the number of subjects at each level. In addition, Level 7 subjects (N=60), equivalent to B2 level on CEFR, were excluded from the analysis for the second research question given that they had reached graduation benchmark and accounted for a small percent of the total sample. On the basis of CEFR, of the 725 subjects (excluding 60 B2 level subjects), 110 were observed to have an A1 level, 223 an A2 level, and 392 a B1 level upon entry to the university.

College Student English Proficiency Test

This study adopted the *College Student English Proficiency Test* (CSEPT) as the initial reference points for placing students into different levels/groups by proficiency, which were converted to CEFR later on for data analysis purpose. The CSEPT, designed by the *Language Training and Testing Center* (LTTC)^{see Endnote 1} for higher education institutes in Taiwan, is an English proficiency test for EFL college students. The purpose of the test was to evaluate university students' English proficiency; primarily targeting students' receptive skills including listening, reading and grammar. The test fulfills the need of analyzing the outcomes of English language teaching and learning. The Primary Level CSEPT was made available in 1997 followed by the Secondary Level in 1998 (LTTC, 2007). Table 1 presents the measurements of the CSEPT and its equivalent, the *Common European Framework of Reference for Languages* (CEFR, Council of Europe, 2011) as illustrated by LTTC (n.d.). The primary level is the equivalent of CEFR B1, a level at which the test questions measure intermediate level of proficiency. The secondary level is the equivalent of CEFR B2 and measure the English proficiency of intermediate to advanced level learners. The test has been adopted by many technical schools and colleges and universities in Taiwan (Pan & Newfields, 2012). It is intended to measure language learners' receptive skills, such as listening and reading proficiency within the context of everyday and campus life.

Table 1
Level Groups and Their Proficiency at SU for Freshmen

General English at SU				
CEFR	SU's GE Level	Student's CSEPT Score	Subjects (N)	Notes
A1	1	~119	110	Level 1 – Level 4: Extra 2 hours of remedial instruction; Self-access to learning resources
A2	2	120~144	111	
	3	145~169	112	
B1	4	170~200	119	
	5	201~219	166	
	6	220~239	107	
B2	7	240~259	60	Benchmark for non-English majors
	8	260~344	NA	Benchmark for English related majors
Total:			785	

Notes.

1. During the time of this study, a level nine course was not offered and was not implemented until 2016 for freshman.
2. English related majors include: English, Foreign Language Instruction and Translation and Interpreting.

Adopting CSEPT at SU

In 1997, SU was one of the universities to adopt CSEPT as a mandatory test for all students to measure their gains in English proficiency. The university adopted CSEPT for three reasons: first, to place students in leveled classes based on their proficiency; second, to document learners' language proficiency so that SU could constantly evaluate the effectiveness of its language curriculum; and third, as an English proficiency benchmark for students to fulfill as a partial graduation requirement. The secondary-level CSEPT test was used for all students at SU and administered to the subjects of this study. This CSEPT test has three sections: listening, grammar, and reading. First, in the listening test, students listen to and understand short conversations in addition to short speeches. The listening test includes a total of 30 questions. For the grammar test, students are required to complete sentences and short passages that consist of 50 questions. Finally, the reading test consists of 30 reading comprehension questions. The total time allowed to complete the CSEPT test is 90 minutes.

In this study, the CSEPT tests were officially administered by the LTTC at SU when the subjects attended a mandatory summer camp before their first semester. The actual CSEPT scores collected during the summer camp were considered the pretest scores (T1). Near the end of the first (T2), second (T3), and third (T4) years of the GE training program, the official CSEPT tests were administered by LTTC as posttests to measure the students' progress in English.

Southern University (SU) and its ability-grouping GE curriculum

Founded in 1966, SU is known for its foreign language pedagogy, with a vision that all students will demonstrate English proficiency to complement their knowledge in their respective majors, such as communication arts, digital content application, international business, international affairs, foreign language teaching, etc. SU believed that through foreign language learning, students would be able to understand global culture and expand their world views. Thus, each college student was required to take an adaptive three-year GE program before they graduated. The program was designed to ensure students' English language proficiency by the time they exited the program. For example, students at SU were eligible to become exempt from some credit hours as soon as they completed the highest level of English proficiency (level nine) or when they demonstrated high English proficiency (CSEPT test score over 345). For example, if a student's level of English proficiency was at eight when admitted to SU, he would be required to take two years of EFL to exit the program. In other words, he would only have to complete 16 credit hours. As for level upgrading, two rules applied. The first was to upgrade

students to one level automatically after one year of GE training, regardless of their updated CSEPT scores. The other was that students could apply to be upgraded, to the appropriate level, based on their updated CSEPT scores.

The regular GE program design was conceptualized in a student proficiency-based teaching philosophy and embraced Communicative Language Teaching (CLT) as an approach. The program offered the integration of four skills- listening, speaking, reading, and writing- and leveled materials for each skill. The teaching delivery varied, including listening and speaking training, simulated dialogues, reading skills, and writing practices.

General English is a required subject for all students at SU. Beginning in 2013, SU required students to complete a total of 24 credits (approximately six courses with each class worth four credits) in English within the first three college years. However, this excludes an additional two non-credit hours of remedial instruction per week for students under Level Five during freshman and sophomore years. This consideration was based on the assumption that students' exposure to English would gradually increase their EFL proficiency. It is noteworthy that the total hours required by SU were exceptionally high compared to other universities in Taiwan. For example, Chern (2010) reported that a range of four to six credit hours (usually two to three courses) was a common requirement at many universities in Taiwan.

Although the program at SU requires students to take 24 credit hours, not all students receive the same amount of English training. The primary feature of the GE curriculum at SU is that it is adapted to students' level of English proficiency, measured by a recognized English proficiency test. Students whose CSEPT pretest is under 200 (equivalent to CEFR A1 and A2 levels) receive additional two hours of remedial instruction per week, whereas students whose entry level is Level Nine (equivalent to CEFR B2 level) were only required to take 1 year of GE instruction. This allowed those students to take advanced English or English as a Medium Instruction courses as electives. Table 1 illustrates the group levels and the entry levels of freshmen students at SU. Students whose levels were under five in their freshman and sophomore years (i.e., CSEPT test score below 200; the CSEPT test will be introduced later in the text) received an additional two hours per week, for remedial instruction.

The language curriculum for the subjects was tailored to meet their different needs. In addition to the remedial hours, SU provided each student self-access to language learning consultation and resources in the Language Diagnostic and Consulting Center (LDCC). In the LDCC, students can consult teachers about their learning styles and strategies, as well as practicing language with computer assisted learning programs. In order to provide incentives for students to study English on their own time, the record of students' self-access learning progress was considered part of their overall course performance.

To ensure teaching quality and consistency, SU implemented a structured curriculum with the same textbooks being used by all teachers at each level as determined by the level coordinator in consultation with instructors. Exams were also created and administered in a similar fashion. For each level, teachers were expected to be consistent with their content materials and assessments. The CSEPT washback effect was minimal because SU did not tailor the curriculum to prepare students to take the CSEPT, as the school-based exams evaluated both the receptive and productive skills of language learners, including speaking and writing. Every semester, faculty meetings were held several times a semester for staff to discuss their teaching with other colleagues, including the authors of the study.

Data Collection and Analysis

The subjects took the official CSEPT administered by LTTC as a pretest for placement purposes at the beginning of their freshman year. After approximately one year of GE training, in May, at the end of their freshman year, the subjects took an alternate version of CSEPT for the one-year posttest. After two years of GE training, the subjects took another CSEPT for the two-year posttest. Finally, in a three-year English program, the subjects took the last official CSEPT, the three-year posttest. The subjects' pretest scores were regarded as their English proficiency before the intervention of GE classes at SU. The posttest scores were considered a measure of the subjects' progress in English after taking the GE classes.

To analyze the data that had a within-subjects design, one-way repeated measures ANOVAs were performed to investigate whether all the subjects' observed English proficiency scores changed significantly over time (RQ1). This same analysis was then done on three groups from the sample to answer RQ1: A1 upon entry, A2, B1. Main effect sizes (time and proficiency scores), for RQ1 and RQ2, were reported via partial-eta-squared (Lakens, 2013). Post hoc pairwise comparisons were employed to assess significance and effect size (via Cohen's *d*-average) of differences between two measurements, e.g. pretest and year one posttest. The magnitude thresholds for *d*-average are the same as with Cohen's *d* (see Cohen, 1988: .2-small - .5-medium - .8-large) way repeated measures ANOVA were referred to as a within-subject ANOVA for the same group of subjects. Since RQ2 was framed as 3 independent hypotheses, α -level for statistical significance left at .05. For all ANOVAs, the Mauchly's tests were significant ($p_s < .01$), indicating that the sphericity assumption was violated, Greenhouse-Geisser corrections were therefore applied.

Results

First Research Question

Table 2 presents the descriptive statistics of subjects' proficiency scores over time. Via one-way repeated measures ANOVA testing, a significant association was observed among the four proficiency scores across time [$F(2.53, 1982) = 223.34, p < .01$, partial-eta-squared=.22]. There were 5 significant ($p_s < .01$) observed pairwise post hoc comparisons: T4 > T3 (d-average=.1); T4 > T2 (= .12); T4 > T1 (= .42); T3 > T1 (= .33); T2 > T1 (= .3). T3 > T2 ($p = .14$; d-average=.02) was nonsignificant.

Table 2
Descriptive Statistics of the Scores over Time

	<i>M</i>	<i>SD</i>	<i>N</i>
Pretest	188.82	57.27	785
Freshman posttest	206.12	57.00	785
Sophomore posttest	207.49	57.10	785
Junior posttest	213.00	56.47	785

Second Research Question

The second research question was analogous to the first except for the creation of 3 independent samples based on observed CEFR proficiency level upon entry into the program: A1, A2, B1.

Table 3 presents the descriptive statistics of the A1 group's proficiency scores over time. Via one-way repeated measures ANOVA testing, a significant association was observed among the four proficiency scores across time [$F(2.19, 238.22) = 39.98, p < .01$, partial-eta-squared=.27]. There were 5 significant ($p_s < .01$) observed post hoc comparisons: T4 > T2 (d-average=.32); T4 > T1 (= .92); T3 > T2 (= .22); T3 > T1 (= .84); T2 > T1 (= .64). T4 > T3 ($p = .21$; d-average=.12) was nonsignificant.

Table 3
Descriptive Statistics of the Scores: A1 Group

	<i>M</i>	<i>SD</i>	<i>N</i>
Pretest (T1)	103.08	26.22	110
Freshman posttest (T2)	120.73	30.95	110
Sophomore posttest (T3)	127.67	31.93	110
Junior posttest (T4)	131.56	35.69	110

Table 4 presents the descriptive statistics of A2 group's proficiency scores over time. Via one-way repeated measures ANOVA testing, a significant association was observed among the four proficiency scores across time [$F(2.58, 572.49) = 111.71, p < .01$, partial-eta-squared=.34]. There were 6 significant ($p_s < .01$) observed post hoc

comparisons: T4 > T3 (d-average=.25); T4 > T2 (= .42); T4 > T1 (=1.17); T3 > T2 (= .15); T3 > T1 (= .87); T2 > T1 (= .79).

Table 4
Descriptive Statistics of the Scores: A2 Group

	<i>M</i>	<i>SD</i>	<i>N</i>
Pretest (T1)	147.91	23.53	223
Freshman posttest (T2)	168.73	28.90	223
Sophomore posttest (T3)	173.43	34.76	223
Junior posttest (T4)	182.25	35.28	223

Table 5 presents the descriptive statistics of B1 group's proficiency scores. Via one-way repeated measures ANOVA testing, a significant association was observed among the four proficiency scores across time [$F(2.55, 995) = 76.32, p < .01, \eta^2 = .16$]. There were 5 significant ($p_s < .01$) observed post hoc comparisons: T4 > T3 (d-average=.16); T4 > T2 (= .1); T4 > T1 (= .65); T3 > T1 (= .48); T2 > T1 (= .61). T2 > T3 was nonsignificant with a direction contrary to expectation.

Table 5
Descriptive Statistics of the Scores: B1 Group

	<i>M</i>	<i>SD</i>	<i>N</i>
Pretest (T1)	222.15	24.93	392
Freshman posttest (T2)	237.99	26.61	392
Sophomore posttest (T3)	235.86	32.75	392
Junior posttest (T4)	240.98	32.64	392

Discussion

This study first aimed at uncovering whether the students made progress over time after enrolled in the three-year GE training program with ability grouping between-class. The findings demonstrated significant gains in the subjects' CSEPT scores from the first year to the third year. Time and proficiency scores shared 22% of the variance, and post hoc comparisons revealed 5 significant differences where the later test score average was higher. These observations suggested that students, in the aggregate, had made somewhat continuous progress over time since their first-year enrollment in the three-year GE program designed based on ability grouping principles at SU. This finding was constrained and limited by the lack of a comparison with a control group or non-ability grouping treatment group.

Ability grouping supporting L2 proficiency attainment over time was also suggested by the existing literature in domestic (Wen, 2011) and international contexts (Khazaenezhad et al., 2012). Significant gains were found in the majority of students who received long-term leveled GE instruction. Unlike Wen (2011) who focused on the effect of one-year leveled GE instruction, the current study demonstrated gains over three years. Therefore, this inquiry has contributed to the case for ability grouping in both Taiwanese and other contexts.

In spite of the proficiency gains, students' motivation may have lessened after the first year of GE instruction (e.g., T3 > T2 – nonsignificant; T4 > T3 – d-average=.09/very weak effect) due to the lack of integrated, as well as instrumental, motivation. Warden and Lin (2000) posited that Taiwanese students at a technical college had undergone this very phenomenon. As pointed out by Hua and Beverton (2013), GE courses in Taiwan were made compulsory to increase the nation's global competitiveness. However, if the courses did not relate to their major field of study throughout the program's duration, learning English might not have offered any specific value to the students. The subjects of this study might not have seen the value of their efforts and eventually lost interest in learning English.

L2 Proficiency Gains across Different L2 Proficiency Levels upon Entry

This study also proposed to investigate how groups at different CEFR levels upon entry varied in relation to proficiency gains. The A2 group as presented above had the strongest gains over time as evidenced by all 6 possible post hoc comparison being significant, in the expected direction and the highest observed partial-eta-squared. The A1 group also made gains, but they were not as pronounced as the A2 group. The B1 group had the weakest observed effect between time and proficiency scores with one post hoc comparison (T2 > T3) being in a direction contrary to expectation. This difference was nonsignificant however.

A similar result found in Kim's (2012) study was also observed regarding the subjects with an A1 entry level in this study. Kim (2012) found that mid- and lower-level students struggled with ability grouping and suggested that ability grouping alone would not lead to significant improvements in students' English proficiency. In order for ability grouping to create an environment where significant gains could be made, Kim claimed that it would be necessary to have a combination of curricula that corresponded to the students' learning styles, interests, and abilities.

Lastly, Level A1 and B1 students' attitudes when taking the CSEPT may have influenced the findings. A1 students' lack of achievement may be due to their low self-esteem as they probably knew they were the lowest group. On the other hand, perhaps a plateau effect occurred among B1 learners. As most of the students whose entry level was B1 passed the required graduation benchmark (240 for non-English majors – B2) in the first year of GE learning, these subjects may not have been taking the subsequent tests seriously, leading to underperformance.

Conclusion

Adopting a quasi-experimental (within-subject) design, this study contributed to the understanding of the observed longitudinal language gains of learners who received General English (GE) instruction designed around ability-grouping principles. EFL students with an A2 entry level experienced an ongoing progress when the stratified English curriculum was adaptive to offer remedial instruction and to challenge their current level by upgrading annually. Below are implications for EFL education and suggestions for future research.

Implications for EFL education

Ability grouping has been commonly used in EFL college settings including Taiwan, particularly when implementing GE courses. The implication is that policy makers need to re-conceptualize ability grouping as a way to increase language proficiency (Lee & Lin, 2013). The ultimate goal of ability grouping is not to widen the proficiency gap among language learners but to offer different kinds of scaffolding for different levels of students. Therefore, a well-designed leveled (ability grouping) curriculum helps learners to challenge their current levels. Rather than seeing students' diverse levels as a problem, teachers and policy makers can regard it as an opportunity to make the curriculum more adaptable for learners at all levels. As Kim (2012) noted, the effects of ability grouping can be enhanced or lessened depending on materials used, teaching hours, assessments, and resources provided by a university. Administrators need to consider how the leveled curriculum is implemented and adapted, and for what purpose. A leveled curriculum requires an integration of school-related learning resources such as remedial instruction and self-access learning into the curriculum. It also allows teacher collaboration within the same level to share their experiences. This allows the school to provide various accommodations for the needs of students with differing proficiency levels.

The current study also illustrated how A2 upon entry students, as a group, had progressed most since their enrollment in a long-term program designed based on ability grouping. Perhaps, the supplementary remedial intervention and the use of the school's self-access learning resources worked best for them. The effect of these types of resources can be positive for many students with a similar entry level. However, the language progress was least pronounced for B1 upon entry students, particularly those who have reached the English proficiency benchmark for graduation. These students may be more focused on their professional studies rather than concentrating on English language learning.

Implications for Future Research

As this study was quasi-experimental in design, future researchers may consider conducting studies involving a control group to compare gains via an experimental design. These studies could also include other treatment groups to compare other class organization strategies, e.g., heterogenous (in relation to proficiency) grouping. Like this study, future studies may continue to use CEFR to systematically report learners' levels of proficiency in their research designs. By doing so, researchers can make cross-study comparisons to show the effects of leveled GE instruction in various contexts. Finally, as the findings of this study suggest, A2 learners progress more than those of the other levels, future studies could further investigate the phenomenon of why A2 learners make smooth gains, whereas those of the other levels do not. Studies could also delve into the ceiling effect for B1 learners as implied by this study. Qualitative studies with interviews or observations could be conducted to explore learners' perceptions of effectiveness of ability grouping as a curricula and program design scheme.

Limitations of this study

Our design for the ability groups has several limitations. First, the CSEPT only evaluated students' listening, reading and grammar (usage norms) proficiency. In contrast, the course design and materials at SU were integrated with all four language proficiency skills (speaking, writing, listening, reading). Although the GE courses at SU placed a strong emphasis on the subjects' speaking and writing skills, students' production skills were not measured by the CSEPT. Second, the program's automatic progression, for level X to X + 1, regardless of CSEPT score could have influenced the observed findings. The final and, perhaps, most important limitation was the lack of a control group. The findings of this current study cannot be used to suggest direct causality.

Endnotes

1. The CSEPT's psychometrics have been assumed as credible for some time in the Taiwanese context given its long history of development. The CSEPT's governing body, the LTTC is partners with several international English proficiency testing groups such as Cambridge language assessment. Given these observations, the CSEPT's validity and reliability was accepted on its face. LTTC website: <https://www.ltcc.ntu.edu.tw/language/testing.htm>
2. Beginning in May 2016, all test takers who reached the English proficiency benchmark could be exempted from taking further CSEPT tests. However, this is not applicable to this study.

Acknowledgements

1. The Editorial Board wishes to thank: Dr. Daniël Lakens for his input on Cohen's d-average, and Dayna Jost for her role as a guest production editor for this paper.
2. The authors would like to express their sincere gratitude to the two anonymous reviewers for their constructive comments that helped to improve the quality of the manuscript. They also would like to thank Dr. Cathrine Beaunae for English language editing.
3. This study was one of an integrated research project funded by Wenzao Ursuline University of Languages, Kaohsiung, Republic of China. (Grant #: GERS 104002).

References

- Chen, A. (2011). Parents' perspectives on the effects of the primary EFL education policy in Taiwan. *Current Issues in Language Planning*, 12(2), 205-224.
- Chen, I. W.-L., & Hsieh, J. J.-C. (2011). English language in Taiwan: An examination of its use in society and education in schools. In A. Feng (Ed.), *English language education across greater China* (pp. 70-94). Bristol: Multilingual Matters.
- Chern, C. (2002). English language teaching in Taiwan today. *Asia-Pacific Journal of Education*, 22(2), 97-105.

- Chern, C. (2010). General English programs at universities in Taiwan: Curriculum design and implementations. *Chang Gung Journal of Humanities and Social Sciences*, 3(2), 253-274.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale, NJ: Erlbaum.
- Feng, A. (2012). Spread of English across Greater China. *Journal of Multilingual and Multicultural Development*, 33(4), 363-377. doi:10.1080/01434632.2012.661435.
- Feng, M., & Chang, H. (2010). Effects of ability grouping and development of English graduation benchmarks in a national technological university in Taiwan. *Journal of Applied Foreign Languages*, 13, 71-101.
- Hallam, S., & Ireson, J. (2003). Secondary school teachers' attitudes towards and beliefs about ability grouping. *British Journal of Educational Psychology*, 73(3), 343-356.
- Hua, T.-L., & Beverton, S. (2013). General or vocational English courses for Taiwanese students in vocational high schools? Students' perceptions of their English courses and their relevance to their future career. *Education Research Policy Practice*, 12(2), 101-120.
- Ireson, J., Hallam, S., Hack, S., Clark, H., & Plewis, I. (2002). Ability grouping in English secondary schools: Effects on attainment in English, mathematics and science. *Educational Research and Evaluation*, 8(3), 299-318.
- Ireson, J., & Hallam, S. (2001). *Ability grouping in education*. London: Paul Chapman Publishing House.
- Jung, M. (2000). Ability grouping and equity of educational opportunity. *The Journal of Curriculum Studies*, 18, 275-297.
- Khazaenezhad, B., Barati, H., & Jafarzade, M. (2012). Ability grouping as a way towards more academic success in teaching EFL – A case study of Iranian undergraduates. *English Language Teaching*, 5(7), 81-89.
- Kim, Y. (2012). Implementing ability grouping in EFL contexts: Perceptions of teachers and students. *Language Teaching Research*, 16(3), 289-315.
- Krashen, S. D. (1985). *The input hypothesis: Issues and implications*. New York: Longman.
- Kulik, J. A. (1992). *An analysis of the research on ability grouping: Historical and contemporary perspectives*. Storrs, CT: National Research Center on the Gifted and Talented.
- Lakens, D. (2013). Calculating and reporting effect sizes to facilitate cumulative science: a practical primer for t-tests and ANOVAs. *Frontiers in Psychology*, 4, 1-12.
- Language Training and Testing Center (2007). Research reports on College Students English Proficiency Test—Primary level. Retrieved on July 9th, 2017, from <https://www.lttc.ntu.edu.tw/academics/csept/CSEPT%E7%AC%AC%E4%B8%80%E7%B4%9A%E5%B8%B8%E6%A8%A1%E5%BB%BA%E7%BD%AE%E7%A0%94%E7%A9%B6%E8%A8%88%E7%95%AB%E5%A0%B1%E5%91%8A.pdf>.
- Language Training and Testing Center (n.d.). CESPT and CEFR comparison chart. Retrieved on July 9th, 2017, from https://www.lttc.ntu.edu.tw/CEFRbyLTTC_tests.htm.
- Lee, C., & Lin, C. (2013). A method of English ability-grouping teaching in a university of technology. *International Journal of Information and Education Technology*, 3(2), 268-272.
- Lee, C.-L., & Su, C. C. (2009). The effectiveness of English ability-grouping teaching: A case study of a university of technology in central Taiwan. *Journal of Humanities and Social Science of NHCUE*, 2(1), 233-253.
- Liu, H. J. (2008). An analysis of the effects of ability grouping on student learning in university-wide English classes. *Feng Chia Journal of Humanities and Social Sciences*, 16, 217-249.
- Pan, Y., & Newfields, T. (2012). Tertiary EFL proficiency graduation requirements in Taiwan: A study of washback on learning. *Electronic Journal of Foreign Language Teaching*, 9(1), 108-112.
- Slavin, R. E. (1987). Ability grouping and student achievement in elementary schools: A best-evidence synthesis. *Review of Educational Research*, 57(3), 293-336.
- Slavin, R. E. (1990). Achievement effects of ability grouping in secondary schools: A best-evidence synthesis. *Review of Educational Research*, 60(3), 471-499.
- Trautwein, U., Koller, O., & Kammerer, E. (2002). The effects of within-class and between-class ability grouping on academic self-concepts, reported classroom participation, and perceived social integration. *Psychologie in Erziehung Und Unterricht*, 49(4), 273-286.

- Warden, C. A., & Lin, H. J. (2000). Existence of integrative motivation in an Asian EFL setting. *Foreign Language Annals*, 33(5), 535-545.
- Wen, S.-M. (2011). An analysis of the implementing outcomes of ability grouping of Freshman English in a university of technology. *Journal of National Taichung University of Education*, 25(2), 65-80.

About the Authors

Chiu-hui (Vivian) Wu is an associate professor in the Department of English and the Director of Center for English Language Teaching at the Wenzao Ursuline University of Languages, Kaohsiung, Taiwan. Her research interests include: intercultural education, English language education and qualitative research.

Chia-jung Tsai is an assistant professor in the Center for English Language Teaching at Wenzao Ursuline University of Languages, Kaohsiung, Taiwan. Her areas of research include learning strategies, vocabulary teaching, and materials design.

Yi-Min Chiu is currently an English lecturer at Wenzao Ursuline University of Languages and a doctoral candidate at Department of Foreign Languages and Literature, National Cheng Kung University, Taiwan. Her main research interests lie in peer response in L2/FL writing and English for Academic Purposes (EAP).

Language Input Effects on L2 Composition Peer review Feedback

Clay Williams*

Akita International University

Abstract

To date, very little research on L2 peer review of L2 essay composition has focused on the issue of whether the choice of L1 or L2 use in the peer review process facilitates or hinders students' transmission and reception of productive commentary on their essay drafts. This study was designed to ascertain the impact of language choice in written peer review sessions on both the types of commentary made and essay authors' propensity to incorporate peer commentary into subsequent drafts. Data was collected from essay assignments in an English writing course at an English-medium university in Japan using written peer reviews conducted in both Japanese and English in a cross-balanced design. The data were analyzed to measure the impact of the language used on the number and types of comments made, as well as to ascertain the relative impact of peer commentary on the subsequent revisions. The study found sizeable differences according to the language a peer review was conducted in, suggesting that L2 written peer reviews may be more beneficial at identifying/rectifying paragraph-level and structural issues, whereas L1 peer review was slightly better for correcting rhetorical and logical issues.

Keywords: L2 writing; peer review; essay writing pedagogy; L1/L2 effects

Introduction

Given the increased prominence of English communication in the East Asian educational domain in recent years, it is hardly surprising that more and more universities are offering courses on advanced English essay writing. Increasing competition in the job market has increased the need for demonstrated English skills when applying for employment (Reed, 2002). Recent years have witnessed a sudden growth in the number of tertiary-level programs and whole institutions wherein English is the lingua-franca for all instruction. While a few institutions have long and illustrious histories (e.g., International Christian University in Japan), the last decade has seen the birth of such institutions as Akita International University (Japan), University of Nottingham Ningbo (China), Tan Tao University (Vietnam), and Xing Wei College (China), all offering full degree programs taught entirely in English. In addition to entirely English-medium institutions, other established regional universities have begun offering specific degree programs taught largely or entirely in English, such as Waseda University (Japan), Ho Chi Minh International University (Vietnam), and Asian Pacific International University (Thailand). Outside of the Asia-Pacific region, likewise, there is a global trend of increased number and prominence of English-medium programs (Dearden, 2014). According to an ICEF Monitor report (Trend alert, 2012), English is already the lingua franca of many Middle Eastern universities, there are over 4500 courses being taught in English across continental Europe, and the number of English-medium courses/programs is on the rise in Africa and South America, as well.

* Email: williams@aiu.ac.jp. Tel.: +81-18-886-5820. Address: English Language Teaching Practices, Graduate School of Global Communication and Language, Akita International University, Yuwa, Akita-city 010-1292, JAPAN

This steep rise in English-medium content course availability has had a deep impact on both the type and level of English language preparatory courses being offered, and the area of academic writing has been no exception to this phenomenon. Whereas, not very many years ago, most L2 English writing programs functioned at a low level, mostly focusing on issues of grammatical accuracy, these new English-medium programs require writing courses which prepare students to write for English-medium content courses across a broad array of academic disciplines. This has effectively raised the standards for writing course content and curricular goals in East Asia, as has the increased English proficiency of the students enrolled in such programs. These curricular changes have resulted in many of these preparatory courses consciously modelling themselves after the sorts of writing programs that L2 English international students would encounter in universities in L1 English countries. This gives rise to some unique issues, as most prior research in L2 English writing examining high-proficiency writing has come from the ESL context, and not an EFL context. The sudden proliferation of advanced-level writing programs in EFL contexts—wherein students come from the same L1 and national background—enables an opportunity for experimentation in pedagogical design and efficacy that would not be possible in lower proficiency level nor with students with mixed L1s and nationalities.

This paper investigates peer review methodology in light of the new classroom dynamics created by the proliferation of high-level (i.e., focusing on production of academic and/or professional essay writing skills as opposed to a more general focus on sentence and paragraph-level vocabulary and grammar skills) L2-English writing courses in East Asia. As many academic writing courses make use of written peer review as part of the drafting/revision process for essay writing, the question arises as to whether there would be any immediate advantages, disadvantages, or qualitative differences derived from conducting peer reviews in the students' L1 or L2.

Literature Review

Peer Review

One of the fundamental, main-stay pedagogical approaches to most writing composition programs and courses is the concept of peer review. Originally developed for use in L1 composition teaching, and fitting squarely into the cooperative learning strategies domain, this has been a time-tested teaching technique, and its positive effects on L1 composition student performance have been reported in studies too voluminous to fully catalogue here (e.g., Beaven, 1977; Gere & Abbot, 1985). In the L2 composition domain, as well, it has a distinguished track-record, and many researchers have attested to its effectiveness in application to L2 writing courses (e.g., Allison & Ng, 1992; Arndt, 1993; Keh, 1990; Lockhart & Ng, 1993; Tsui & Ng, 2000). The perceived benefits of peer review are many: it is thought to help students in developing the ability to appropriately analyze and revise their own writing (Zhang, 1995); it may increase learner participation (Mendonça & Johnson, 1994); and it is supposed that peer feedback is less threatening than teacher feedback (Jacobs, Curtis, Braine, & Huang, 1998). However, these views have not gone without challenge. Nelson and Murphy (1993) observed significantly more instances of students being overly focused on surface problems at the expense of deeper textual issues, as well as greater reluctance to implement peer commentary on the part of L2 writers as opposed to L1 writers. Their status as second language speakers is assumed to make students more reticent towards accepting classmate feedback. Additionally, the tendency towards teacher-centered classroom practices and general deference extended to teachers in East Asia can make students wary of accepting peer commentary as authoritative, or even as helpful (Ferris, 2003; Nelson & Murphy, 1993). Wu (2006) found in a study on Chinese L1 students of English that teacher feedback had a measurably greater impact on student writing performance than did peer feedback.

There has been much investigation delving into various aspects of peer review in the L2-English context which has enlightened our understanding of the strengths and potential limitations of peer review activities in L2-English composition courses. For instance, Liao and Lo (2012) found that the relative quality and types of peer review commentary was largely dependent upon students' L2 proficiency levels. While both high and low-level proficiency learners' comments were dominantly used to identify problems, higher-level proficiency reviewers' comments provided significantly more detail in both discussion of the problem and in suggesting means of

improvement. Salih (2013) conducted post-peer-review debriefing sessions and interviews to compare student expectations for peer review comments with the actual patterns emerging from peer review activities. They found that, despite the writers' expectations that peer reviewers would focus on grammatical correction, in reality, the dominant type of comments delivered were regarding essay structure, and reviewers recounted their primary focus was in maintaining the clarity of the feedback. Yu and Lee (2015) investigated the factors which determine individual participation in group peer feedback activities, and they found that the primary determining factor was student motivation, which itself was affected directly by sociocultural context. Some of the studies on peer review in L2 contexts have turned up less definitive results, but are nevertheless valuable in shaping our understanding of the pedagogical value of the activity. For instance, LoCastro (2000), investigating whether peer review would follow the dominant discursive norms of the L1 and L2, found the results inconclusive; however, the study also revealed that there seemed to be a clear effect of instruction in dropping L1 discursive behaviors.

One of the more extensively-covered aspects of peer review in the L2-English classroom domain has been the issue of the extent to which recent technological advancements could significantly improve students' performance and acceptance of classmates' suggestions. Much of the research has found computer-mediated learning to be an effective means of L2 study, as it can lower affective variables and enhance motivation (e.g., Coniam & Wong, 2004; Strenski, Feagin, & Singer, 2005), there has naturally been some curiosity as to whether this would impact the efficacy of peer review in L2 writing classes. Some researchers (e.g., Crank, 2002; Liu & Sadler, 2003) have found significant improvements to peer review performance when peer review is conducted via computer platforms in asynchronous computer-mediated communication; whether or not those effects are permanent has come under question. Xu (2007) found that performance boosts fade over time, and suggests that the temporary increase in productivity simply reflects student curiosity and excitement over the new technology, and diminishes as students become accustomed to its use.

The Issue of Language Medium in Peer Review

The motivating question for this study was whether the language in which written peer essay review was conducted would prompt students to write better commentary and/or be more accepting of peer suggestions (as evidenced by inclusion in subsequent drafts). As basic as the question appears on its face, it appears that there has been almost no research to date touching upon the issue. While it would be impossible to state conclusively why something does not receive broader treatment in the academic community, the author's suspicion is that the issue has been ignored because, at least in East Asia, the issue was mostly moot until relatively recently. As teaching practices tend to be pragmatic at their core, if a group of students was not very skilled in the L2, it would be unsurprising for an instructor to allow students to use their L1 for peer review purposes, so as to enable them to fully explain their thoughts on each other's writing. Likewise, if students had the requisite skills to conduct a meaningful peer review in the L2, many teachers might be hesitant to allow students to conduct the review in their L1, and thus to lose such a precious opportunity to use and expand upon L2 writing skills. In the ESL context, likewise, there often exists the need to conduct peer review in L2 English regardless of student proficiency because classes are frequently composed of mixed groups of students from various language/cultural backgrounds, with English being the only common vehicle of communication. It is only in the new, emerging context of the English-medium international university, wherein monolithic blocks of students from the same L1-background study through the medium of L2-English, that the question becomes valuable. These students unquestionably have the skills to conduct L2 peer review; however, as they use English day-in and day-out during their university experience, there would likely be less compulsion to make every minute of class time count. Classes are also overwhelmingly populated by students of similar cultural and linguistic backgrounds, so use of L1 for peer review becomes a realistic option. As both options seem equally practical and valid, it becomes necessary to consider whether or not there would be any qualitative differences in feedback and revision performance between written peer review conducted in students' L1 or L2.

While prevailing practices have emphasized use of L2 in written peer review sessions, there are reasons to suspect that L1 peer commentary may in some ways be more effective towards producing a better final essay. Studies such as that by Nelson and Murphy (1993), as well as Wu (2006), indicate some level of resistance by L2

writers to take the critiques of other L2 writers very seriously. While writers can understandably be concerned about the limitations of subject expertise and knowledge by peer reviewers, these negative perceptions are sometimes further exacerbated by reviewers' lack of L2 skill constraining their ability to articulate legitimate points about the writings they are tasked with reviewing. Simply put, if peer reviewers respond in the shared L1, they may be able to express their concerns about the essay's structure, coherence, logic, etc. with considerably greater coherence and detail than is allowed by the L2. This greater specificity by reviewers could result in a stronger argument for change, thereby more likely being accepted into subsequent revisions by the original author.

There appears to have been little study of language effects on peer review. One exception was Huang's (1996) study of oral peer review in L2 English classes. Huang found an asymmetry in the foci of students' commentary on each other's papers, depending on the language used to moderate the discussions. Peer feedback delivered in the students' L1 (Mandarin Chinese) were more specific, and focused mostly on issues of language usage. By contrast, while groups delivering feedback in the L2 (English) were more general in their commentary, they nevertheless managed to discuss a broader range of issues (e.g., language use, essay reasoning, and rhetorical strategies). Both languages displayed positive effects on student feedback sessions: the use of L1 was perceived as being more effective in eliciting deeper commentary and appeals to implement peer feedback; however, the L2 sessions, while not as focused, nevertheless elicited more communal support among students. Since this study focused only on *oral feedback*, it is still unknown whether language choice in *written feedback* elicited via peer review would lead to differences either in the types of comments made or in its persuasiveness (as measured by prompting the authors to change their papers in subsequent drafts).

Research Questions

The study described herein was designed to ascertain whether advanced students of English composition would show any significant differences in peer review performance depending on whether the peer reviews were conducted in their L1 or L2. The specific questions which this study sought to answer are:

- 1) What is the relationship between the language that peer feedback is delivered in and the type of feedback delivered?
- 2) In individual feedback categories, what is the relationship between the language that peer feedback is delivered in and the degree to which authors are willing to accept comments and integrate them into subsequent revisions?
- 3) In holistic categories (based upon Huang, 1996) of mechanical vs. rhetorical commentary, is there a relationship between the language used for peer review and the types of comments that authors are more willing to accept and integrate into subsequent revisions?

Methods

Participants

The study was conducted at a small, English-medium international university in northern Japan. At this university, all students take 1-3 semesters of foundational intensive English coursework before being mainstreamed into regular content (degree-seeking) coursework. Upon mainstreaming, the students are still required to take two English composition-writing courses as graduation requirements. These courses are usually taken within the first two semesters after finishing the foundations program and beginning their degree programs. The participants in this study were 39 students in two different sections (20 in one class section, and 19 in the other) of the first required composition-writing course. All students had Japanese as their L1, and all had been enrolled at the university for 1-3 semesters prior to taking the course (most were in their second semester). At the time, TOEFL IPT scores of at least 500 were required in order to exit the foundation classes and begin regular coursework; however, most students had already exceeded that minimum score before embarking on the

foundations coursework, and exit criteria have since been modified. While individual TOEFL scores are protected by law, and thus inaccessible to the researcher, at the time of the study, the average TOEFL IPT score in such writing classes would have been around 530, with a range from a low of 500 (as the minimum score for entry) to a high of 650. The students had all taken at least one academic writing course focused on production of multi-paragraph essays in the preceding semester, and as part of that course, had already gained some experience with peer review procedures.

Data Collection

The course was designed as an introduction to analytical, academic essay writing. For each assignment, students would read a handful of essays and/or short stories (grouped by a theme) and develop a unique interpretation through literary analysis. They would then write a 4-5-page essay arguing/defending their interpretation. Each essay-drafting period would involve two opportunities for peer review. The peer review sessions were conducted in-class, and were facilitated by a 16-question response sheet (see Appendix 1) asking a mix of short- and long-answer questions focusing on paragraph-level and essay-level critique and response. All peer review during the first essay was conducted in English, and was preceded by explanations, examples, and group work designed to get students familiar with both the form and the peer review dynamic, as well as to firmly define the expectations and boundaries of acceptable and unacceptable peer commentary.¹

The experiment started from the second required essay, and lasted through the third and final essay (for a total of two essays and four peer review sessions). The essay response sheet was translated by a bilingual native speaker of Japanese, and from the first peer review session for paper #2, class “A” was given the Japanese peer review sheet (see Appendix 2) and was encouraged to respond to each other in Japanese, whereas class “B” continued to peer review in English. This was held constant through both peer review sessions for paper #2. For the third paper, the languages used for the peer review in each class were switched, with class “A” conducting both peer reviews in English, and class “B” conducting both peer reviews in Japanese, thus counterbalancing the presentation of English and Japanese peer review. The course instructor explained to both classes that this peer review commentary was being used in a study to determine what (if any) impact the language of the peer review would have on their performance, and encouraged students to respond in the appropriate language; students were not given any additional information on how the data collected would be analyzed. All peer review sheets, corresponding draft essays, and final essays were handed in at the end of each assignment period. Photocopies were made to enable final grades and instructor feedback to be distributed to the students while the researcher retained the original forms. Students were told that, upon request, they could reclaim their original copies, though none followed up on this offer.

Analytical Procedure

After the end of the semester, all documents were analyzed by a team of two bilingual L1-Japanese/L2-English graduate students according to a prepared rubric. First, they focused on identifying all peer review comments according to the types listed in Table 1. The first 5 categories focus on word, sentence, and paragraph-level issues of structure. They include mechanical issues (e.g., highlighting a misspelled word or incorrect grammar), introduction/thesis (e.g., one of the most common comments regarded theses which were either overly broad or failed to make an argument), topic statements (e.g., pointing out incorrectly placed, missing, or topic statements irrelevant to the following paragraph content), body paragraphs (e.g., unclear structure or examples that do not clearly support the topic), and conclusions (e.g., failure to restate one’s position). Categories 6-9 focus on more global issues of coherence and appropriateness. These include phrasing issues (e.g., issues of word choice and degree of formality/informality in writing), logic issues (e.g., non-sequitur arguments), persuasiveness issues (i.e., an inability to convince the audience of one’s point), and global comments (i.e., the reviewer’s overall view of the paper). As such, while the feedback types are subcategorized more extensively, the first 5 categories can be considered to correspond directly to Huang’s (1996) category of language/accuracy focus, and categories 6-9 correspond to Huang’s second (broader) category of rhetorical/logical focus. Category 10 measures the extent to which the peer review exercise was used for purely social reasons (e.g., “Hi! How’s it going?”), and as such,

stands largely outside of either category. The two student evaluators worked together and were required to agree on classifications. In the event of disagreements, the author would make the tie-breaking vote. Next, all instances of each category were cataloged and numbered according to language used (i.e., Japanese or English).

An important caveat needs to be mentioned before continuing to the results. Seven students were dropped from the study and results from their work were not factored into the analysis. Six of the seven were rejected because they had missed one or more peer review sessions and/or failed to hand in one of the required papers. Results are only tabulated using students who attended all peer review sessions and turned in all required work. The other student responded almost entirely in the wrong language during one of the peer review sessions, and was thus omitted from the analysis.

Table 1
Individual Feedback Categories in Peer Review

Word/ Sentence/ Paragraph-level Analysis	Global Issues (Coherence/Appropriateness)	Social Commentary
1) Spelling & Grammar	6) Phrasing Issues	10) Personal Comments
2) Introduction/thesis	7) Logic Issues	
3) Topic statements (in-body paragraphs)	8) Persuasiveness Issues	
4) Body Paragraph	9) Global Comments	
5) Concluding paragraph structure		

Once the peer review comments had been analyzed and categorized, the data was used to address the three motivating research questions. The first question—regarding the relationship between the language that peer feedback is delivered in and the type of feedback delivered—could be answered in two ways: either (1) by counting the number of subjects receiving explicit feedback according to each of the two categories used by Huang (1996) and then comparing (via a 2x2 Chi-square test) the number in each feedback category according to whether the feedback was delivered in Japanese or in English; or, (2) by counting the total number of instances of peer feedback in each of the two categories, and comparing by language (again, via a 2x2 Chi-square test). Both methods were employed here. When analyzing by ‘subject,’ it should be noted that it would be possible for one subject to populate more than one cell and thereby violate an assumption of the test, but subsequently, the unit of analysis was the instance of feedback, so the Chi-square test was still employed noting this limitation.

Research question number two—measuring any association between the language of peer review and the likelihood of individual categories of feedback being accepted into subsequent revisions—required an analysis of the draft vs the final form of the assignment (handed in for grading) to determine if the peer review comments had been accepted and used. With this information, the data could be analyzed via 2x2 Chi-square tests for each category individually.

In responding to the third question—concerning the relationship between the language of peer feedback and types of peer commentary which authors are more likely to accept and integrate into subsequent revisions—the data collected recording the number of instances of peer commentary and the number of comments which were implemented into subsequent paper drafts was simplified into the two broad categories of mechanical comments (sentence or paragraph-level) and logical/rhetorical comments (essay-level) according to Huang (1996). Once the data was compiled, a comparison of the degree of uptake in each category could be made in each language individually, and the categories of feedback could also be directly compared (individually) across languages.

Results

Research Question 1: Language Choice and Type of Feedback

The total number of subjects receiving each manner of feedback is shown in Table 2. Frequencies of mechanical and rhetorical feedback rendered in Japanese and English were compared via a Chi-squared test ($\alpha = 0.05$). There was no significant difference in the numbers of subjects receiving each type of feedback: $X^2(1, N = 156) = 0.06, p < 0.80$.

Table 2

Numbers of Students Receiving Feedback by Type and Language

Type of feedback	English: Count # of Students receiving commentary by type	Japanese: Count # of Students receiving commentary by type
1) Spelling & Grammar	11	17
2) Intro / thesis	12	9
3) Topic statements	11	10
4) Body paragraph structure	11	10
5) Concluding paragraph structure	10	9
TOTAL: Word/Sentence/Paragraph-level Analysis (categories 1-5)	55	55
6) Phrasing issues	9	8
7) Logic issues	6	6
8) Persuasiveness issues	4	5
9) Global comments	5	3
Total: Global Issues (Coherence/ Appropriateness) (categories 6-9)	24	22
10) Social Commentary		

The total number of incidents of each manner of feedback in each language is shown in Table 3. An analysis of the association between language used in the peer review and the feedback type according to Huang's (1996) two categories of word/sentence/paragraph-level analyses and global issues yielded no significant differences $X^2(1, N = 234) = 0.39, p = 0.53$, so while comments were somewhat more voluminous in Japanese than in English, the relative proportion of mechanical to rhetorical comments did not vary according to the language used in the peer review.

It is also worth noting that Japanese (L1) critiques of all types tended to be longer and more detailed. As a small example of the difference between the two, we can look at two examples, by the same reviewer, taken from the "logic issues" category. In a review written in English, the student writes simply, "Body paragraphs should be more organized," with no more detail. Reviewing a different paper in Japanese, the reviewer makes a relatable observation, but this time expands into significantly more detail, "[Paragraph order should be reorganized according to your 2 categories: communication and information. Now, the body paragraphs are not connected well]." This kind of specificity and detailed critique allows students to be much more confident in the accuracy and validity of peer commentary, thus possibly explaining why Japanese holistic comments were so much more widely implemented into the essay drafts.

Table 3
Numbers of Instances of Feedback by Type and Language

Type of feedback	Number of instances of English feedback	Number of instances of Japanese feedback
1) Spelling & Grammar	13	23
2) Intro / thesis	16	18
3) Topic statements	17	21
4) Body paragraph structure	14	19
5) Concluding paragraph structure	12	17
TOTAL: Word/Sentence/Paragraph-level Analysis (categories 1-5)	72	98
6) Phrasing issues	12	15
7) Logic issues	8	9
8) Persuasiveness issues	4	6
9) Global comments	6	4
Total: Global Issues (Coherence/Appropriateness) (categories 6-9)	30	34
10) Social Commentary	0	4

The topic of personal comments also bears further discussion. While the use of peer review activities for personal interaction falls outside the scope of this study, and the category was simply used to classify commentary that was social in nature and therefore did not respond directly to the paper being reviewed, I still found it rather curious that there were no such comments made in L2 English. While it was somewhat predictable that L1 would lend itself more to socializing (assuming that students would normally speak to one another in Japanese outside of class), it must be noted that these students are all extremely proficient in English by Japanese academic standards, and make social use of English on a daily basis at the school (where roughly 20% of the student body is composed of international students). The absence of any sort of personal comments made during English-language peer reviews suggests that the students were subconsciously treating English as a transactional medium, and interaction was reserved for Japanese. While the four examples garnered are hardly enough to base any firm conclusions on, the use of L1 and L2 for interactive and social purposes in written peer review activities would be an issue well worth devoting attention to.

Research Question 2: Acceptance of Comments for Integration in Later Revisions vs. Language of Feedback in Individual Feedback Categories

In analyzing comments to determine whether or not the essay authors incorporated them into their subsequent essay drafts, more differences start to emerge. We can see the raw numbers for feedback integration according to whether feedback was delivered in L1 or L2 in Table 4. Direct analysis of the degree of association between L1 or L2 use in peer review and the subsequent degree of integration of peer review comments into the final paper for each of the individual subtypes of feedback (via isolating each horizontal line on Table 4 and using a 2x2 Chi-square test) showed several interesting trends. Comments regarding topic statements were significantly more likely to be implemented when feedback was delivered in English: $X^2(1, N = 38) = 3.75, p = 0.05 (\Phi=0.31,$

medium effect size); as were comments about essay conclusions: $X^2(1, N = 29) = 5.15, p = 0.02$ ($\Phi=0.42$, medium effect size). Comments regarding body paragraph structure delivered in English approached significance: $X^2(1, N = 33) = 2.95, p < 0.09$ ($\Phi=0.29$, small effect size). Peer feedback delivered in Japanese, however, yielded rates of incorporation in subsequent drafts approaching significant difference when comments were about phrasing issues: $X^2(1, N = 27) = 2.79, p < 0.10$ ($\Phi=0.32$, medium effect size); and logic issues: $X^2(1, N = 17) = 2.84, p = 0.09$ ($\Phi=0.40$, medium effect size). The analysis of personal comments was suspended as 3 of the 4 comments were social in nature, and did not require any sort of decision which would yield a measureable change in the paper. Other categories failed to produce any significant difference between English and Japanese feedback (Spelling & Grammar: $X^2(1, N = 36) = 1.00, p < 0.32$; Intro/thesis: $X^2(1, N = 34) = 0.17, p < 0.68$; Persuasiveness: $X^2(1, N = 10) = 0.27, p < 0.60$; Global: $X^2(1, N = 10) = 1.67, p < 0.20$).

Table 4

Acceptance/Integration of Feedback into Subsequent Draft by Type and Language

Type of feedback	English feedback: Integrated	English feedback: NOT Integrated	Japanese feedback: Integrated	Japanese feedback: NOT Integrated
1) Spelling & Grammar	11	2	16	7
2) Intro / Thesis	10	6	10	8
3) Topic statements	14*	3	11	10
4) Body paragraph structure	12**	2	11	8
5) Concluding paragraph structure	10*	2	7	10
6) Phrasing issues	7	5	13**	2
7) Logic issues	3	5	7**	2
8) Persuasiveness issues	2	2	4	2
9) Global comments	2	4	3	1
10) Social Commentary	0	0	1 (**+3)	0

Notes.

*statistically significant difference

**approaching significance

***3 of the 4 personal comments did not involve any decisions on the paper itself

Research Question 3: Relationship Between Language of Peer Review and Types of Commentary Integrated into Subsequent Revisions

By simplifying the feedback categories—word/sentence/paragraph-level analyses and global issues, more patterns emerge. Table 5 provides the total number of integrated and non-integrated comments in each category according to the language of peer review. Direct comparison between language of peer review feedback and degree of integration into the final paper across the board (i.e., adding together comments from both categories vs uptake analyzed according to language of peer review) yielded no significant results: $X^2(1, N = 234) = 1.41, p = 0.49$. However, an analysis of the degree of uptake according to type of comments (i.e., mechanical vs rhetorical) in English peer reviews revealed that comments and recommendations about mechanical issues were significantly more likely to be integrated into the subsequent version of the paper than were comments regarding

rhetoric: $X^2(1, N = 102) = 10.63, p = 0.005$ ($\Phi=0.32$, medium effect size). Japanese peer review yielded a significantly higher proportion of uptake of comments concerning rhetorical issues: $X^2(1, N = 132) = 5.82, p = 0.05$ ($\Phi=0.21$, small effect size). Direct comparison of the degree of uptake of comments concerning mechanical issues according to the language of peer review revealed a significantly higher degree of uptake when the comments were delivered in English: $X^2(1, N = 170) = 9.8, p = 0.007$ ($\Phi=0.24$, small effect size), whereas comparison of uptake of comments about rhetorical issues showed a significant advantage for comments made in Japanese: $X^2(1, N = 64) = 7.43, p = 0.02$ ($\Phi=0.34$, medium effect size).

Table 5

Acceptance/Integration of Feedback into Subsequent Draft by Simplified Type and Language

Type of feedback	English feedback: Integrated	English feedback: NOT Integrated	Japanese feedback: Integrated	Japanese feedback: NOT Integrated
Word/Sentence/Paragraph-level Analysis (categories 1-5)	57	15	55	43
Global Issues (Coherence/ Appropriateness (categories 6-9))	14	16	27	7

Discussion

There are some useful discoveries from the experiment which may help elucidate how students conduct and respond to written peer feedback according to whether they use L1 or L2. Comparing the instances of feedback, one does find some support for Huang's (1996) study of oral feedback, as feedback delivered in the students' L1 did produce more comments regarding word, sentence, and paragraph-level issues, which could be classified in Huang's terminology as "language issues." This type of focus can also be used to explain the near-significant (i.e., $p < 0.10$) difference in the number of comments regarding the structure of the conclusion paragraph.

However, it is when we start comparing the relative willingness of students to incorporate peer feedback that we start to see a definite trend. Peer feedback delivered in L2 regarding word, sentence, and paragraph-level issues seemed more likely to be accepted and implemented into the next draft, as evidenced by the significantly higher rate of comment incorporation for comments in L2-English regarding topic statements, body paragraph structure, and conclusion paragraphs. By contrast, comments regarding phrasing issues or logic issues were significantly more likely to be accepted if delivered in L1-Japanese—even though the number of comments made in English and Japanese were quite similar. Thus, the evidence suggests that in written peer review activities, much like Huang (1996) found in oral feedback sessions, L1 feedback will focus more on issues of language (i.e., at the word/sentence/paragraph level); however, commentary of this type is much more likely to be implemented when delivered in the L2. While no real differences in frequency of comments regarding issues of language use, rhetoric, and reasoning were found between L1 and L2 use, students were more likely to incorporate such suggestions when delivered in their L1.

Conclusion

These findings can be of some use for curricular planning because they suggest that students are unconsciously focusing on opposite areas of essay analysis depending upon the language of the peer review. If these results are representative, it would follow that teachers could strategically use L2 peer review to good effect for focus on language form issues, and that it would be more effective to allow peer reviews focusing primarily on issues of rhetorical effectiveness to be conducted in an L1.

It must be noted that the study described herein is limited by nature, being a relatively small test group from a single nation/language background, therefore it would be advisable to ascertain the applicability of such

results to the larger East Asian or global context via additional testing in other classroom contexts. These results cannot yet be considered as indicative of broader L1/L2 issues, but rather as local (i.e., Japanese/English, university-level, etc.) effects. While the results *might* fall within a larger trend, it is important not to generalize the results of 32 students in northern Japan without plenty of verification from a wide array of locales and L1/L2 combinations. Additionally, the small sample size introduces a power concern (seen herein by all significant effects only having small or medium effect size) which could only be eliminated through more expansive testing of this type. Furthermore, the institution where the study was administered, being a rather small English-medium university with a high national ranking, there is admittedly a high chance for selection bias to have influenced the results. In order to be able to generalize these results beyond their immediate context, it would be useful to try to replicate the study in other areas of the Asia Pacific region. Whether the results captured herein are specific to Japanese learners (or even a subsection thereof) or represent a general trend in L2 learners of essay writing, it would be invaluable to the field to amass more varied information on the interrelation of language and written peer review commentary. While national, regional, and/or linguistic differences may well emerge from such expanded study, this knowledge could then enable writing teachers to target peer review activities to best suit the individualized learning aims of classes.

Notes

¹Problems here are rare, but it is very important to make sure that peer reviews are not used as a forum for airing grievances or executing vendettas).

²Translation from Japanese by one of the graduate students involved in the initial sorting.

References

- Allison, D., & Ng, P. (1992). Developing text revision abilities. In M. Lau and M. J. Murphy (Eds.), *Developing Writing: Purposes and Practice* (pp. 106-130). Hong Kong: Institute of Language in Education.
- Arndt, V. (1993). Response to writing: Using feedback to inform the writing process. In M.N. Brock and L. Walters (Eds.), *Teaching Composition Around the Pacific Rim: Politics and Pedagogy* (pp. 90-116). UK: Multilingual Matters.
- Beaven, M.H. (1977). Individualized goal setting, self-evaluation, and peer evaluation. In C.R. Cooper and L. Odell (Eds.), *Evaluating Writing: Describing, Measuring, Judging* (pp. 135-156). Urbana, IL: National Council of Teachers of English.
- Coniam, D., & Wong, R. (2004). Internet Relay Chat as a tool in the autonomous development of ESL learners' English language ability: An exploratory study. *System*, 32(3), 321-335.
- Crank, V. (2002). Asynchronous electronic peer response in a hybrid basic writing classroom. *Teaching English in the Two Year College*, 30(2), 145-155.
- Dearden, J. (2014). *English as a medium of instruction – a growing global phenomenon: Phase 1*. London: British Council.
- Ferris, D.R. (2003). *Response to student writing: Implications for second language students*. Mahwah, NJ: Routledge.
- Gere, A.R., & Abbot, R. (1985). Talking about writing: The language of writing groups. *Research in the Teaching of Writing*, 19, 362-385.
- Huang, S.Y. (1996). L1 or L2 peer response sessions? Differences in verbal interaction between a writing group that communicates in Mandarin Chinese and one that uses English. *Tunghai Journal*, 38(1), 131-151.
- Jacobs, G.M., Curtis, A., Braine, G., & Huang, S.Y. (1998). Feedback on student writing: Taking the middle path. *Journal of Second Language Writing*, 7, 307-317.
- Keh, C.L. (1990). Feedback in the writing process: a model and methods for implementation. *ELT Journal* 44(4), 294-304.
- Liao, H.C., & Lo, Y.F. (2012). Peer review comments provided by high- and low-proficiency L2 learners: A comparative study. *International Journal of English Linguistics*, 2(5), 45-54.

- Liu, J. & Sadler, R.W. (2003). The effect and affect of peer review in electronic versus traditional modes on L2 writing. *Journal of English for Academic Purposes*, 2, 193-227.
- LoCastro, V. (2000). Evidence of accommodation to L2 pragmatic norms in peer review tasks of Japanese learners of English. *JALT Journal*, 22(2), 245-270.
- Lockhart, C., & Ng, P. (1993). How useful is peer response? *Perspectives*, 5(1), 17-29.
- Mendonca, C.O., & Johnson, K.E. (1994). Peer review negotiations: revision activities in ESL writing instruction. *TESOL Quarterly*, 28(4), 745-769.
- Nelson, G.L., & Murphy, J.M. (1993). Peer response groups: Do L2 writers use peer comments in revising their drafts? *TESOL Quarterly*, 27, 135-142.
- Reed, J. (2002). The pedagogical challenges for Western ESL teachers in Asia. *TESL Ontario: Contact*, 28(4), 1-8.
- Salih, A.R.A. (2013). Peer response to L2 student writing: Patterns and expectations. *English Language Teaching*, 6(3), 42-50.
- Strenski, E., Feagin, C.O., & Singer, J.A. (2005). Email small group peer review revisited. *Computers and Composition*, 22(2), 191-208.
- Trend alert: English spreads as teaching language in universities worldwide (2012, July 11). *ICEF Monitor*, Retrieved from <http://monitor.icef.com/2012/07/trend-alert-english-spreads-as-teaching-language-in-universities-worldwide/>
- Tsui, A.B.M., & Ng, M. (2000). Do secondary L2 writers benefit from peer comments? *Journal of Second Language Writing*, 9(2), 147-170.
- Wu, W.S. (2006). The effect of blog peer review and teacher feedback on the revisions of EFL writers. *Journal of Education and Foreign Languages and Literature*, 3, 125-138.
- Xu, Y. (2007). Re-examining the effects and affects of electronic peer reviews in a first-year composition class. *The Reading Matrix*, 7(2), 1-21.
- Yu, S.L., & Lee, I. (2015). Understanding EFL students' participation in group peer feedback of L2 writing: A case study from an activity theory perspective. *Language Teaching Research*, 19(5), 572-593.
- Zhang, S. (1995). Re-examining the affective advantage of peer feedback in the ESL writing class. *Journal of Second Language Writing*, 4(3), 209-222.

Appendix 1 **English Peer Review Form**

Peer Review Worksheet:

Introduction Paragraph

1. Write what you believe the thesis is.
2. Where is it located? Is it in the expected place (the very last sentence of the Introduction)?
3. How does the writer lead up to the thesis? Is the information useless, helpful, boring, interesting, alluring, off-putting? Is the information directly connected to the thesis? Explain your answer.

Body Paragraphs

4. List the individual topics for each paragraph.
5. Are there any paragraphs that do not deal with a specific topic?
6. How well do the topic sentences for each body paragraph represent what those paragraphs contain?
7. How well do the body paragraphs serve as evidence for the thesis? Do they directly connect to the thesis? Are any points unexplained?
8. Which paragraphs seem useless, aimless, or need reorganizing? Explain any problems.
9. How well does each paragraph transition to the next? Do the paragraphs seem disjointed or carefully arranged?

Conclusion Paragraphs

10. How long is the conclusion?
11. Is the thesis *restated* in some way in the Conclusion?
12. How alike/different is it from the thesis in the Introduction?
13. Where is the thesis located?

Evaluation Summary

14. What are three excellent aspects of this paper?
15. If you had to make three recommendations for change, what would they be? Name them in order of importance.
16. How much effort do you think went into this draft?

Appendix 2

Japanese Peer Review Form

相互評価シート

序文

1. この論文の主題を述べなさい
2. 主題はどこに書かれていますか。主題は序文の最後の部分にありましたか。
3. 筆者はどのように主題へと導いていますか。序文の導入部に書かれている情報は役立ちましたか。または（不要、退屈、興味深い、魅惑的、的外れ）ものでしたか。それらの情報は主題に直接的に関与していましたか。自分の意見を述べなさい。

本文

4. 各段落の論題を述べなさい
5. それぞれの論題に対し直接関与していない段落はありましたか
6. 各段落においてそれぞれの論題は明確に示されていましたか
7. 本文はこの論文の主題を証明するものとしての役割を果たしていますか。主題に直接繋がるものですか。説明が不十分な部分はありましたか
8. 不必要、目的のない段落がある場合はそれらの問題を指摘しなさい
9. 各段落はそれぞれ次の段落に円滑に繋がっていましたか。各段落は分裂しているものでしたか。それとも注意深く繋がっていたものでしたか

結論

10. 結論はどのくらいの長さでしたか
11. 主題と結論に関連性はありましたか
12. 序章の主題と結論を比較し、似ている部分、異なっている部分を述べなさい
13. 主題はどこに書かれていますか

評価

14. この論文において優れた見地を3つ述べなさい
15. 校正すべき部分を3つ、最重要箇所から順に述べなさい
16. この論文の下書きはどの程度練られていたか自分の意見を述べなさい

About the Author

Clay Williams is an associate professor in the English Language Teaching graduate program at Akita International University. He is the author of *Teaching English in East Asia: A Teacher's Guide to Chinese, Japanese, and Korean Learners*, as well as other books and articles about English pedagogy in the East Asian region.

Learning to Spell in English by Chinese Students: A Cross-sectional Study

Donald R. Bear*

Iowa State University, USA

Sam von Gillern

Texas A&M University, USA

Wei Xu

Davidson Institute: University of Nevada, Reno, USA

Abstract

This study investigates the English spelling of students in grades 2 through 8 in Mainland China. A review of spelling and cross-linguistic research in spelling is presented. The orthographic development of 273 students was assessed *with validated spelling inventories* (Sterbinsky, 2007) to sample developmental features across three layers of English orthography. The sample size and the detailed analyses make this study unique. The 13 features in the analyses spanned development from beginning consonants to roots. Feature analyses predicted grade level, and there were significant differences across grades. Students made predictable spelling errors that reflected a similar developmental sequence to native English speakers. The error analyses across grades adds to the body of cross-linguistic literature illustrating how English language learners develop basic literacy skills in a similar manner to English-L1 children (Chiappe, Siegel, & Gottardo, 2002; Ford, Invernizzi, & Huang, 2014). This study indicates that as some English learners advanced in grades, their L1 has less of an impact on their spelling in English. When orthographic knowledge is examined across the alphabet, pattern, and meaning layers and cross-linguistically, researchers and educators can identify features students have mastered and what they are learning as a guide for a sequence of instruction and monitoring growth.

Keywords: grades 2-8, spelling development, orthographic knowledge in second language, Chinese-speaking English learners, EFL, cross-language transfers

Introduction

Students' spelling error patterns illuminate the development of their orthographic knowledge (Berninger, Abbott, Nagy, & Carlisle, 2009; Ehri, 2000; Ehri & Roberts, 2006; Templeton & Morris, 2000). This study contributes to the study of English spelling errors and orthographic knowledge in a cross-linguistic study of errors made by Chinese-speaking English L2 learners. Researchers have found that English learners gain their orthographic knowledge in a similar manner to their English L1 peers across the alphabet, pattern, and meaning layers of text while also using their knowledge of writing acquired from their experiences learning another writing system. Like all students, rates of acquisition among English learners vary with experiences (Helman, 2004; Invernizzi & Hayes, 2004; Yeong, Fletcher, & Bayliss, 2014). However, most of the research in developmental spelling and orthographic knowledge has been conducted with students whose primary language (L1) is English. In this study of second through eighth graders, the orthographic development of students who are building on Chinese are examined for developmental trends over these seven grades.

* *Email: drbear@iastate.edu. Address: 415 Buena Vista Avenue, Santa Cruz, CA 95062.*

The theoretical perspective in this study is that spelling reflects what learners know about the three layers of English orthography: the alphabet layer when sound-symbol correspondences are learned, the orthographic pattern layer, and the meaning in morphology layer. The sequence of learning specific features has been described in terms of five stages or phases. The five stages of spelling describe the spelling development and orthographic knowledge of English spelling (Henderson, 1980; Templeton & Bear, 2018). The first stage is the emergent stage when students are learning about phonological awareness, and the concepts of print and the spatial-temporal match between what we say and read. After the emergent stage, during the letter name-alphabetic stage, students focus on sound-symbol, and in English, letter-sound correspondences inherent in the alphabetic principle (Templeton & Morris, 2000). Students learn how to spell short vowels and many consonant digraphs and blends. This stage closes when students make generalizations about basic short vowel sounds that include commonalities in sound across a short vowel, and a tacit understanding of the closed syllable pattern containing short vowel, i.e., the highly phonetic short-vowel CVC pattern as in *bat*, *ball*, and *blank*. Having mastered these features, they begin to examine spelling patterns, a time when more abstract patterns are learned. This developmental stage is called the within word pattern stage, it is a time that has been described as an orthographic stage. In the pattern layer of English, students learn about long vowel (e.g., *meet/CVVC*, *time/CVCe*) and complex vowel patterns (e.g., *ou/around*, *ow/clown*). During the last two stages, students make the meaning connection as they learn about inflectional morphology (e.g., *hop/hopping*, *hope/hoping*, *talk/talking*), and then derivational morphology with roots (*ter*, *spect*, *duct*).

This study examines the spelling of English foreign language (EFL) learners in Mainland China from a developmental perspective (Templeton & Bear, 1992), an approach that has not been undertaken with Chinese speakers learning English. From this relatively large sample of errors, we wanted to know how the three layers of orthographic knowledge (alphabet, pattern, and meaning layers) are reflected in these students' spelling. The purpose of this study is to share an analysis of a wide range of orthographic features that may reveal developmental patterns in the acquisition of features. With this purpose in mind, three areas of investigation were undertaken:

1. What are notable developmental patterns in English spelling by Chinese EFL learners?
2. What cross-linguistic influences may impact the English spelling of Chinese EFL learners?
3. How does spelling achievement advance with schooling experience by grade level?

We begin with a brief description of developmental spelling, orthographic knowledge, and assessment.

Developmental Orthography

Orthographic Knowledge and Spelling Development

Phonological, orthographic, and morphological knowledge are described as components or layers of written English that are essential to learning to read, spell, and write (Berninger, Abbott, Nagy, & Carlisle, 2009; Blachman et al., 2013; Cho, McBride-Chang, & Burgess, 2005; Nagy & Townsend, 2012; Perfetti, 2007; Stanovich, 2000; Taft, 2003). A wide range of research that includes psycholinguistic, speech and language, anatomical, and neurolinguistic studies strongly suggests that underlying spelling and reading is a shared foundational knowledge (Beeson, Rising, Kim, & Rapsak, 2010; Cutler, Treiman, & van Ooijen, 2010; Ellis, 1997; Perfetti, 1997; Rapp & Lipka, 2010).

Developmental sequences have long been observed in reading (Biemiller, 1970; Invernizzi & Hayes, 2004) and spelling (Bahr, Silliman, & Berninger, 2009; Read, 1975; Templeton, 2003; Templeton & Bear, 2018), and these changes in reading and spelling have been described as stages or phases of learning to read and spell (Chall, 1983; Ehri, 1997, 2014; Wolf, 2007; Yin, Anderson, & Zhu, 2007). While there is evidence that spelling and reading share a foundation across the alphabet, pattern, and meaning layers, some researchers suggest that spelling and reading are different processes (Caravolas, Hulme, & Snowling, 2001; Ellis & Cataldo, 1990). It is suggested by others that spelling and reading achievement are highly correlated (Ehri, 1997; Ellis, 1997; Foorman & Petscher, 2010; Invernizzi & Hayes 2004; Mehta, Foorman, Branum-Martin, & Taylor, 2005; Templeton & Morris, 2000); for example, in a recent longitudinal study of Cantonese speakers, orthographic and morphological knowledge contributed to spelling (Yeung, Ho, Chan, & Chung, 2013).

A reciprocal relationship has been observed between reading and spelling with the consistent finding that spelling instruction can affect reading achievement, vocabulary, and morphological knowledge (Conrad, 2008; Ehri, 2006; Ehri, 2014; Graham, Harris, & Chorzempa, 2002; Graham & Santangelo, 2014; Nunes, Bryant, & Bindman, 2006; Richards et al., 2006; Rosenthal & Ehri, 2008; Taft, 2003; Treiman, 1998). Likewise, when students learn about inflected and derivational morphology during the last two stages, a reciprocal relationship between reading and orthographic knowledge has been observed in studies of morphological knowledge, word knowledge, and reading comprehension (Carlisle, 2007; Corson, 1997; Nagy, 2007; Nagy, Berninger, & Abbott, 2006).

Developmental Spelling Stages in English

The three layers of orthographic development, alphabet, pattern and meaning, have also been described in terms of developmental phases or stages (Ehri, 2015; Templeton & Morris, 2000). In the assessments used in this study with 273 students, 15 features were examined that match the developmental continuum of the five spelling stages of development described by Henderson (1981) and his students (Templeton & Bear, 2018).

This study builds on the work of Charles Read (1971, 1975, 1986), who initially found that children's invented spellings are not random but are based on the learners' developmental theories of English spelling. Henderson and his students (Henderson, 1981, 1992; Henderson & Templeton, 1986; Morris, 2001) built on Read's work and described five stages of spelling that run parallel to the three layers of English orthography, alphabet (sound), pattern, and meaning (Henderson, 1974, 1992; Henderson & Templeton, 1986; Morris, 2001; Templeton & Morris, 2000). These five stages of development have been useful to understand children's orthographic development and knowledge (Lee & Scanlon, 2015) and are described here as a framework for the qualitative analyses of Chinese speakers' spelling in English. We look at these five stages quantitatively by the number of features students spell; we also collect and analyze qualitatively the spelling errors and examine errors for possible translanguage contrasts.

The *emergent* stage of spelling begins before school (ages 1–5). The spelling during this stage is mostly pre-phonetic, and the writing consists mostly of scribbles. During the latter part of the emergent stage, students have learned the names of several letters of the alphabet, and they may represent the beginning or most prominent sounds in a word (L for *elephant*). The reading of children in this stage is based on what they know about the text from pictures or having heard the story before (Ehri, 1997; cf. Biemiller, 1970).

In the *letter name-alphabetic spelling* stage (ages 5-8), children learn letter-sound correspondences, and they use the name of a letter to represent the sounds of the letter, such as spelling *mess* as MS and *help* as HLP because the names of the letters *s* and *l* are pronounced as /es/ and /el/ (Treiman & Cassar, 1997). Spelling errors often reflect how spellers use articulation to spell (i.e., how a sound is articulated, or feels in the mouth, influences a learner's spelling) (Read, 1971, 1975). During this stage, children use a letter name strategy to spell vowels. Long vowels are usually easy to represent because the names of the vowels match the letter names (HOP for *hope*). To spell short vowels, students in this stage often spell the short vowels with the long vowel name closest in articulation (FES for *fish*) (Read, 1975). For example, the /ɪ/ in *fish* and the letter name *e*, which is pronounced /i/, are articulated in a similar place; they are both high-front vowels (Finegan, 2008). By the end of this stage, children have a full understanding of the alphabet layer of English orthography and are capable of spelling most short vowels and consonant digraphs and blends, including preconsonantal nasals (Templeton & Morris, 2000).

Mastering the alphabetic principle (i.e., that certain letters correspond with certain sounds) is learned in kindergarten and first grade, and is highly related to concept of word in text and the ability to match syllabic units (Morris, Bloodgood, Lomax, & Perney, 2003; cf. Goswami, 2006). Once letter-sound correspondences are learned in English, the pattern layer comes into play first with students developing the idea that the CVC, closed syllable pattern, is a short vowel spelling.

In the *within word pattern stage* (ages 7-10), students learn the orthographic patterns for spelling long and ambiguous vowel patterns in English. Their spelling errors during this stage reflect growing knowledge of orthographic patterns, particularly long vowel patterns as when they spell *team* as TEME or TEEM. Gradually, they learn the correct spelling of long vowel words that include the consonant-vowel-consonant + e-marker (CVCe; i.e., *drive*), consonant-vowel-vowel-consonant (CVVC; i.e., *train*), and consonant-vowel-vowel (CVV; i.e., *say*) patterns, such as when they spell *team* as TEEM or TEME, and gradually they learn to spell nearly all single

syllable words correctly (Invernizzi, Abouzeid, & Gill, 1994). Learning to spell long vowel patterns in English grows in parallel to learning to spell low frequency consonant di- and trigraphs in initial and final positions. As children's sight vocabularies grow and the regularities of long vowel patterns are learned, complex vowel patterns are learned. Learning these aspects of the developmental sequence can take some students to the middle of third grade, though many students learn common long vowel patterns by the end of first grade. In the latter part of this stage, children learn about vowels that are neither long nor short and which may be ambiguous or inconsistent, such as the *ou* in *mouth, cough, through, and tough* (Gehsmann & Templeton, 2011, 2012).

Students in the fourth stage, *syllables and affixes spelling* (ages 8-18), expand their orthographic knowledge as they stretch into the meaning layer. As they learn to spell polysyllabic words they experiment with morphological elements like inflections (e.g., *stopped/hoped, shopping/hoping*), and past tense endings; e.g., that *-ed* signifies past tense regardless of the suffix's three pronunciations (/t/, /d/, or /ed/). They become familiar with the structure of syllables and principles of spelling including consonant doubling and the structure of open and closed syllables in relationship to the vowel; e.g., *pilot/napkin* and exceptions like *pivot*. The meaning layer is also seen in their learning the spelling and meaning of morphological elements like prefixes and suffixes.

Students may move into the last stage, *derivational relations spelling* as early as grade 4 or 5 (age 9 or 10); although the majority of them enter this stage in middle, high school, or even college. As the name implies, derivational relations spellers become aware of the derivational relations among words in terms of roots, origins, and meaning. They discover that words such as *compete, competition, and competitor* are connected in meaning, and therefore are similar in spelling pattern, even though they sound slightly different (Templeton, 2003). Students benefit from discovering spelling-meaning connections because it enables them to continue expanding their vocabulary throughout their lives. Hence, this is a lifelong stage (Henderson, 1990).

The concepts of phases or stages can be useful conceptually to describe a gradual progression in learning specific features. For example, students learn first about short and then long vowel patterns in parallel to their learning to spell consonant blends and digraphs. The relationships between reading and writing will vary with the structure of the orthography (Carlisle, 2010; Helman, Delbridge, Parker, Arnal, & Jara Mödinger, 2015) and the progression through the three layers varies with the particular orthography (Helman, 2004; Shen & Bear, 2000; Templeton & Morris, 2000).

The sequence of development reflected in these five stages is evident in the spelling development of children learning English as a second language where the impact of students' primary oral and written languages has been examined (Ford, Invernizzi, & Huang, 2014; Helman, 2004). These translanguage errors are useful for assessment and instruction among learners of different languages, in this case, Chinese students learning English as a foreign language. The next section presents a discussion of Chinese orthography and spelling to suggest features that may be seen in the cross-language spelling of students in grades 2-8.

Chinese and Pinyin

Pinyin functions as an indispensable tool like the alphabetic principle in Western writing systems (Share, 1995). Pinyin is introduced to Chinese children in first grade and typically learned by the end of the year. They become fluent in pinyin and continue to use it throughout their primary school years. In their early years of writing when their character knowledge is limited, students write in pinyin to substitute for the characters they do not know (Cheung & Ng, 2003). After learning pinyin, students receive exposure to pinyin in conjunction with the Chinese characters to help them read and learn the logographic characters and understand the phonetic pronunciation of the characters.

Progression through the three layers of alphabet, pattern and meaning and a sequence of spelling development in Mandarin has been observed in patterns of spelling. Spelling accuracy increased significantly across grade levels, and the quality of the errors over the grades suggested a progression that related to the sound-pattern-meaning layers of Mandarin orthography (Shen & Bear, 2000). In an analysis of 7,000 spelling errors classified into 15 categories from the writing of 1,200 children in grades 1 through 6 in Mainland China, nearly 80 percent of the spelling errors had some phonological base, decreasing from 96 percent in first grade to 53 percent in sixth grade. Errors classified as pattern-type errors increased from 4 percent in first grade to 33 percent in sixth grade. Meaning-type errors jumped from 0.3 percent in first grade to 11 percent in sixth grade. Similar findings for responses to morphological instruction support the utility of addressing this morphological

layer of Mandarin (McBride-Chang et al., 2008). The development of the three layers reflects the depth and frequency of characters in Mandarin orthography.

Cross-linguistic Spelling Development among English Learners

Building on English-L1 developmental spelling research, and using it as a source of comparison, cross-linguistic researchers have produced a growing body of literature on the influences of English learners' L1 on their English spelling and orthographic knowledge and development. Researchers have found that English learners make orthographic errors that often occur when a single phoneme is represented by different letters in different languages (Cook, 1997; Fashola, Drum, Mayer, & Kang, 1996). For instance, Fashola et al. (1996) found that Spanish-L1 children spelled *happy* as *japi* because /h/ is represented by the letter *j* in Spanish and by the letter *h* in English. In Japanese, Cook (1997) found that Japanese-L1 English learners confused /l/ and /r/ sounds in their spelling reflecting phonological differences in languages (Thompson, 2001). The following examples illustrate how English learners make phonetic errors, which often occur when a sound does not exist in the L1. Wang and Geva (2003a) found that Cantonese-L1 children substituted *sh* with *s* and *th* with either *s* or *z* due to the absence of these two phonemes in Cantonese. Similarly, Morris (2001) discovered that French-L1 children tended to spell *house* as OUSE, which is likely due to the absence of /h/ in the French language (Walter, 2001). At the same time, research suggests that L1s can be facilitative, depending on the nature of the two writing systems. In general, studies have shown that learners with an alphabetic L1 background usually do better on word identification and phonological awareness tasks than those with a non-alphabetic L1 (Leong, Cheng, & Tan, 2005; Wang, Perfetti, & Liu, 2005).

Some studies, such as Hamada and Koda (2008) and Wang, Koda, and Perfetti (2003), have illustrated learner performance differences on tasks due to varied L1 backgrounds. A number of studies have also documented that learners with the same written L1 background vary in word recognition abilities and phonological tasks given their variation in exposure to the alphabetic principle (Leong, Hau, Cheng, & Tan, 2005). Bertelson, Chen, and de Gelder (1997) found that students from Mainland China outperformed students from Hong Kong on tasks related to phonemic awareness, which is likely due to the Mainland learners' familiarity with pinyin that students in this study from Hong Kong lacked. In a recent meta-analysis of fMRI studies, language and writing processing networks in Chinese are largely similar to alphabetic language processing but there are distinct differences related specifically to Chinese (Wu, Ho, & Chen, 2012).

The facilitative effect of pinyin is also shown in research by Leong, Hau, Cheng, and Tan (2005), who investigated phonological sensitivity, and Cheung, Chen, Lai, Wong, and Hills (2001) in a study of phonological awareness among pre-readers in Hong Kong, Guangzhou in Mainland China, and New Zealand and by McDowell and Lorch (2008) among Chinese speaking students 15-19 years-old learning English.

To date, there are only a handful of studies that have examined the development of cross-linguistic spelling. Within this limited number of studies, researchers have found that English language learners develop their basic literacy skills in a similar manner to English-L1 children (Chiappe, Siegel, & Gottardo, 2002; He & Wang, 2009), and their progression may be influenced by various factors, including their L1 (Wang & Geva, 2003a, 2003b), with a modest impact for phonological influences on learning English as a foreign language. Early studies revealed that younger children made more L1-influenced errors than older children, and as their grade level advanced, the number of L1-influenced errors decreased; likewise, as English proficiency grew, English learners used more of their knowledge of English and spelled more English words correctly (Fashola et al., 1996; Ferroli & Shanahan, 1993;).

Longitudinal studies of the evolution of specific features in spelling by emergent bilinguals have also been illuminating (McBride-Chang, Liu, Wong, Wong, & Shu, 2012). In a study by Wang and Geva (2003a), younger bilingual learners (Grade 1) made more predictable L1-specific phonological errors in their spelling than their English-L1 counterparts. By the end of Grade 2, these ESL spellers performed as well as their English-L1 peers. Nassaji (2007) analyzed the development of English spelling and orthographic knowledge of a Farsi-L1 student's spelling errors from his daily journals and free writings over a four-year period, from grade one through four. Nassaji found that the child's spelling followed a stage-like sequence in a similar fashion to English L1 speakers, in which the child developed more complex spelling abilities and accuracy over time.

To summarize, previous cross-linguistic research indicates that English learners develop their basic literacy skills in a similar way to English-L1 children, and as their English proficiencies grow their spelling abilities also develop. As students' knowledge of English develops, English learners make fewer L1-influenced errors. To examine the three components or layers of English spelling cross-sectionally across grades 2–8 by features and developmental stages, this study focused on how Chinese, Mandarin-L1 students learning English as a foreign language spell both quantitatively and qualitatively.

Method

Participants

273 students from two primary schools and two middle schools in northeastern China participated in the study (Grade 2: N = 41, Grade 3: 32, Grade 4: 40, Grade 5: 40, Grade 6: 40, Grade 7: 40, Grade 8: 40). They ranged in age from seven to 14. At the time of data collection, each participant had at least one academic year of English instruction as a foreign language as part of the regular curriculum. They received one hour of English instruction each day. All participants spoke Mandarin Chinese at school and at home.

Measures

Students' spelling was assessed with two spelling inventories (Bear, Invernizzi, Templeton, & Johnson, 2016) designed to measure the students' spelling knowledge. Inventories consist of 25 words (grades 2 & 3) and 26 words (grades 4–8), both of which have undergone a validation study (Sterbinsky, 2007). Two spelling inventories were used to provide developmentally appropriate words and spelling features for the participants who varied by age, ranging from second grade to eighth grade. While these spelling inventories are different, they largely measure the same spelling features, ranging from initial consonants to inflected endings, and the more advanced spelling inventory also measuring more challenging spelling features, such as syllable junctures and roots. In the classroom setting, the classroom teacher followed the inventory instructions, read the directions in Chinese and read each word, followed by a sentence in English using the word, and then the word again. Participants were asked to spell the best they could even when they did not know how to spell a word.

Sterbinsky (2007) conducted a reliability and validity analysis of both spelling inventories in which 647 students completed the Primary Spelling Inventory and 862 students completed the Elementary Spelling Inventory. The results indicate that both inventories are “reliable instruments and valid predictors of student achievement” (Sterbinsky, 2007, p. 19).¹ Twelve features are studied in this inventory in a sequence that reflects the three layers (alphabetic, pattern, and meaning) of English orthographic knowledge, and the corresponding stages of spelling. Emergent stage spelling was not evident in this sample, so the features we examined begin with (1) beginning consonants, (2) ending consonants, (3) short vowels, (4) consonant digraphs and (5) blends (e.g., *bed*, *ship*, *when*, *lump*). In the pattern layer of English spelling during the within word pattern stage of spelling development, students' knowledge of (6) long vowel patterns (*float*, *train*, *place*, *drive bright*, and *throat*), and (7) other vowel patterns including diphthongs (*spoil*) and r-influenced vowels (*serving*) are examined. Upper level spelling knowledge for the syllables and affixes stage begins with (8) inflected endings (*serving*, *chewed*, *carries*, and *marched*), then includes an examination of other features with words with (9) syllable junctures and (10) unaccented syllables (*shower*, *bottle*, *favor*, *ripen*, *cellar*). Finally, derivational morphology and knowledge of (11) more difficult affixes and (12) roots (*pleasure*, *fortunate*, *confident*, *civilize*, and *opposition*) are examined.

Data Analysis

Students' spellings were analyzed for words and features spelled correctly, and scores for the total features and words spelled correctly were obtained for each student. Two trained college students at a research institution in the western United States first coded the data. An experienced researcher then coded 70 students' spelling (10 in each grade) to confirm the errors had been identified accurately. The handwriting posed a greater than usual challenge for the scorers, and there were times when the children used a character to spell a word. Discrepancies in scoring were discussed between the two scorers, and when agreement could not be reached, the first and second authors were consulted. The scoring guide for the features has been used extensively in research and

teaching (Bear, et al., 2016), and the mean scores and standard deviations for this study are presented in Tables 1 and 4. The feature types and their developmental order are presented from top to bottom for the primary inventory in Table 1 and the elementary inventory in Table 4. The specific features and words are presented in the following discussion of the data.

For informal, post hoc analyses, spelling stages were matched to power scores, the number of words spelled correctly. These criteria were set developmentally (Bear et al., 2016), and a range of development was described for grade levels by power scores. While the two different spelling inventories measure many of the same spelling features, the results of the two spelling inventories were not compared from one inventory to the other. This approach was utilized to illustrate how each age range performed on their respective spelling inventories.

Results

Grades 2 and 3

As described earlier, students' spelling abilities were analyzed by the total number of words and features spelled correctly. This section will briefly examine descriptive statistics, and then explore specific features in detail. The list of specific spelling features addressed in the results section is not exhaustive, but rather focuses on notable findings, and features that characterize the stages of development. Table 1 presents the means and standard deviations from independent *t*-tests for second and third graders. Students in both grades experienced difficulty with short vowels. More students in grade 3 struggled when spelling the final consonants than students in grade 2 ($t = 2.17$, $df = 71$, $p = .03$, $d = .60$).² No significant differences ($p \leq .05$) were observed in the other features examined. We continue with an analysis of the categories of orthographic development.

Table 1
Means and Standard Deviations of Spelling Features in Grades 2-3

	2 nd (n = 41)			3 rd (n = 32)	
	Max.	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Words Correct	26	2.22	2.29	2.06	1.72
Features Correct	56	21.90	5.69	21.53	4.74
Total Score	82	24.12	7.55	23.59	6.14
Features					
Initial Consonants	7	6.56	.63	6.44	.91
Final Consonants ³	7	5.78	1.29	5.16	1.11
Short Vowels	7	2.22	1.51	1.97	.82
Digraphs	7	3.54	1.82	3.91	1.51
Blends	7	2.44	1.30	2.21	1.24
Long Vowels	7	.34	.62	.50	.76
Diphthongs, and R-influenced Vowels	7	.27	.63	.56	.76
Inflected Endings	7	.76	1.02	.79	.97

Note. Total N=73.

Consonants

As noted in Table 1, students in second and third grade spell beginning and final consonants with great accuracy. This illustrates their expertise with this feature, which is much higher than their spelling of other features. The words for these features are regular and familiar; only the *r* in *rob* might be difficult. The beginning consonants were spelled accurately by second graders 94% of the time, compared to 83% for the third graders.

Short vowels

Short vowel sounds were difficult for the second- and third-grade students to spell correctly. The most difficult short vowel was *u* as in *gum*, with only 2.74% of the students spelling it correctly. Forty-one percent of the students spelled *gum* as *gam*. This is likely because the *u* in *gum* (represented phonemically as /ʌ/, which is similar to /ə/) “is sometimes replaced by /a/” (Chang, 2001, p. 311). This case of misspelling is also likely influenced by the place articulation, as all of these are central vowels; /ʌ/ and /ə/ are mid-central vowels and /a/ is a low-central vowel (Finegan, 2008).

The *e* as in *sled* was the second most difficult short vowel sound for the 2nd and 3rd grade participants. This is likely related to the lack of a distinct /ɛ/ phoneme in Chinese (Duanmu, 2006). The most frequent replacement was *a*. The letter *A* was also the most frequent substitution for *e* as in *pet*. These instances may be influenced by how /ɛ/ in *sled* sounds like the letter name *a*, which is pronounced /e/; both /ɛ/ and /e/ are front central vowels (Finegan, 2008). Table 2a displays the most prevalent misspellings of short vowels for the second and third graders.

Table 2a
Misspelling of Short Vowels by Children in Grades 2-3

Short Vowels	% Feature Correct	Misspellings (%)
fan	65.75	fun (10.96); fain (4.11)
pet	26.03	pat (26.03); pait (5.48)
dig	53.42	deg (9.59); dg (6.85)
rob	24.66	raob/raop (15.07); roub/roubo/rout (8.22)
gum	2.74	gam (41.10); garm (4.11)
sled	10.96	slad (13.70), slaid (9.59)
stick	28.77	sdk (12.33); sdek/sdekc/sdeke (9.59)

Note. N=73. In the Misspellings column, similar types of errors were combined as noted.

Consonant digraphs

Compared to spelling short vowel sounds, students spelled more digraphs correctly. A notable exception was the /θ/ sound as in *thorn* and *third*, which had lower rates of accuracy. The other digraphs did not impose a significant obstacle for most students. This may be because pinyin has similar digraphs (e.g., *ch* and *sh*), but lacks a *th* digraph, which is represented in the spelling inventory and in this case, matches the /θ/ phoneme. Most students replaced it with either *s* or *f*, two common substitutions for the /θ/ sound among Mandarin speakers (Rau, Chang, & Tarone, 2009). This substitution makes sense as all three sounds are voiceless fricatives, and thus, only differ in the place of articulation. The students replaced an unfamiliar sound, a /θ/ (an interdental fricative), with a familiar sound, either an /s/ (an alveolar fricative) or an /f/ (a labiodental fricative). Table 2b shows the most frequent misspellings of digraphs.

Consonant blends

Students had trouble with the majority of the blends except *sl* as in *sled* and *fr* as in *fright*. Table 2c displays this analysis. Students often inserted a vowel between consonants and wrote some single-syllable words as two-syllable words. This is likely because there are no consonant clusters in Chinese (Chang, 2001). As shown in Table 4, students added either *i* or *a* between the *s* and the *l* in *sled*, and *e* or *u* between the *b* and the *l* in *blade*; however, this addition was not consistent.

There are a few reasons why students' might have spelled *cr* as *kr*. One is that there is also a /k/ sound in pinyin that is represented by the letter *k*. Consequently, they may have made an error based on their orthographic knowledge in pinyin. Another possible reason is that in English, the /k/ phoneme can be spelled as

c, *cc*, *ck*, *k*, and *ke*. While learning to spell in English as a foreign language in their early years, these students were less familiar with the varied spellings of the /k/ phoneme in English where there is a greater frequency of initial *c*-blend compared to a *k*-blend.

Interestingly, students often replaced the *tr* in *tries* with *ch*, affrication, a logical linguistic confusion that is often observed at this point in students' development. Additionally, many students also replaced the *dr* in *drive* with a *j*. This phenomenon represents how when alveolar stops are followed by an *r*, palatalization can occur. Thus, *t* becomes /tʃ/ (or *ch*) and *d* becomes /dʒ/ (or *j*). This is an interesting finding that aligns with previous studies of the development of L1 English children (Henderson & Beers, 1980; Read, 1975).

Table 2b
Misspelling of Digraphs by Children in Grades 2-3

Digraphs	Feature Correct (%)	Misspellings (%)
sh ine	78.08	sun/sunny (12.33); sai/san/saie/sam/sane (6.58)
co ach	67.12	kouh/holh/coldh/cdh,/coh (9.59)
ch ewed	71.23	cug/cud/coud/cund (9.59) treed/tryd/trd (5.48)
w ish es	60.27	wsas/ws/wesds/wsaz/weiseis (6.85); wrhz/wxhz/wszh (5.48)
th orn	5.48	saon/sone/son/sonjn/sorn/sosoht/sen/sanyin/san/sany/saone/ sonn/seen/sloor (42.47); fao/faoen/foth/faon/fon/foog/fan/foan/foin/forh/foun/ford/ fane/found/fonu/foon/fng (30.14)
sh outed	72.60	soutb/sotd/saoied/saotd/saofeit/sahtd/std/soutd/suo/suts (15.07)
th ird	31.51	sed/serd/sede/st/srot/srt/sd/seed/srerd/sot/snes/sered/srid/ sterl/seth/stas/sorasd/sride/sead/sends/sensd/sred (30.14); 3ed (2.74)

Note. N=73. In the Misspellings column, similar types of errors were combined, and the subsequent percentage illustrates the total percentage of spellings with that type of error.

Table 3
Misspelling of Blends by Children in Grade 4

Blends	Feature Correct (%)	Misspellings (%)
lump	2.50	lam (25.00); larm (5.00)
fl oat	85.00	fol (5.00)
tr ain	70.00	tain (10.00); tian (5.00)
pl ace	97.50	--
dr ive	62.50	diver (7.50)
br ight	47.50	blight (7.50); grit (7.50)
sp oil	37.50	sbout (5.00); sbor (5.00)

Note. N=40. Only the two most frequent misspellings are reported. Misspellings only occurring once are not included in this table.

Long vowels, diphthongs, and r-influenced vowels

On average, second and third graders spelled less than one long vowel and one diphthong and r-influenced vowel correctly (see Table 1). Some vowels were spelled with two or three vowels; e.g., *CHUAIS/CHUAIZ* for *tries* and *GEEM/GAIME* for *dreams*. Some words are more frequent and spelled more accurately than pattern would belie. For example, the word *hope* was spelled correctly by 8 of the 63 second and third graders. The word *dream* was not spelled correctly by any child. The chance for error is much greater for *dream* than a word like *wait*; the word *dream* is less common and is a more difficult word to spell correctly with the *dr* affricate, a vowel digraph, and a final nasal that can be confused by English learners more so than with native English speakers (Helman et al., 2012). This pattern continues with the vowel patterns in the *other vowels* category; and we see that no second or third graders spelled the following words in this category correctly: *chewed*, *crawl*, *thorn*, *shouted*, *spoil*, and *growl*. The word *third* was spelled correctly by nine children who may have learned the word through reading and instruction. Specifics for a number of the features in this category are discussed below; we have chosen to discuss the most salient features given space limitations.

Inflected endings

This was a difficult feature for students to spell correctly. None of the second and third grades spelled *chewed*, *shouted*, and *camped*, correctly and only one student spelled *wishes* and *tries* correctly. Fewer than one of seven features were spelled correctly; third grade mean was .79 for spelling words with inflected endings correctly, except for the frequently occurring word *riding*, spelled correctly by 13 students, and the word *clapping* (7), a word with a fairly unambiguous vowel and syllable pattern.

Developmental spelling analyses

As seen in Tables 2a and 2b, the mean score was slightly above 2 words spelled correctly. This score has been used as a power score to relate to stages of spelling development, and for these students a score of 2 indicates that a student is in the letter name–alphabetic stage of spelling. For context, this stage also corresponds roughly to a lexile level (see lexilie.com) of up to 300, a beginning reader, first grade level. Consider that the first four words (*fan*, *bet*, *dig*, *rob*) were the easiest words to spell correctly, accounted for the vast majority of errors, and reflected knowledge of short vowels. The feature analyses as noted in Tables 2a and 2b also indicate a beginning knowledge of short vowels (2.22/7) compared to final consonants (5.78/7).

Qualitative analyses across features revealed similar patterns of the letter name–alphabetic stage in which students use a phonetic principle that incorporates articulatory information to spell, a strategy shared by English-L1 students. For instance, similar to English-L1 children, some of these Chinese students spelled like early letter name–alphabetic stage spellers when they deleted vowels. Examples include spelling *dig*, *bled*, and *stick* as *dg*, *bld*, and *sdk*, respectively. Participants also spelled *dr* as *gr*, which is similar to their English-L1 counterparts. To spell vowels, students in the middle of this stage may have employed the letter-name strategy with the alphabetic principle and used their knowledge of a vowel’s name to spell long and short vowels (e.g., *HOP* for *hope*, *GREM* for *dream*, or *PAT* for the short vowel in *pet*) (Helman et al., 2012).

Grades 4–8

The spelling inventory included several more difficult words including the last two words, *civilize* (0 correct spellings) and *opposition* (4 correct spellings). Table 3 presents the means and standard deviations of these analyses. As expected, the total words spelled correctly and feature score totals were highly correlated in a Pearson correlation ($r = .94$, $p < .001$). Statistical analyses confirm the findings of previous research in English as a second language that learners gain orthographic knowledge as their English language proficiency advances; in this instance, spelling achievement predicted grade level, the proxy for experience with English, with the total feature and words correct score accounting for 44 percent of the variance ($F(1,198) = 154.79$, $p < .000$), a large effect size (Cohen’s d of .98 based on r -squared). A factor analysis of the raw scores of each feature was computed across fourth through eighth grade levels. The Kaiser-Meyer-Olkin measure of sampling adequacy indicates that the factor analysis is appropriate for these data (.91). The general sequence of features followed the predicted developmental sequence except for the consonant digraph feature. The first factor included the first eight features except for the consonant digraph category and together, they accounted for 49.9% of the variance. The second factor included for the last two features, harder suffixes and roots, accounted for 10.39% of the variance.

Table 4
Means and Standard Deviations (Grades 4-8, n = 160)

	4 th (n = 40)		5 th (n = 40)		6 th (n = 40)		7 th (n = 40)		8 th (n = 40)		
	Ma x	M	SD	M	SD	M	SD	M	SD	M	SD
Words Correct	25	3.50	2.19	8.83	2.82	10.28	2.09	9.25	2.47	11.65	2.42
Features Correct	62	22.23	4.93	36.58	5.41	39.48	4.60	36.23	6.09	42.58	4.98
Total Score	87	25.73	6.90	45.40	7.88	49.75	6.54	45.48	8.25	54.23	7.21
Features											
Initial Consonants	2	1.88	0.36	2.00	0.00	1.98	0.16	1.75	0.44	2.00	0.00
Final Consonants	5	3.75	1.02	4.85	0.36	4.93	0.27	4.83	0.38	4.95	0.22
Short Vowels	5	2.85	0.97	3.83	0.96	4.15	0.48	3.90	0.55	4.08	0.53
Digraphs	6	4.83	1.01	5.10	0.63	5.00	0.88	5.25	0.81	5.53	0.60
Blends	7	4.05	1.17	5.40	1.13	6.25	0.71	5.38	0.90	5.68	0.80
Long Vowels	5	1.00	1.07	3.10	0.90	3.38	0.63	3.48	0.72	4.00	0.45
Diphthongs and r- influenced vowels	7	0.65	0.80	2.23	1.23	3.20	1.22	2.35	1.23	3.25	1.43
Inflected Endings	5	0.98	0.87	2.50	1.30	3.45	0.93	2.63	1.39	3.78	0.89
Syllable Junctures	5	1.28	0.72	3.60	0.78	3.63	0.81	3.18	1.08	3.60	0.71
Unaccented Final Syllables	5	0.83	0.88	2.28	0.91	2.30	0.79	2.25	1.28	2.80	0.97
Harder Suffixes	5	0.15	0.36	1.58	1.08	1.03	0.95	1.03	0.89	2.08	1.05
Bases or Roots	5	0.00	0.00	0.13	0.40	0.20	0.61	0.23	0.48	0.85	1.03

Short vowels

Students in grades 4 through 8, spelled 3.76 of the 5 short vowel features correctly. Like so many of the second and third graders, 30 percent of the fourth graders, and 63 percent of the fifth graders spelled the *u* in *lump*, with an A. A quarter of the students in grades 4 and 5 spelled the *e* in *bed* with an A, also a common error among students in grades two and three, and students among English-L1 students in the middle letter-name stage (Bear et al., 2016). Closeness in articulation of the short vowel with the long vowel letter name may account for this error (Read, 1975).

Consonant blends and digraphs

Blends and digraphs were spelled correctly more often than the vowels; 5.14 of 6 blends and 5.35 of 7 digraphs were spelled correctly. The majority of the students spelled at least one of the features in the blends (e.g., *m* as *mp* and *t* as *tr*). One possible reason for these errors is the absence of consonant clusters in Chinese (Chang, 2001). Again, these errors are expected of students in the letter name – alphabetic stages by English L1 learners (Bear et al., 2016).

Similar to their peers in grades 2 and 3, students in grades 4 through 8 had difficulty spelling unaspirated sounds, such as *p* in *spoil*. Expectedly, some of students spelled *sp* as *sb*. This may be related to the voicing differences in *p* and *b*; additionally, the vowel sound after the onset is voiced, which could have affected the students' interpretation of the previous consonant sound. Students in grade 4, in particular, had trouble with the blends. However, unlike students in the lower grades, these students did not often insert a vowel between

consonants. This result might suggest that as their experience with English increased, the influence of a phonetic strategy from their L1 faded. Table 4 illustrates the misspelling of blends by grade 4 students.

Additionally, the *mp* in *lump* was difficult for students in grades 4-8. Only 12% of these students spelled *lump* correctly attributable to both errors in the *u* and the preconsonantal nasal in *mp*. Again, this is likely because of the lack of final consonant clusters in Chinese. Learners often insert a reduced vowel or as we found, particularly with fourth grade students, “simplify the cluster...by dropping the last consonant” (Chang, 2001, p. 312).

Long vowels

Almost three of the five long vowels were spelled correctly in fourth through eighth graders. The majority of the students found it difficult to spell *oa* in *float* and *igh* in *bright*. Most students either used a letter-name alphabetic stage misspelling (*flot*), the most common error of sixth and seventh graders (23%), or they applied their English knowledge of the CVCe pattern to produce the /o/ sound with *flote*, an error made by 14% of the fourth through eighth grade. Students also spelled *float* with other possible variants such as *ou* and *ow*. Students who spelled *float* with *ow* may have been using their knowledge of other sight words, like the /ow/, long o in *flow*, *mow*, *row*, *tow*. Using two vowels to spell the long o reflects students’ knowledge of orthographic patterns in English, a move beyond the more linear, phonetic strategies used during the letter name – alphabetic stage.

The misspelling of *igh* as in *bright* can be understood in a similar way. The most common substitution for *bright* was *brite*, just as *flote* was a common error for *float*, (12% by fourth through seventh grade) which may reflect their experimenting with orthographic patterns of long vowels. The students’ spelling errors are similar to those of English-L1 students in the early within word pattern stage (Templeton & Bear, 2018). As older students made fewer of these types of errors, these results again indicate that as students’ knowledge of English expanded, they tended to make fewer L1-influenced errors. There was an interesting jump in the correct spelling of *bright* by 8th graders (92.50%), a change from 55% in seventh grade. Perhaps this word was taught directly or learned through reading. In contrast, *float* was spelled correctly by just a few eighth graders and none in grades four and six.

Diphthongs, r-influenced vowels, inflected endings, and syllable junctures

Students’ lower accuracy rates for the spelling of “other vowels” that included diphthongs and r-influenced vowels, may be related to the fact that in pinyin there are no vowel digraphs. In pinyin, all vowels sound the same in length. Across all grades, many misspellings of *other vowels* were substitutions of short vowels for ambiguous vowels, such as spelling *in* in this case, the r-influenced vowel in *marched* was spelled *mucht*, an error which is also made by English-L1 students in the early within word pattern stage (Helman, 2004). Furthermore, the students in grade 4, had particular difficulty identifying the inflected ending in *marched* (hence the misspelling *much* by over 57%), a word that has a complex final consonant cluster /ɪʃt/ that is not permitted in Chinese phonology. However, as students progressed they were more accurate in their identification of sounds (*marcht*) and use of inflected endings (*marched* spelled correctly by 60% of the students in grade 4).

In the last two stages of spelling with the analysis of syllable junctures, the consonant doublet was difficult to spell across all later grade. Gradual competency spelling words with doublets increased for spelling these words correctly at 34.8% for the word *shopping*, 21.2% for *carries*, and 25.8% for *bottle*. The word *cellar* was gradually spelled more accurately each year, up to 32.50 percent by eighth grade, an unsurprising error rate for this unfamiliar word.

The final five words on the spelling inventory asked students to spell five words with less frequent suffixes, and five word roots. Less frequent suffixes (*-ent*, *-tion*) and word roots (*civ*, *fid*) were rarely spelled correctly. Nineteen students spelled *pleasure* correctly, two spelled *fortunate* correctly, and no children spelled the last word, *confident*, correctly.

Developmental spelling analyses

The mean number of words spelled correctly for the fourth through eighth graders in this study was 8.70 which corresponds to the later part of the within word pattern stage of spelling development, a 420-820 extended lexile level range, and a second-grade reading level (Bear et al., 2016). There was a range from 3.5 words spelled

correctly for 4th graders to 10.275 for 8th graders. Accordingly, there were many students who did not spell the short or long vowel features correctly (3.76/5; 2.99/5, respectively). See Table 4 for mean scores by grade. There were difficulties spelling short vowels, particularly *e* and *u*, and they spelled more consonant digraphs and blends correctly than short vowels. Students in grade 4 also had more trouble with blends than their older, more experienced peers. Like the younger spellers in grades 2 and 3, the greater the error rate, the greater the number of variants they produced. The eighth graders were refining their knowledge of other vowels (e.g., *our*, *ow*) at the same time that they were learning about syllable junctures and inflected morphology. They were spanning the focus from basic vowel patterns of single syllable words to a beginning analysis of morphemes and unaccented syllables.

Discussion

The examination of students' spelling showed them to master features in an order that reflects the three orthographic layers and the five stages of spelling. The second- and third-grade children showed greater proficiency with features associated with the alphabetic layer of English orthography (e.g., consonants, short vowels, digraphs, and blends), than with the pattern layer features (e.g., long vowel patterns), and the meaning layer (e.g., inflected endings, base and roots). These children did quite well spelling beginning and ending consonants, features of the alphabetic layer, and they did better spelling short vowels, blends, and digraphs than features associated with the pattern and meaning layers of English orthography. Native English-speaking children likely find short vowels easier to learn than these Chinese speaking EFL students as they match familiar sounds with a single letter. Chang (2001) has shown how Chinese speakers have a variety of difficulties mastering short vowels, which may be influenced by the greater number "of vowel contrasts in English than in Chinese, so English vowels are closer to each other in terms of position of articulation than Chinese vowels. This means that more effort is required to distinguish them" (p. 311). The better performance of Chinese students with digraphs over short vowels is perhaps influenced by this phenomenon and also suggests that their knowledge of consonants and consonant digraphs (e.g., *ch*, *sh*, *zh*) in pinyin may have served as a springboard to developing their orthographic knowledge of some digraphs in English.

After one or two years of instruction in English, most second and third grade students were centering on the alphabetic layer and in the letter name – alphabetic stage as evidenced in the way they spelled features associated with the alphabetic layer of English orthography (e.g., consonants, short vowels, blends, and digraphs). Overall, the results show that these Chinese students developed greater proficiency with the alphabetic layer before the orthographic pattern and meaning layers of English orthography, and while this pattern illustrates the general trend, the data also illustrate interesting differences between native English speakers and Chinese English learners in their development and proficiency of short vowels and digraphs, as illustrated above.

In the present study, the 173 fourth through eighth grade students' spelling abilities developed as their grade levels advanced. Spelling ability among students in grade 4 was less developed than that of students in later grades. Notably, students gradually progressed in their proficiency from the alphabetic stage to the pattern stage of English orthography. Students in fourth grade struggled substantially more than their older peers with spelling features related to the pattern stage (e.g., long and other vowels). Scores for spelling long vowels, diphthongs, and r-influenced more than tripled from fourth to fifth grade. This spike in development illustrates a shift from the alphabetic to pattern layer of English orthography, as students were developing their abilities to recognize and utilize long and other vowel patterns in which the same sound can be represented with different orthographic patterns (e.g., *fare*, *fair*). Furthermore, there are noticeable differences when comparing grade 4 students to grade 8 students in terms of ability to spell features associated with the meaning layer (such as suffixes and roots). However, their knowledge of both the pattern and meaning layers of English orthography is still in development. Ultimately, this illustrates how these older students also develop their knowledge of English orthography in a similar manner to native English speakers.

Analysis of the data reveals a few more interesting differences among students in different grades. While students in grade 8 outperformed students in the early grades, in most areas, students in grade 6 led students in both grades 5 and 7 on spelling three features: blends, diphthongs, and r-influenced vowels (like *ew*, *oi*, *or*), and inflected endings. This outcome might be influenced by the preparation students in grade 6 were receiving for

the entrance examination for middle school; their constant review of English may have helped their performance on the assessment.

There were occasions across all grades that students could spell some more difficult words than easy words. For instance, among second and third graders, more students spelled *riding* correctly (Grade 2: 6/41; Grade 3: 7/32) than *chewed* (Grade 2: 0/41; Grade 3: 0/32). Among students in grades 4 through 8, the word *bottle* was spelled correctly more often than an easier word, like *float* (Grade 4: 10 vs. 0; Grade 5: 30 vs. 1; Grade 6: 36 vs. 0; Grade 7: 26 vs. 0; Grade 8: 27 vs. 7). This outcome may be because these more difficult words might occur more frequently than the easy ones in their instructional context, and thus, students learn them earlier because of greater exposure. Nonetheless, taking the results from the spelling of students in grades 2 and 3 into consideration, it can be concluded that the influences of students' L1 became less of an issue as their experience with English advanced.

This study adds empirical evidence to the body of cross-linguistic literature that has found that English learners develop basic literacy skills in a similar manner to their English-L1 peers (Chiappe, Siegel, & Wade-Wooley, 2002). More specifically, the present study lends support to previous research that has suggested that learners of written English acquire their orthographic knowledge in a similar manner and vary primarily in the rate of acquisition (Invernizzi & Hayes, 2004).

While this study contributes to our understanding of how young Chinese students develop their English literacy skills, it also has its limitations. One notable limitation was that individuals were only measured once, and thus, individual patterns and development over time could not be examined. In this study, grade level was a proxy for experience and achievement. Following students over time would provide unique insights into development. Additionally, the specific vocabulary covered in the curricula is also unknown, and thus, it is unknown if the students are already familiar with the words on the spelling inventories. This information may provide a deeper understanding of why some difficult words saw greater rates of accuracy than easier words (e.g., greater accuracy spelling *riding* than *stick* or *fright* for the students in grades 2 and 3). More information about instruction may account for the achievement of the sixth-grade sample. Finally, the two results of the two spelling inventories were analyzed separately in this study, as to focus on how each age range performed on their respective spelling inventories. However, the analysis could have been conducted differently and compared the same spelling features across age groups and spelling inventories, which may have revealed patterns related to the children's spelling development and proficiency with various spelling features over a larger age range.

Future research examining how the spelling of individuals and groups develops over time would be valuable as it may indicate what patterns and phenomena are difficult to capture through examining a single spelling sample at a specific point in time. Additionally, research that compared different pedagogical practices (e.g., various combinations of vocabulary, phonics, word study, etc.) and how those practices influence spelling development would also be helpful. While examining the spelling development of Chinese students is important for what is learned about word knowledge, it would be beneficial to explore more deeply comparisons between spelling development and other components of literacy, such as word recognition, comprehension, morphological knowledge, phonological awareness, fluency, and writing. Investigating these relationships might provide unique insights into Chinese students' developmental processes in literacy and provide additional perspectives on the similarities and differences of Chinese speaking English learners and their English L1 peers. The advanced students in this study would be interesting to study as they learn beginning morphological orthographic features, something that can be learned from secondary students; see Carlisle, 2010; McBride-Chang et al., 2008.

This study suggests some important implications. At the alphabet layer, there were some specific sounds, letters, and patterns that caused difficulty for participants. Learners may benefit from specific instruction related to these difficult features. The short vowels *e* and *u* were difficult, and minimal pair exercises that help learners discriminate these vowel sounds with others may be beneficial (e.g., contrasting *pet* to *putt* or *bed* to *bad* or *mud* to *mod*). Drawing attention to the difficult short vowel sounds and juxtaposing them with other vowels may help students discriminate these sounds from others and ultimately, read and spell these sounds with greater mastery.

During the pattern layer of analysis, participants' spelling illustrated how they experimented spelling long vowel patterns (e.g., spelling *float* as FLOTE). Thus, explicit instruction on long vowel patterns may be helpful for students, particularly showing them patterns like the CVVC, CVV, CVCe patterns. Blends were particularly

difficult for younger students, which is understandable given the lack of consonant clusters in Chinese (Chang, 2001). Thus, explicit instruction in early grades that highlights pure blends vs. blends with a slight inserted vowel (e.g., *blade* - /bled/ vs. /bøled/) may draw attention to the differences in these sounds and concepts for younger students and promote mastery at a younger age. Once long vowels patterns are mastered, students are ready to examine other vowels in English. During vocabulary lessons, students will benefit from seeing the words written as part of the instruction (Rosenthal & Ehri, 2008).

Vocabulary learning expands at the meaning layer in the examination of inflected and derivational morphology (Nagy & Townsend, 2012). It has been emphasized that students' reading, including vocabulary and comprehension, will improve when accompanied by spelling instruction (Cutler et al., 2010; Foorman & Petscher, 2010; Graham & Santangelo, 2014). In an integrated and generative form of instruction, students can begin with inflected morphology and related word families in general, and domain-specific academic vocabulary (Templeton et al., 2015).

Conclusion

The orthographic development in English of Chinese L1 students illustrates interesting similarities and differences in comparison to English L1 learners. This study aligns with previous studies as it illustrates that Chinese EFL learners develop their English skills in a similar fashion to English L1 peers (Chiappe, Siegel, & Gottardo, 2002; Helman & Bear, 2007). However, differences in acquisition were also found, which are likely related to the differences in phonology between English and Chinese and students' familiarity with pinyin. This study builds upon previous research by providing specific information and analyses of Chinese students' English orthographic development, which is particularly valuable given the efforts of the Chinese education system to promote English skills for their students as well as the increasing presence of Chinese speakers in the United States and around the world. Moving forward, when researchers and educators understand students' orthographic knowledge from developmental and cross-linguistic perspectives, they are better prepared to identify which features students have mastered, and they have a guide for a sequence of instruction that includes features that might be confusing cross-linguistically.

Notes

- ¹ Internal consistency was evaluated and overall reliability was established with a coefficient of .915 (Cronbach's alpha). Analyses for reliability by item discrimination and difficulty, and internal consistency provide evidence that these instruments differentiate between relatively higher and lower performing students reliably. Test-retest data for these two inventories indicate similar reliability when English learners, special education, and gifted students are included. Reliability estimates for both inventories were acceptable (Sterbinsky, 2007).
- ² According to Cohen (1988), effect sizes (*d*) of .2, .5, and .8 are small, medium, and large, respectively.
- ³ Only significant finding.

References

- Bahr, R. H., Silliman, E. R., & Berninger, V. (2009). What spelling errors have to tell about vocabulary learning, pp. 177–210. In C. Wood & V. Connelly (Eds.), *Contemporary perspectives on reading and spelling*. London: Routledge.
- Bear, D. R., Invernizzi, M. A., Templeton, W. S., Johnston, F. (2016). *Words Their Way: Phonics, vocabulary, and spelling instruction* (6th ed.). Boston, MA: Allyn & Bacon.
- Beeson, P. M., Rising, E., Kim, E. S., & Rapcsak, S. A. (2010). A Treatment Sequence for Phonological Alexia/Agraphia. *Journal of Speech, Language, and Hearing Research*, 53(2), 450–468. doi:10.1044/1092-4388(2009/08-0229)
- Berninger, V.W., Abbott, R.D., Nagy, W., & Carlisle, J. (2010). Growth in phonological, orthographic, and morphological awareness in grades 1 to 6. *Journal of Psycholinguistic Research*, 39(2), 141-163. doi:10.1007/s10936-009-9130-6

- Bertelson, P., Chen, H-C, & de Gelder, B. (1997). Explicit speech analysis and orthographic experience in Chinese readers. In H-C, Chen (Ed.), *Cognitive process of Chinese and related Asian languages* (pp. 27–46). Hong Kong: Chinese University Press.
- Biemiller, A. (1970). The development of the use of graphic and contextual information as children learn to read. *Reading Research Quarterly*, 6(1), 75-96.
- Blachman, B. A., Schatschneider, C., Fletcher, J. M., Murray, M. S., Munger, K. A., & Vaughn, M. G. (2013, August 12). Intensive reading remediation in Grade 2 or 3: Are there effects a decade later? *Journal of Educational Psychology*. Advance online publication. doi:10.1037/a0033663
- Caravolas, M., Hulme, C., & Snowling, M. J. (2001). The foundations of spelling ability: Evidence from a 3-year longitudinal study. *Journal of Memory and Language*, 45(4), 751-774. doi:10.1006/jmla.2000.2785
- Carlisle, J. F. (2010). Effects of instruction in morphological awareness on literacy achievement: An integrative review. *Reading Research Quarterly*, 45(4), 464-487. doi:10.1598/RRQ.45.4.5
- Chang, J. (2001). Chinese speakers. In M. Swan & B. Smith (Eds.), *Learner English* (2nd ed., pp. 310-324). Cambridge, United Kingdom: Cambridge University Press.
- Cheung, H., Chen, H.C., Lai, C.Y., Wong, O.C., & Hills, M. (2001). The development of phonological awareness: Effects of spoken language experience and orthography. *Cognition*, 81(3), 227-241. doi:10.1016/S0010-0277(01)00136-6
- Cheung, H., & Ng, L. K. H. (2003). Chinese reading development in some major Chinese societies: An introduction. In C. McBride-Chang & H-C. Chen (Eds.), *Reading development in Chinese children* (pp. 3-17). Westport, CT: Greenwood Press.
- Chiappe, P., Siegel, L.S., & Gottardo, A. (2002). Reading-related skills of kindergartners from diverse linguistic backgrounds. *Applied Psycholinguistics*, 23(1), 95–116. doi:10.1017/S014271640200005X
- Chiappe, P., Siegel, L., & Wade-Wooley, L. (2002). Linguistic diversity and the development of reading skills: A longitudinal study. *Scientific Studies of Reading*, 6(4), 369–400. doi:10.1207/S1532799XSSR0604_04
- Chow, B. W-Y., McBride-Chang, C., & Burgess, S. (2005). Phonological Processing Skills and Early Reading Abilities in Hong Kong Chinese Kindergartners Learning to Read English as a Second Language. *Journal of Educational Psychology*, 97(1), 81-87. doi:10.1037/0022-0663.97.1.81
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Earlbaum Associates.
- Cook, V. J. (1997). L2 users and English spelling. *Journal of Multilingual and Multicultural Development*, 18(6), 474–488. doi:10.1080/01434639708666335
- Cronnell, B. (1985). Language influences in the English writing of third- and sixth-grade Mexican-American students. *Journal of Educational Research*, 78(3), 168–173.
- Cutler, A., Treiman, R., & van Ooijen, B. (2010). Strategic deployment of orthographic knowledge in phoneme detection. *Language and Speech*, 53(3), 307-320. doi:10.1177/0023830910371445
- Duanmu, S., 2006. Chinese (Mandarin): Phonology. In K. Brown (Ed.), *Encyclopedia of Language and Linguistics*, (2nd ed., pp. 351-355). Oxford, UK: Elsevier Publishing House
- Ehri, L. C. (2006). Alphabetic instruction helps children learn to read. In R. M. Joshi & P. G. Aaron (Eds.), *Handbook of orthography and literacy* (pp. 649–678). Mahwah, NJ: Lawrence Erlbaum.
- Ehri, L. C. (1997). Learning to read and learning to spell are one and the same, almost. In C. A. Perfetti, L. Rieben, & M. Fayol (Eds.) *Learning to spell: Research, theory, and practice across languages* (pp. 237–269). Mahwah, NJ: Erlbaum: Lawrence Erlbaum Associates.
- Ehri, L. C. (2014). Orthographic mapping in the acquisition of sight word reading, spelling memory, and vocabulary learning. *Scientific Studies of Reading*, 18(1), 5–21.
- Ehri, L., & Roberts, T. (2006). The roots of learning to read and write: Acquisition of letters and phonemic awareness In D. K. Dickinson & S. B. Neuman (Eds.), *Handbook of Early Literacy Research* (Vol. 2, pp. 113-131). New York, NY: Guilford Press.
- Ellis, N. (1997). Interactions in the development of reading and spelling: Stages, strategies, and exchange of knowledge. In C.A. Perfetti, L. Rieben & M. Fayol (Eds.), *Learning to spell: Research, theory, and practice across languages* (pp. 271 – 294). Mahwah, NJ: Lawrence Erlbaum Associates.

- Ellis, N., & Cataldo, S. (1990). The role of spelling in learning to read. *Language and Education*, 4(1), 1-28. doi:10.1080/09500789009541270
- Fashola, O. S., Drum, P. A., Mayer, R. E., & Kang, S. J. (1996). A cognitive theory of orthographic transition: Predictable errors in how Spanish-speaking children spell English words, *American Educational Research Journal*, 33(4), 825–843. doi:10.3102/00028312033004825
- Ferrolli, L., & Shanahan, T. (1993). Voicing in Spanish to English spelling knowledge transfer. *National Reading Conference Yearbook*, 42, 413–418. is <http://eric.ed.gov/?id=ED358739>
- Finegan, E. (2008). *Language: Its structure and use* (5th ed.). Boston, MA: Thomson Wadsworth.
- Foorman, B.R., & Petscher, Y. (2010). Development of spelling and differential relations to text reading in grades 3–12. *Assessment for Effective Intervention* 36(1), 7-20. doi:10.1177/1534508410379844
- Ford, K. L., Invernizzi, M. A., Huang, F. (2014). Predicting first grade reading achievement for Spanish-speaking Kindergartners: is early literacy screening in English valid? *Literacy Research and Instruction*, 53(4), 269–286. doi:10.1080/19388071.2014.931494
- Gehsmann, K., & Templeton, S. (2011/2012). Of stages and standards in literacy: Teaching developmentally in the age of accountability. *Journal of Education*, 192(1), 5–16.
- Graham, S., Harris, K.R., & Chorzempa, B.F. (2002). Contribution of spelling instruction to the spelling, writing, and reading of poor spellers. *Journal of Educational Psychology*, 94(4), 669-686. doi:10.1037/0022-0663.94.4.669
- Graham, S., & Santangelo, T. (2014). Does spelling instruction make students better spellers, readers, and writers? A meta-analytic review. *Reading and Writing*, 27, 1703–1743. doi: 10.1007/s11145-014-9517-0
- Hamada, M., & Koda, K. (2008). Influence of first language orthographic experience on second language decoding and word learning. *Language Learning*, 58(1), 1–31. doi:10.1111/j.1467-9922.2007.00433.x
- He, T-H., & Wang, W-L. (2009). Invented spelling of EFL young beginning writers and its relation with phonological awareness and grapheme-phoneme principles. *Journal of Second Language Writing*, 18(1), 44–56. doi:10.1016/j.jslw.2008.06.001
- Helman, L. (2004). Building on the sound system of Spanish: Insights from the alphabetic spellings of English language learners. *Reading Teacher*, 57, 452–460.
- Helman, L. (2005). Using literacy assessment results to improve teaching for English language learners. *The Reading Teacher*, 58(7), 668-677. Newark, DE: International Reading Association.
- Helman, L. A., & Bear, D. R. (2007). Does an established model of orthographic development hold true for English learners? In D. W. Rowe, R. Jimenez, D. L. Compton, D. K. Dickinson, Y. Kim, K. M. Leander, & V. J. Risko (Eds.), *56th Yearbook of the National Reading Conference* (pp. 266–280).
- Helman, L. A., Bear, D. R., Templeton, S., Invernizzi, M. A., & Johnston, F. (2012). *Words their way with English learners: Word Study for Phonics, Vocabulary, and Spelling Instruction* (2nd ed). Boston, MA: Allyn & Bacon.
- Helman, L. A., Delbridge, A., Parker, D., Arnal, M., & Jara Mödinger, L. (2015). Measuring Spanish orthographic development in private, public, and subsidised schools in Chile. *Assessment in Education: Principles, Policy & Practice*, 1-26. <http://dx.doi.org/10.1080/0969594X.2015.1038217>[dx.doi.org]
- Henderson, E.H. (1974). Correct spelling—An inquiry. *The Reading Teacher*, 28(2), 176–179.
- Henderson, E. H. (1981). *Learning to read and spell: The child's knowledge of words*. DeKalb, IL: Northern Illinois Press.
- Henderson, E.H. (1990). *Teaching spelling* (2nd ed.). Boston: Houghton Mifflin.
- Henderson, E. H. (1992). The interface of lexical competence and knowledge of written words. In S. Templeton & D. Bear (Eds.), *Development of orthographic knowledge and the foundations of literacy: A memorial Festschrift for Edmund H. Henderson*. (pp. 1-30). Hillsdale, NJ: Lawrence Erlbaum.
- Henderson, E. H., & Beers, J. (Eds.) (1980). *Developmental and cognitive aspects of learning to spell: A reflection of word knowledge*. Newark, DE: International Reading Association
- Henderson, E. H. & Beers, J. W. (1980). *Developmental and cognitive aspects of learning to spell*, (Eds.). Newark, DE: International Reading Association.
- Henderson, E.H., & Templeton, S. (1986). The development of spelling ability through alphabetic, pattern, and meaning. *Elementary School Journal*, 86(3), 304–316.

- Hoeffel, E., Rastogi, S., Ouk Kim, M., & Shahid, H. The United States Census Bureau, (2012). *The Asian population: 2010-2010 census briefs*. Retrieved from website:
<https://www.census.gov/prod/cen2010/briefs/c2010br-11.pdf>
- Institute of International Education. (2010, November 15). *Open doors 2010 fast facts*. Retrieved from http://www.iie.org/Research-and-Publications/Open-Doors/Data/~/_media/Files/Corporate/Open-Doors/Fast-Facts/Fast Facts 2010.ashx
- Institute of International Education. (2013, November 11). *Open doors 2013 fast facts*. Retrieved from http://www.iie.org/~/_media/Files/Corporate/Open-Doors/Fast-Facts/Fast-Facts-2013.ashx
- Invernizzi, M., Abouzeid, M., & Gill, J. T. (1994). Using students' invented spellings as a guide for spelling instruction that emphasizes word study. *The Elementary School Journal*, 95(2), 155-167.
- Invernizzi, M., & Hayes, L. (2004). Developmental-spelling research: A systematic imperative. *Reading Research Quarterly*, 39(2), 216-228. doi:10.1598/RRQ.39.2.4
- Lee, S. H., & Scanlon, D. M. (2015). The effects of the Interactive Strategies Approach on at-risk Kindergarten spellers. *Reading and Writing*, 28, 313-346. doi:10.1007/s11445-014-9526-z
- Leong, C. K., Hau, K.T., Cheng, P.W., & Tan, L.H. (2005). Exploring two-wave reciprocal structural relations among orthographic knowledge, phonological sensitivity, and reading and spelling of English words by Chinese students. *Journal of Educational Psychology*, 97(4), 591-600. doi:10.1037/0022-0663.97.4.591
- Leong, C. K., Tan, L. H., Cheng, P. W., & Hau, K. T. (2005). Learning to read and spell English words by Chinese students. *Scientific Studies of Reading*, 9(1), 63-84. doi:10.1207/s1532799xssr0901_5
- Li, H., & Rao, N. (2000). Parental influences on Chinese literacy development: A comparison of preschoolers in Beijing, Hong Kong, and Singapore. *International Journal of Behavioral Development*, 24(1), 82-90. doi:10.1080/016502500383502
- Lin, D., McBride-Chang, C., Shu, H., Zhang, Y., Li, H., Zhang, J., Aram, D., & Levin, I. (2010). Small Wins Big: Analytic Pinyin Skills Promote Chinese Word Reading. *Psychological Science*, 21(8), 1117-1122. doi:10.1177/0956797610375447
- Liow, S.J.R., & Lau, L.H.S. (2006). The development of bilingual children's spelling in English. *Journal of Educational Psychology*, 98(4), 868-878. doi:10.1037/0022-0663.98.4.868
- McBride-Chang, C., Liu, P.D., Wong, T., Wong, A., & Shu, H., (2012). Specific reading difficulties in Chinese, English, or both: Longitudinal markers of phonological awareness, morphological awareness, and RAN in Hong Kong Chinese children, *Journal of Learning Disabilities*, 45(6) 503-514. doi:10.1177/0022219411400748
- McBride-Chang, C., Tardif, T., Cho, J.R., Shu, H., Fletcher, P., Stokes, S.F., & Leung, K. (2008). What's in a word? Morphological awareness and vocabulary knowledge in three languages. *Applied Psycholinguistics*, 29(3), 437-462. doi:10.1017/S014271640808020X
- McDowell, H.J., & Lorch, M.P. (2008). Phonemic awareness in Chinese L1 readers of English: Not simply an effect of orthography. *TESOL Quarterly*, 42(3), 495-513. doi:10.1002/j.1545-7249.2008.tb00143.x
- Mehta, P., Foorman, B.R., Branum-Martin, L., & Taylor, W.P. (2005). Literacy as a unidimensional multilevel construct: Validation, sources of influence, and implications in a longitudinal study in grades 1 to 4. *Scientific Studies of Reading*, 9(2), 85-116. doi:10.1207/s1532799xssr0902_1
- Morris, L. (2001). Going through a bad spell: What the spelling errors of young ESL learners reveal about their grammatical knowledge. *Canadian Modern Language Review*, 58(2), 273-286. doi:10.3138/cmlr.58.2.273
- Nagy, W., & Townsend, D. (2012). Words as tools: Learning academic vocabulary as language acquisition. *Reading Research Quarterly*, 47(1), 91-108.
- Nassaji, H. (2007). The development of spelling and orthographic knowledge in English as an L2: A longitudinal case study. *Canadian Journal of Applied Linguistics*, 10(1), 77-98.
- Perfetti, C.A. (1997). The psycholinguistics of spelling and reading. In C. A. Perfetti, L. Rieben, & M. Fayol (Eds.), *Learning to spell: Research, theory, and practice across languages* (pp. 21-38). Mahwah, NJ: Lawrence Erlbaum Associates.
- Perfetti, C.A. (2007). Reading ability: Lexical quality to comprehension. *Scientific Studies of Reading*, 11(4), 357-383. doi:10.1080/10888430701530730

- Rapp, B. & Lipka, K. (2010). The literate brain: The relationship between spelling and reading. *Journal of Cognitive Neuroscience* 23(5), 1180–1197. doi:10.1162/jocn.2010.21507
- Rau, D.V., Chang, H-H. A., & Tarone, E.E. (2009). Think or sink: Chinese learners' acquisition of the English voiceless interdental fricative. *Language Learning*, 59(3), 581–621. doi:10.1111/j.1467-9922.2009.00518.x
- Read, C. (1971). Pre-school children's knowledge of English phonology. *Harvard Educational Review*, 41(1), 1–34.
- Read, C. (1975). *Children's categorizations of speech sounds in English* (Research Report No. 17). Urbana, IL: National Council of Teachers of English.
- Read, C. (1986). *Children's creative spelling*. Boston, MA: Routledge.
- Rosenthal, J., & Ehri, L. (2008). The mnemonic value of orthography for vocabulary learning. *Journal of Educational Psychology*, 100(1), 175-191. doi:10.1037/0022-0663.100.1.175
- Share, D. (1995). Phonological recoding and self-teaching: sine qua non of reading acquisition. *Cognition*, 55(2), 151–218. doi:10.1016/0010-0277(94)00645-2
- Shen, H., & Bear, D. R. (2000). The development of orthographic skills in Chinese children. *Reading and Writing: An Interdisciplinary Journal*, 13, 197–236.
- Stanovich, K.E. (2000). *Progress in Understanding Reading: Scientific Foundations and New Frontiers*. New York, NY: Guilford.
- Sterbinsky, A. (2007). Words Their Way Spelling Inventories: Reliability and validity analyses. Memphis, TN: Center for Research in Educational Policy, University of Memphis. Retrieved from http://rti.pearsoned.com/docs/RTIsite/Reliability_and_Validation_Study_Report.pdf
- Taft, M. (2003). Morphological representation as a correlation between form and meaning. In G. H. Assink & D. Sandra (Eds.), *Reading complex words: Cross language studies* (pp. 113-137). New York: Kluwer Academic.
- Templeton, S. (2003). Spelling. In J. Flood, D. Lapp, J. R. Squire, & J. M. Jensen (Eds.), *Handbook of research on teaching and the English language arts* (2nd ed., pp. 738-751). Mahwah, NJ: Lawrence Erlbaum Associates.
- Templeton, S., & Bear, D. (Eds.). (1992). *Development of orthographic knowledge and the foundations of literacy: A memorial Festschrift for Edmund H. Henderson*. Hillsdale, NJ: Lawrence Erlbaum.
- Templeton, S., & Bear, D. (2018). Word study research. In Lapp, D. and Fisher, D. (Eds.), *Handbook of Research in the Teaching of English* (4th ed., pp. 207-232). NY: Routledge.
- Templeton, S., & Morris, D. (2000). Spelling. In M. Kamil, P. Mosenthal, P.D. Pearson, & R. Barr (Eds.), *Handbook of Reading Research: Vol. 3* (pp. 525-543). Mahwah, NJ: Lawrence Erlbaum Associates.
- Thompson, I. (2001). Japanese speakers. In M. Swan & B. Smith (Eds.), *Learner English* (2nd ed., pp. 296-309). Cambridge, United Kingdom: Cambridge University Press.
- Treiman, R., & Cassar, M. (1997). Spelling acquisition in English. In C.A. Perfetti, L. Rieben, & M. Fayol (Eds.), *Learning to spell: Research theory, and practice across languages* (pp. 61–80). Mahwah, NJ: Lawrence Erlbaum Associates.
- Walter, C. (2001). French speakers. In M. Swan & B. Smith (Eds.), *Learner English* (2nd ed., pp. 52-72). Cambridge, United Kingdom: Cambridge University Press.
- Wang, M. & Geva, E. (2003a). Spelling acquisition of novel English phonemes in Chinese children. *Reading and Writing: An interdisciplinary Journal*, 16(4), 325–348. doi:10.1023/A:1023661927929
- Wang, M. & Geva, E. (2003b). Spelling performance of Chinese children using English as a second language: Lexical and visual-orthographic processes. *Applied Psycholinguistics*, 24(1), 1–25. doi:10.1017/S0142716403000018
- Wang, M., Koda, K., & Perfetti, C.A. (2003). Alphabetic and nonalphabetic L1 effects in English word identification: A comparison of Korean and Chinese English L2 learners. *Cognition*, 87(2), 129–149. doi:10.1016/s0010-0277(02)00232-9
- Wang, M., Perfetti, C.A., & Liu, Y. (2005). Chinese-English biliteracy acquisition: Cross-language and writing system transfer. *Cognition*, 97(1), 67–88. doi:10.1016/j.cognition.2004.10.001
- Wolf, M. (2007). *Proust and the squid: the story and science of the reading brain*. New York, New York: Harper.

- Wu, C-Y., Ho, M-H. R., Chen, & S-H. A. (2012). Congruent with alphabetic language networks, but subnetworks that are language specific: A meta-analysis of fMRI studies on Chinese orthographic, phonological, and semantic processing. *NeuroImage* 63(1), 381-391. doi:10.1016/j.neuroimage.2012.06.047
- Yeong, S. H. M., Fletcher, J., Bayliss, D. M. (2014). Importance of phonological and orthographic skills for English reading and spelling: A comparison of English monolingual and Mandarin–English bilingual children. *Journal of Educational Psychology*, 106(4), 1107–1121. doi.org/10.1037/a0036927
- Yeung, P., Ho, C.S., Chan, D.S., & Chung, K.K. (2013). Modeling the relationships between cognitive-linguistic skills and writing in Chinese among elementary grades students. *Reading and Writing*, 26(7), 1195-1221. doi:10.1007/s11145-012-9411-6.

About the Authors

Donald Bear is Professor Emeritus in Literacy Studies at Iowa State, and University of Nevada, Reno where he directed literacy centers and taught at all levels. His research explores the synchrony of literacy development from beginning concept of word in prekindergarten to how morphology underlies academic vocabulary learning.

Sam von Gillern is a Clinical Assistant Professor of Literacy at Texas A&M University. He enjoys facilitating experiences for preservice teachers to connect with the community to promote learning opportunities for college students and children alike. His research interests include second language acquisition, game-based learning, and digital literacies.

Wei Xu teaches Chinese, Japanese, and Advanced Asian Studies at The Davidson Academy of Nevada. Her PhD degree is in Curriculum and Instruction with a focus on applied linguistics. Her research interests include pragmatics and literacy development in 2nd language acquisition.

Exploring Young English Learners' Perceptions of the Picture Word Inductive Model in China

Xuan Jiang*

St. Thomas University, Miami, USA

Abstract

The Picture-Word Inductive Model (PWIM) is one of the new teaching methods introduced in English Language Education (ELT) in China, in order to develop English learners' communicative competence. However, studies conducted on PWIM from the perspectives of English learners are underrepresented and no research has been done about using PWIM with English learners in China. Considering this gap, I attempted to answer the question: What are young English learners' perceptions of PWIM? I deployed observations, questionnaires, and interviews to explore what fourth and seventh graders perceived to be the strengths and drawbacks of PWIM. My findings suggest that many students listed and exemplified numerous strengths of PWIM, and a few mentioned its weaknesses. The students expressed that PWIM had a positive effect on their English learning.

Keywords: picture-word inductive model, perceptions, vocabulary, mixed methods, exploratory

Introduction

English has been taught as a core and compulsory subject in China for decades, making China the largest population of English learners and users: approximately 440 million English-learning and English-using people in China (Crystal, 2008, pp. 4-5). And for decades, English teaching in China promoted a strong linguistic focus on grammar, reading, and translation, with a method called "teacher-centered textbook-analysis-based grammar-translation" (Yang, 2000, p. 19). This traditional approach, however, is not the only English teaching approach implemented nowadays, because English learners in China need more skills (e.g., writing and speaking) and communicative competence. Communicative competence is defined as "a certain level of language proficiency, speech and social-cultural set of knowledge, skills and abilities that enable to vary acceptably and appropriately their communicative behavior in a communicative way" (Fahrutdinova, Yarmakeev, & Fakhrutdinov, 2014, p. 36). Picture-Word Inductive Model (PWIM), as a new English Language Teaching (ELT) approach that seemingly had never been used in China before, might be essential to the ongoing reformed English education, because PWIM potentially enables learners to manage the meaning and use of new words, empowers learners from passive to active learning by speaking and writing, and helps learners write paragraphs step by step from adding up words, phrases and sentences.

In terms of research studies, PWIM is not well represented. Until now, there have been only a few research studies of PWIM (Calhoun, 1999; Feng, 2011; Swartzendruber, 2007; Wong, 2009), and none of them have been conducted with any English as a Foreign Language (EFL) learners in Mainland China. With the empirical and contextual gap mentioned above, the purpose of this study was to explore young EFL learners' perceptions of PWIM in China. The implication of this study included using this research piece as a mechanism to develop a

* Email: xjiang@stu.edu. Tel.: 1-786-838-9482. Address: School of Arts and Education, Saint Thomas University, 16401 NW 37 Avenue, Miami Gardens, FL 33054, USA.

view about PWIM as a newly-tried ELT method from learners' perspectives within a policy-driven framework, so that educational leaders and practitioners can consider PWIM as a new teaching approach being introduced and implemented widely in Mainland China. In addition, applying the findings and discussions of this paper to different contexts, English teachers may have better approaches to developing appropriate and effective ways of presenting and practicing PWIM with students, in order to help students develop their language skills and communicative competence.

Literature Review

English Teaching in China

English teaching in China has its peculiar cultural, social and political context; moreover, China has experienced its own history of English textbooks and syllabi changes. Historically, English has been taught as a compulsory subject with a strong emphasis on grammar, translation, and reading, mainly through direct instruction approaches (Yang, 2000). Such traditional teaching approaches are characterized by systematic study of grammar, extensive use of Chinese-English translation, and persistent memorization of syntactic patterns and vocabulary (Hu, 2002). The traditional approach, however, has failed to develop an adequate level of EFL learners' communicative competence (i.e. the ability to use English for authentic verbal and textual communications) in China. Millions of EFL learners taught by this traditional approach can read, but cannot speak or write well. As a result, new English teaching approaches including Communicative Language Teaching (Yang, 2014) and Task-based Learning and Teaching (Zheng & Borg, 2014), focusing on transforming learners from passive to active agents, have recently entered EFL classrooms in China.

As mentioned earlier, some innovative teaching approaches have recently been introduced into English education in China. However, school principals and teachers may feel uncomfortable and unsafe when they are informed of change. This feeling occurs when "change...upsets the pattern" people are accustomed to, thereby thrusting people into new perceptions and challenging people's way of dealing with life (Evans, 1996, p. 27). Such feelings are more likely to register with those educational practitioners who are already satisfied with the situation they are in (Feng, 2011).

Students may similarly be accustomed to the traditional teaching methods and resistant to change. Their resistance to change may also be rooted in Chinese culture which has been largely influenced by Confucius (Jiang, 2011), and Confucian philosophy has led to the teacher-centred approach. It has also, as Scollon (1999) pointed out, correspondingly exerted an ideological impact on students' ways of thinking, teachers' and students' roles, objectives of education, and behaviours of teachers and students. Under the influence of the Confucian education philosophy, what teachers say is authoritative and true, and these truths should be repeated and remembered instead of being questioned by students. Values including "docility, passivity, and conformity" are what education demands of its students in this Confucian heritage culture (Yen, 1987, p. 52). Students under the influence of Confucian heritage culture for years are more likely to be passive thinkers without any consciousness of inductive reasoning (Scollon, 1999). Scollon (1999) undertook a case study to identify the cultural constructs that underlie the participants' viewpoints on the purposes of education in Chinese classrooms and that influence the behaviors of students and teachers. She observed that the cultural notions stemming from Confucius affect learning processes tacitly (Scollon, 1999). Chinese EFL learners influenced by historically dominant Confucian traditions tend to be disabled from using rhetorical reasoning (Jiang, 2011). Nevertheless, it deserves exploring and experimenting whether and how these learners can be transformed with rhetorical reasoning after certain instructional practices, such as PWIM.

Picture-Word Inductive Method

PWIM is a teaching approach based on Calhoun's (1999) research on early literacy. The fundamental tenet of PWIM is its use of "pictures as a stimulus for language experience activities" in classrooms to teach young beginners learning to read and write (Joyce, Weil, & Calhoun, 2009, p. 130). Calhoun suggested that PWIM be

used to teach several skills as an integrated whole simultaneously, beginning with the phonetic and structural components of language, for example, the pronunciation and the spelling of a word. Students using PWIM see the item, listen to the teacher's pronunciation, and then pronounce the word to reinforce word recognition. PWIM also covers explicit instruction and induction, as well as an immediate assessment of students' needs and comprehension. Calhoun suggested that a teacher can arrange any individual activity, as well as small-group and large-group activities, with a PWIM format. She also suggested that PWIM can be modified and applied to older beginners (Calhoun, 1999).

PWIM has many successful examples of applications in classrooms (Calhoun, 1999; Joyce & Showers, 2002; Joyce et al., 2009). Calhoun claims that its successful applications are closely related to prescribed and detailed steps for implementation, an instructional sequence incorporating cycling and recycling through the following 10 steps (Calhoun, 1999):

1. Select a picture.
2. Ask students to identify what they see in the picture.
3. Label the picture parts identified. (Draw a line from the identified object or area, say the word, write the word; ask students to spell the word aloud and then to pronounce it.)
4. Read and review the picture word chart aloud.
5. Ask students to read the words (using the lines on the chart if necessary) and to classify the words into a variety of groups. Identify common concepts (e.g., beginning consonants, rhyming words) to emphasize with the whole class.
6. Read and review the picture word chart (say the word, spell it, say it again).
7. Add words, if desired, to the picture word chart and to the word banks.
8. Lead students into creating a title for the picture word chart. Ask students to think about the information on the chart and what they want to say about it.
9. Ask students to generate a sentence, sentences, or a paragraph about the picture word chart. Ask students to classify sentences; model putting the sentences into a good paragraph.
10. Read and review the sentences and paragraphs (p. 23).

PWIM embraces the development of visual perceptions, which is vital to children's literacy acquisition (Astorga, 1999; Clay, 2001; Joyce, Calhoun, & Hopkins, 2002). During instruction with the PWIM, students are shown a picture and are asked to identify items in the picture or "shake out" the words of the picture. The picture as a visual image has a significant role in PWIM to develop children's literacy. Joyce and Weil (2004) showed how pictures may serve as important stimuli for connecting learners' life experiences to their language learning in the classroom. The teacher writes each identified word on chart paper outside the picture and draws a line from the word to the item in the picture, essentially creating a picture-word dictionary which the students can employ to connect words with corresponding pictures. Joyce and Weil (2004) claim that such a kind of connection contributes to developing new ideas and learning vocabulary.

Wong (2009) demonstrated that vocabulary learning is also achieved through connections via dual channels of speaking and writing. 'Shaking out' the words and spelling the words, she says, brings "awareness to the letters in the word and connection between the spoken words onto written text" (Wong, 2009, p. 9). Reading out, spelling out, and writing down these words on the picture-word dictionary is the first step for children to know the phonetic and morphological form of the words. After they become more familiar with this mode of instruction, they start to write phrases, short sentences and later long ones about the pictures independently or with the assistance of teachers and more competent peers. This writing process involves learning form, meaning and use of those words. Gradually, these vocabularies should be stored, as Calhoun (1999) suggests, in students' long-term memory and eventually become a part of their prior knowledge, which will be used to learn new words.

Inquiry-oriented induction is another key feature of PWIM. Inductive thinking or induction promotes learners' awareness of language development (Astorga, 1999; Clay, 2001; Joyce et al., 2002). Induction in PWIM refers to developing new ideas through building parallelism among unrelated information, ideas, and artifacts (Joyce, Weil, & Calhoun, 2004). This type of thinking assists students noticing and inferring with patterns and relationships within the language—elements which should enable them to apply and transfer such learning to novel words. Joyce, et. al (2004) also suggest that such induction empowers students to generalize language rules, for example, how to structure sentences as in building up the sentence. Students may draw generalizations after numerous PWIM activities, and then cycling and recycling the sequences mentioned above.

There is only one quantitative research study (Swartzendruber, 2007) found in the literature review. Swartzendruber (2007) conducted a quasi-experiment study with 35 second graders in a Midwestern U.S. city with both English as a Second Language (ESL) learners and native English speakers (NESs). Results indicated that the experimental group was statistically significant better than the control group on the final assessment. Scaffolding and explicit connections to concepts and words appeared to be beneficial to both ESLs and NESs, with respect to vocabulary acquisition. The only qualitative research was Feng's (2011), in which the participant teachers reported that their fourth to sixth grade students had increased English vocabulary through PWIM and cooperative learning.

PWIM has more space to develop in terms of research studies. First, there are only a few research articles and books about PWIM studies since Calhoun's (1999) research, most of which are written or co-written by Calhoun or Joyce. Evidently, Calhoun and Joyce are the two major proponents of this model, thus PWIM narrated and evaluated by them cannot avoid possible biases.

Second, PWIM is not the only intervention in the learners' language development in those research studies, so it cannot be inferred from the results that PWIM alone contributes to the learners' language development. For example, parents also seem to play a role in the scenario, who “need to read at least five books each week to their children” in English as their native language (Joyce et al., 2009, p. 137). Third, research studies of PWIM are even fewer when the studies are narrowed down to ESL or EFL learners. There is only one quantitative research study (Swartzendruber, 2007) found in the literature review. The only qualitative research found was in Feng's (2011) dissertation. Thus, one might infer that ESL and EFL research studies of PWIM are underrepresented.

Research Question

The focus of this study was to explore young EFL learners' perceptions of PWIM in China EFL classrooms. The primary research question was: What are young English learners' perceptions of PWIM's strengths and weaknesses?

Methodology

To explore the learners' perceptions, I observed, conducted open-ended questionnaires and interviewed 18 fourth graders in an elementary school and 30 seventh graders in a secondary school to examine how they understood and perceived PWIM. I designed interview questions based on the observation and questionnaire answers. The interview data were the main source to answer the primary and secondary research questions.

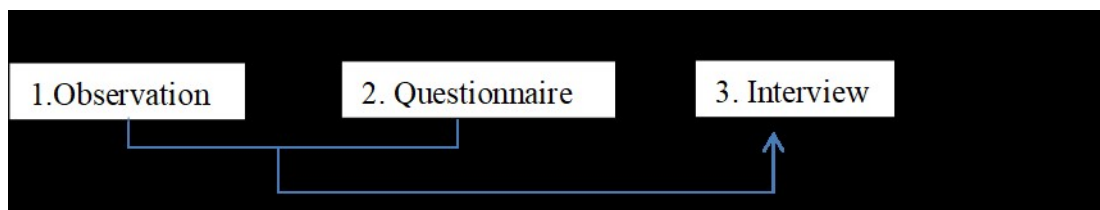


Figure 1. Procedures of data collection

Participants

In the current study, I chose participants who were early elementary pupils at the stage of learning to read and early adolescents in middle schools who already knew how to read and read to learn in two public schools in a coastal city in Mainland China. They were all Chinese. All participants had already received a year or four years of English education. The fourth graders received four 40-minute English classes per week. The seventh graders received five 45-minute English lessons every week.

After my in-class observation during the PWIM trial period for seven weeks, I used nested sampling. I asked the seventh-grade English teacher to hand out a letter to parents about my research, one parental consent form and one child consent form in both Chinese and English to each student, so the students could send them home for parents' approval. The students submitted the signed forms to their English teacher and then I had 30 participants. For the fourth graders, I agreed with the class headmaster's suggestion by introducing my research study in a parents' meeting. I handed out the parent's letter about my research, one parental consent form and one child consent form in Chinese, introduced myself and my research project and I then answered the parents' questions. They had questions with respect to the reward, the timing of interviews with their children, and what I would ask in the interviews. I answered each query one by one. In return, I asked the parents to discuss the project with their kids after the meeting and to hand in the two consent forms with their signatures if their children indeed wanted to participate with the questionnaire and the interview. Later, I collected signed forms from the fourth graders and ultimately had 18 fourth graders for the questionnaires and interviews.

Procedures

Observation

I observed the two PWIM-trial groups during most of their practice and testing sessions, for approximately 25 hours. I took field notes when I observed each of the two classes during the 7-week-long trial. I did not electronically record anything during the observations. I noted down some moments of students' facial expressions and body language, as well as their interaction with peers and teachers, in case I could ask them in the subsequent interviews. In addition to the reason stated above, observations also provided more relevant and personalized questions for the following interviews. For instance, I asked a student's feeling about PWIM after I saw he raised his hands actively during PWIM lessons.

Questionnaires

After I received parents' and children's permission, I handed out the pre-designed hardcopy PWIM questionnaires to the 18 fourth graders and 30 seventh graders, using language that is understandable to them. I gave each of them a coded ID to be put on their answer sheets. Students wrote down the coded ID that I gave, instead of their names.

Interviews

Immediately after collecting and preliminarily analyzing their answers on the questionnaires, I interviewed the 48 participants in Chinese one-to-one and digitally recorded them. The questions asked were based on what I had observed and their answers on their questionnaires. I could only interview them for a maximum of 10 minutes during class breaks, self-learning classes, and minor classes (i.e., geography, history, music, fine arts, physical exercise, and other classes apart from Chinese, Math or English). Seventh graders were pulled out one by one for maximum 10 minutes each. Fourth graders were pulled out in small clusters of three to five, in an attempt to keep the classrooms as uninterrupted as possible. The observations and questionnaire responses helped me with the interview questions. Due to the design of my research questions, the interviews was the main data source.

Data Analysis

I did data analysis during and after data collection to explore the participants' perceptions of PWIM. According to Merriam (2002, p. 14), qualitatively, "data analysis is simultaneous with data collection." I started preliminary

data analysis after collecting data from the observations, so that I could decide to keep or modify my questions asked on questionnaires. After collecting the data from the questionnaires, I analyzed the data from their answers for generating interview questions. As Stake (2005) stated, analysis refers to giving meaning to first impressions of those texts. I wrote down my first impressions on the margin of the questionnaires. I did preliminary analysis by studying and coding answers on the questionnaires, which gave me a clearer direction about what to ask for the future interviews.

Due to the tight schedule of interviews with both fourth and seventh graders, I did not analyze collected interview data deeply and thoroughly during data collection. However, I framed “new questions” for the following interviews as a result of what had been found (Seidman, 2006, p. 113). By adding new questions to the original ones and modifying some old questions for the next interview, I “made adjustments along the way” (Merriam, 2002, p. 14). It should be noted that I avoided “in-depth analysis of the interview data” until I finished all the interviews, because I tried not to “impose meaning from one participant’s interview on the next” (Seidman, 2006, p. 113). Thereafter I minimized “imposing on the generative” and inductive process of the interviews (Seidman, 2006, p. 113).

After I interviewed all assented participants, I transcribed those interviews via listening to the recording and typing every single word. Then I translated their answers from Chinese into English. Following this transcription, I conducted the coding by writing my first thoughts on the margins of the transcript to read through literal words. After that, I used different color highlights to categorize the coding. The coding categories were words and phrases which represented the regularities, patterns and topics my data covered (Bogdan & Biklen, 2007). I categorized coding to find concepts and patterns and developed them into themes through thematic analysis to help answer the three secondary research questions. Themes are “the relationship between two or more concepts” (Rubin & Rubin, 2005, p. 194). The relationship includes “why something happened, what something means, or how the interviewee feels about the matter” (Rubin & Rubin, 2005, p. 194). For instance, I develop one theme as “linguistic acquisition” by grouping seven concepts together, including “to summarize the words by unit”, “to memorize those words deeply,” “to build up paragraphs from words,” “connecting pictures and words,” “emphasizing how to use these words,” “acquiring more words” and “good for writing.” By building up themes, I threaded my draft with themes, categorized answers and evidence from those analyzed data.

Results

The findings start with developed common themes from interviews seventh and fourth graders, covering the topics of pictures, connections of pictures and words, affective enjoyments, and practical issues of implementing PWIM in a class of over 55 students. Following this is the topic of fourth graders’ reliance on their Chinese language to learn English.

Associable Pictures

PWIM uses “pictures containing familiar objects and actions” (Calhoun, 1999, p. 21), which was counted as one of its strengths, revealed from the interview answers. One seventh grade girl (pseudonym Zhou) preferred PWIM pictures to one-to-one word-picture cards because PWIM pictures could prompt more words and connections.

It was interesting that when asked to choose pictures of local culture and western culture, the seventh graders did not mind pictures of western cultures. One student (pseudonym Huang) thought of English as a “carrier,” as a vehicle to carry foreign cultures, so that she could learn more about English-speaking cultures, knowledge, and the world from pictures of English-speaking countries. As revealed from the interviews, there was little resistance from the seventh graders to pictures of foreign culture, and vocabulary and the PWIM pictures may be not necessarily culturally mapped for those EFL participants.

Fourth graders had similar opinions. The characters and stories in the pictures were suitable to their age, so students could easily associate with them. One interviewee (pseudonym Yuan) compared pictures in her textbooks with pictures in PWIM, saying that “Pictures in textbook are a little interesting, and boring with the

same style; the new method [PWIM] has extra-curriculum pictures which are greatly interesting and totally fresh to me.” Comparing PWIM pictures and word cards, another fourth-grade interviewee (pseudonym Tang) stated that many words can be found in one PWIM picture versus only one word found in word cards, and a paragraph or story can be developed from one PWIM picture versus one word found in word cards. Seen from both, PWIM pictures are more informative, connective with words and inductive as a source.

A PWIM picture has a theme with multiple backgrounds and characters to explore and create a scenario; a word card has one item in the picture and a word along with it. As shown in Figure 2, compared to conventional pictures, PWIM pictures are more informative and connective with words and inductive as a source.



Figure 2. Samples of a PWIM picture and a word card

Connections between Pictures and Words

With PWIM pictures, most of the interviewees mentioned that they learned, memorized and retrieved those previously learned and newly learned words more quickly, easily, conveniently and efficiently, “not one by one, not by rote” (interviewee pseudonym Xu), possibly due to “connections and imaginations” (interviewee pseudonym Zhu). They could not tell which jumped out of their mind first, pictures or words. Most respondents thought that pictures and words appeared simultaneously. Interviewees also thought that with pictures, the impression of word was much deeper, because words to be considered were associated with the relevant pictures.

Enjoyment

All of the interviewees liked and welcomed PWIM very much. Most of them expected their English teacher to continue using PWIM in the future to consolidate their vocabulary. One student in seventh grade (pseudonym Xu) said, “I was learning in a happy and easy atmosphere, which reduced my psychological burden and pressure,” so he anticipated the next opportunity of his English class with PWIM and would listen to the English teacher attentively. Nevertheless, there were a few students who were indifferent to whether they would have PWIM in the future or not. Some students (e.g., pseudonym Li) thought they already had the strategy or technique of having an image when meeting new vocabulary or writing topics, so there was no need for teachers to implement PWIM any longer.

In sum, PWIM, indeed for numerous interviewees, made the class alive, interesting, engaging and motivating. Students wanted the teacher to write down what they contributed during the word brainstorming and category stages. Students also expected to be picked to read what they had written at the end of each PWIM cycle, that is, the writing stage. PWIM was also welcoming and friendly to those students of unlikely risk-takers or/and with small vocabulary (revealed from the interviews with pseudonym Zhao, Xu, & Gao), because they had a lot of easy words to contribute (e.g., breakfast and lunch for Unit 4).

Practical Issues

Apart from all the merits of PWIM mentioned above, a few students talked about practical issues of PWIM. One student (pseudonym Zhou) thought of long physical distance for the near-sighted peers sitting far away from PWIM pictures. In such a large room using PWIM, those neglected groups were more likely to be distracted and

even left behind. However, the other interviewees showed disagreement of attributing distraction to PWIM, saying much as one interviewee (pseudonym Wu) that “being distracted is an individual phenomenon, not a method weakness.” A few interviewees (e.g., pseudonym Guo) listed noise as one of PWIM’s weaknesses. However, some interviewees showed disagreement in terms of noise being a disadvantage. “Being noisy is not the weakness,” said one interviewee (pseudonym Huang), “It’s students’ personality weakness.”

For the three PWIM cycles, there were many repeated words (words of body parts, color, background—*sky*, *cloud*, etc.) brainstormed. A couple of the interviewees saw such repetition as a weakness and a redundancy (causing cognitive overload). One interviewee (pseudonym Zhang) said that too many repetitious words kept her slightly from thinking of more words from PWIM pictures, and from easily finding newly learned words in the word bank on her category and writing module papers. She also mentioned that if her English teacher could break the table of word bank into two parts, one with high frequency words and one with newly met words, then such a problem would be solved.

Importance of the Chinese Language to the English Language Learning

Fourth graders’ native language played a vital role in their English language learning. They had not developed phonetic awareness yet, nor any linguistic, cognitive or metacognitive strategies in learning English from my observation. That is to say, the Chinese language was the only existing previous knowledge they could rely on in terms of vocabulary learning. Similar to other interviewees, one student (pseudonym Xie) could think of unknown words in the Chinese language first in the brainstorming stage of PWIM and then memorize the English version told by other students and their fourth-grade English teacher.

Participants’ Perceptions of PWIM as to its Strengths and Weaknesses in Summary

In reference to the question asked of the seventh graders about the strengths of PWIM, there were 15 reported concepts from the questionnaire and interview responses: interesting, to study more quickly, to memorize easily, connecting pictures and words, emphasizing how to use these words, acquiring more words, knowing more knowledge, to learn happily, to make the atmosphere lively, to summarize the words by unit, to memorize those words deeply, to build up paragraphs from words, motivating, directly visible, good for writing, and peer learning. The 15 concepts were further categorized into five themes: linguistic acquisition, knowledge acquisition, cognitive development/advancement, emotional/psychological enjoyment, and peer learning.

Similar to the seventh graders, the fourth-graders spoke highly of PWIM because of the interesting pictures, strong connections of pictures and words as an efficient way to manage new words, linkage to their tests, affective enjoyments and strategic development. When it comes to weaknesses of PWIM, a few of the seventh graders wrote in their questionnaires: noisy classroom without order, hard to hear clearly, think attentively or write down all of the words, more time for teachers to prepare for the class, time consuming as to the reading, spelling and writing down of the upcoming words, extra work for teachers, and unorganized handwriting on the PWIM sheets. Figure 3 shows the six concepts and three themes in three different colors.

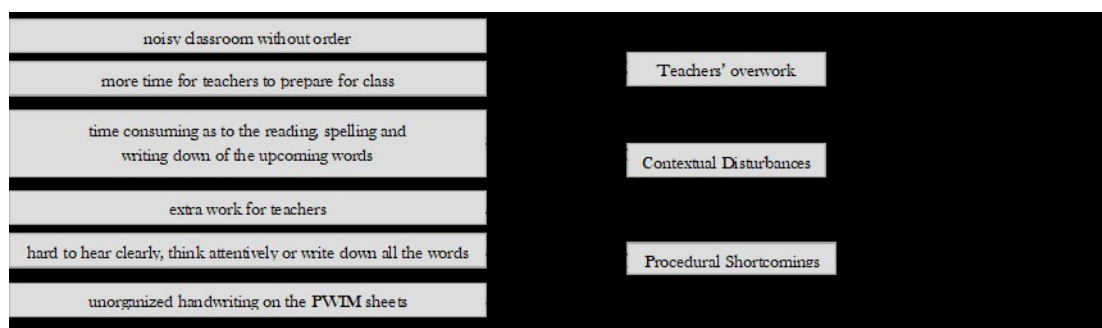


Figure 3. Weaknesses of PWIM reported by the seventh graders

However, during the interview when asked about these weaknesses, student participants stated they would prefer to keep using PWIM by avoiding or improving these weaknesses, as opposed to not using PWIM at all. The fourth graders gave a shorter list of PWIM weaknesses: noisy classroom, too many new words at one time, too many writing examples to think of their own, and so on. The respondents interviewed confessed that they experienced some difficulty in learning, but they felt they would overcome it. They thought that they can process the information and manage them subsequently, which is seen as normal process learning to them.

All the interviewees greatly welcomed PWIM and would use PWIM as English learners and English teachers if possible. Some of them mentioned that they had a habit of connecting pictures and words automatically after experiencing PWIM. Some seventh graders even developed a mind-set to seek a picture to consolidate their unit vocabulary autonomously, but they would love to receive some kind of scaffolding or guidance of using pictures from their English teacher.

Discussion

Similar to the fourth to sixth graders in Feng's (2011) study, the participating fourth and seventh graders (i.e., at the stage of learning to read and of reading to learn in English) spoke highly of PWIM. Their perceived strengths of PWIM included increasing English vocabulary, improving descriptive writing, emphasizing word use, peer learning, and so on. These perceptive merits echo with Calhoun's (1999) descriptions about PWIM and are outcomes beyond systematic study of grammar translation, and rote of syntactic patterns and vocabulary (Hu, 2002).

Moreover, the participants' welcoming disposition about PWIM was able to inspire educators to reflect on Chinese Confucius culture in this modern era. With this historical heritage of Confucian, students can be trained with new ways of thinking, new roles in classroom, and new behaviours in a new student-centred classroom environment, which differ from Scollon's (1999) findings to a certain extent. Such a difference might be due to generational differences born before and after China's Open Door Policy in 1985. Reconciling cultural heritage and modern approaches in ELT is possible, seen from the participants' responses about PWIM.

I can imagine that Calhoun never considered implementing PWIM in such a crowded classroom with such a large number of students. From my observations and the following interviews, I noticed that students sitting far back away from the pictures could not see the pictures and words clearly. The seventh graders could go to see them after class because the pictures were put on the back blackboard, but the fourth graders did not have access after class.

Moreover, noise from peers kept students from hearing clearly. Even with classroom management, neither teacher was able to grant quiet moments for one specific respondent to make him/herself heard. Thus, students were more likely to be distracted and lost. This factor can be a rationale for any unsuccessful implementation of teaching methods, but it plays a more important role in student-centered activities, such as PWIM. With a PWIM format, a teacher can arrange any individual activity, as well as small-group and large-group activities (Calhoun, 1999). However, her concept of large groups may not be as large as a group of over 55 students.

Aside from findings of this research study, there are at least two main limitations which need to be taken into consideration. Firstly, there was no protocol for PWIM for students in secondary schools. Calhoun (1999), as PWIM founder, suggested that PWIM with modification can be used for older language beginners, though she didn't state how. In the current study, I trained two English teachers (one from an elementary school and one from a secondary school), and gave them the 10 steps of PWIM practice in a classroom (Calhoun, 1999, p.43), videos of PWIM demonstrations from YouTube and an implementation log (Calhoun, 1999, p.23) instead of a protocol. Such introductory one-on-one training gave the two teachers conceptual understandings, but not enough instructional strategies or practical preparations.

Secondly, PWIM was used in a crowded classroom with over 55 students, not individually or in small groups. The class size in this research study was larger than many in previous research, so the effectiveness of using PWIM to teach Chinese students English vocabulary may be reflected in the findings. The classes were so

large that the experimental intervention of PWIM may have had different effects to what Calhoun and other researchers have found. Those contextual factors may influence the effectiveness of PWIM.

This research study spanned a three-month period, so the data only reflected such a time period. For the current study, I only observed the group who had PWIM. Even for the two trial groups in Grade 4 and 7, I did not observe all of the PWIM sessions, because of schedule conflicts. For further research studies, it would be better to observe both the trial and non-trial groups, to see whether trial and non-trial classes have the same instructional conditions, the teachers have the same teaching style and the students in two groups play similar roles in teacher-student interactive communications. In that way, it may be further inferred which matters more, using the intervention or how the intervention is used.

In the current study, I interviewed students only instead of involved classroom teachers, in order to understand students' needs and opinions about PWIM. For future research studies about PWIM, both students and teachers should be interviewed, for the sake of understanding the other side in the two-way classroom interactions. Also, data from interviews with teachers, who are educational practitioners facing students every day, may reveal teachers' professional openness and personal dispositions toward new instructional methods, both of which might influence the effectiveness of the intervention. Teachers' voices from bottom up should be noticed and considered by education policymakers.

In the current study, qualitative data were collected from 18 fourth graders and 30 seventh graders, to answer the research question. The 18 fourth graders did not give much information as the 30 seventh grader did in all. One of the reasons might lie in the developmental differences between the age of 9 and 12, specifically, the ability of judging, evaluating and critiquing. The other reason was the fourth graders were pulled out to be interviewed as a cluster of three to five, based on the head teacher's request; thus, they were easily distracted by their peers at the time of interviews, and the seventh graders were pulled out and interviewed individually.

Conclusion

This study has investigated participants' self-perception of PWIM as a newly-tried ELT approach. Via observations, questionnaires and mainly interviews, I found that many participants spoke highly of PWIM and only a few mentioned its weaknesses, some of which were not considered as the method's inherent issues. All of the interviewees expressed the positive (either greatly positive or slightly positive) influence of PWIM on their English learning, with many rationales interfacing with the merits of PWIM.

ELT reformers and policymakers in China may consider the role of English teachers' professional development (to build up their constant exposure to, awareness of and strategic readiness for new teaching approaches) and contextualization (including contextual factors) when introducing new ELT methods. This being said, I hope that findings of this research study will contribute to the existing handful of research studies of PWIM in both ESL and EFL contexts.

References

- Astorga, M. C. (1999). The text-image interaction and second language learning. *The Australian Journal of Language and Literacy*, 22(3), 212-233.
- Bogdan, R. C., & Biklen, S. K. (2007). *Qualitative research for education: An introduction to theories and methods* (5th ed.). Boston, MA: Pearson Education.
- Calhoun, E. (1999). *Teaching beginning reading and writing with the picture word inductive model*. Alexandria, VA: Association for Supervision and Curriculum Development.
- Clay, M. M. (2001). *Change over time in children's literacy development*. Auckland, New Zealand: Heinemann Education.
- Crystal, D. (2008). Two thousand million? *English Today*, 24(1), 3-6. doi: <http://dx.doi.org/10.1017/S0266078408000023>
- Evans, R. (1996). *The human side of school change: Reform, resistance, and the real-life problems of innovation*. San Francisco, CA: Jossey-Bass.

- Fahrutdinova, R. A., Yarmakeev, I. E., & Fakhrutdinov, R. R. (2014). The formation of students' foreign language communicative competence during the learning process of the English language through interactive learning technologies (The study on the basis of Kazan Federal University). *English Language Teaching*, 7(12), 36-46.
- Feng, C. (2011). *The cooperative classroom: Scaffolding EFL elementary learners' English literacies through the picture word inductive model – the journey of three teachers in Taiwan* (Doctoral dissertation). Retrieved from <https://tspace.library.utoronto.ca/handle/1807/27578>
- Hu, G. (2002). Potential cultural resistance to pedagogical imports: The case of communicative language teaching in China. *Language, Culture and Curriculum*, 15(2), 93-105.
- Jiang, X. (2011). Challenges for college-level learners of academic English writing in China. In M. S. Plakhotnik, S. M. Nielsen, & D. M. Pane (Eds.), *Proceedings of the Tenth Annual College of Education & GSN Research Conference* (pp. 95-100). Miami: Florida International University. http://coeweb.fiu.edu/research_conference/
- Joyce, B., Calhoun, E., & Hopkins, D. (2002). *Models of learning: Tools for teaching* (2nd ed.). Philadelphia, PA: Open University Press.
- Joyce, B., & Showers, B. (2002). *Student achievement through staff development*. Alexandria, VA: Association for Supervision and Curriculum Development.
- Joyce, B., & Weil, M. (2004). The picture-word inductive model: Developing literacy across the curriculum. In B. Joyce & M. Weil (Eds.), *Models of teaching* (7th ed.). Boston, MA: Pearson Education.
- Joyce, B., Weil, M., & Calhoun, E. (2004). *Models of teaching* (7th ed.). Boston, MA: Allyn & Bacon.
- Joyce, B., Weil, M., & Calhoun, E. (2009). *Models of teaching* (8th ed.). Boston, MA: Pearson Education.
- Merriam, S. B. (2002). Introduction to qualitative research. In S. B. Merriam (Ed.), *Qualitative research in practice: Examples for discussion and analysis* (pp. 3-17). New York, NY: John Wiley & Sons.
- Rubin, H. J., & Rubin, I. S. (2005). *Qualitative interviewing: The art of hearing data* (2nd ed.). Thousand Oaks, CA: Sage.
- Scollon, S. (1999). Not to waste words or students: Confucian and Socratic discourse in the tertiary classroom. In E. Hinkel (Ed.), *Culture in second language teaching and learning* (pp. 13-27). Cambridge, U.K.: Cambridge University Press.
- Seidman, I. (2006). *Interviewing as qualitative research* (3rd ed.). New York, NY: Teachers College Press.
- Stake, R. E. (2005). Qualitative case studies. In N. Denzin & Y. Lincoln (Eds.), *The Sage handbook of qualitative research* (3rd ed., pp. 443-466). Thousand Oaks, CA: Sage
- Swartzendruber, K. (2007). *The picture word inductive model and vocabulary acquisition*. Proceedings of the 3rd Annual GRASP Symposium (pp. 177-178). Wichita, KS: Wichita State University. Retrieved from <http://soar.wichita.edu/dspace/bitstream/handle/10057/850/grasp%20178.pdf?sequence=1>
- Wong, A. C. (2009). *The influence of picture word inductive model on kindergarten students' development of literacy skills* (Unpublished master's thesis). Ontario Institute for Studies in Education of the University of Toronto, Toronto, Canada.
- Yang, Y. (2000). *History of English education in china (1919- 1998)*. Retrieved from <http://www.eric.ed.gov/PDFS/ED441347.pdf>
- Yang, Y. I. (2014). The implementation of speaking fluency in communicative language teaching: An observation of adopting the 4/3/2 activity in high schools in China. *International Journal of English Language Education*, 2(1), 193-214.
- Yen, R. T. (1987). Foreign language teaching in China: Problems and perspectives. *Canadian and International Education*. 6(1), 48-61.
- Zheng, X., & Borg, S. (2014). Task-based learning and teaching in China: Secondary school teachers' beliefs and practices. *Language Teaching Research*, 18(2), 205-221.

About the Authors

Xuan Jiang is an Assistant Professor in TESOL at St. Thomas University. She teaches TESOL courses in methods, strategies, curriculum design, instruction and assessment to graduates and undergraduates. Her research focuses on cognitive linguistics, contextualization and de-contextualization of borrowed ELT curriculum, and instructional scaffoldings for English learners in content areas.

Critical Pedagogy in EFL Teacher Education in the United Arab Emirates: Possibilities and Challenges

Rana Raddawi*

DePaul University, United States

Salah Troudi

University of Exeter, United Kingdom

Abstract

This study examines the possibilities and obstacles present in adopting a critical approach to English language education in the United Arab Emirates (UAE). Three main elements of critical pedagogy in language teaching were suggested: cultural representation in the curriculum, perceptions of global English(es), and local and global issues. Within a research design informed by an adapted action research methodology, six secondary public schools from three emirates were selected for inclusion in the study. The pre-action stage included questionnaires, interviews, and class observations. An action plan and intervention followed this first phase of data collection. The intervention was informed by Freire's (1996) "Conscientization" approach, Shor's (1992) generative themes, and Heaney's (1995) codification and problem-posing methods of teaching. The action phase consisted of a hands-on workshop for 20 volunteers from the 42 pre-phase participants. The results indicate that while teachers showed interest and even enthusiasm about critical pedagogy, they were aware of a number of obstacles and challenges in applying it in their classrooms.

Keywords: critical pedagogy, English teacher education, secondary schools, United Arab Emirates

Introduction

Globalization in its different manifestations, along with technological development and the internationalization of educational settings, has prompted major educational reforms at both the organizational and school curriculum levels. Furthermore, global economic crises and fierce job market competition are calling for distinction and innovation; the Arab world is no exception. The nature of schooling in Arab countries is urging reform of education based on critical thinking, innovation, and democracy (Akkary, 2014; Al-Suwaidi, 2010; UNESCO report, 2005)

There have been tangible changes affecting educational settings in most Arab nations in order to meet international standards. These have included the adoption of conventional teaching methods and curricula approaches, along with the jargon of education academia (Alrabai, 2016; Badry, F. & Willoughby, 2016; Education System and Curriculum in Dubai and UAE Schools, 2016). The researchers believe, based on their several years of teaching experiences in the Gulf, that a ubiquitous component is still missing in these educational contexts. There is still a lack of discourse that pays attention to the joy and essence of learning, the quality of

* Email: rraddawi@depaul.edu. Tel. 16304028328. Address: Department of Modern Languages, DePaul University, SAC 310, 2320 N Kenmore Avenue, Chicago, IL 60614, USA

teaching, and classroom content and its connection to the outside world including learners' "experiences and history" (Freire, 1996, p. 17). This discourse is the realm of critical pedagogy that can help in introducing these pedagogical elements in the classroom. Critical pedagogy which deals mainly with "politics of difference" (Pennycook, 2001, p. 114) links classroom environment to the wider society. This localized study followed an adopted action research approach to address this perceived lack of critical pedagogy within the ELT (English Language Teaching) space of the UAE.

Literature Review

Critical Pedagogy: History and Meaning

The term critical pedagogy is often associated with the work of Paulo Freire, especially *Pedagogy of the Oppressed* (1996—first published in 1968). In the traditional student-teacher relationship, Freire (1996) characterizes the teacher as the authoritarian figure who transfers decontextualized, impersonal information to passive students, perceived as objects. Freire refers to this type of education as "banking education" where the teacher is the "depositor," the students are the "depositees," and the educational experience itself is "an act of depositing" (see p. 53).

Freire (1996) contrasts banking education with "liberating education" (p. 53) in which teachers do not focus on transmitting information in a mechanical way but rather help students develop their cognitive abilities (see also Giroux, 2011). They maintain that critical evaluation and personal development are inherently human tendencies, so even those trapped by the shackles of traditions can free themselves and develop refined intellects if provided with a nurturing environment. In the learning situation of liberating education, the teacher-student hierarchy is reduced and instead they become co-learners in the classroom where information is shared through dialogue. Even though Freire did not coin the term, his views on educational reform form the foundational pillars of what we now refer to as critical pedagogy. Cho (2013) states that critical pedagogy shares many of its core principles with other critical theories prevalent at the time, such as social constructionism and postmodernism. Nevertheless, there was a gap that warranted the emergence of critical pedagogy.

A major concern of critical pedagogy is the nature of knowledge constructed and transmitted in society and schools. McLaren (2009) captured this concern stating that "critical pedagogy asks how and why knowledge gets constructed the way it does, and how and why some constructions of reality are legitimated and celebrated by the dominant culture while others clearly are not" (p. 63).

At the level of English language teacher education, there is a dearth of literature on how to introduce critical pedagogy to teachers in the Arab world in general and the Gulf region in particular. In her transformative L2 teacher development model (TLTD), which is based on twenty critical pedagogy principles adopted from Crawford (1978), Izadinia (in Wachob, 2009) argues that the practicality and feasibility of developing teacher education programs around the tenets of critical pedagogy are potentially tangible. She fends off criticism of critical pedagogy as practically gloomy, warning against the legacy and effect of the banking model of teacher education.

To date, Wachob's (2009) edited book of critical pedagogy studies conducted in the Middle East is one of the few compilations devoted entirely to the discipline in this part of the world. In the UAE, critical pedagogy is still in its initial stages of classroom practice. To the best of our knowledge, only a few papers examine how the implementation of critical pedagogy can promote strong critical thinking skills in the UAE (Clarke & Otaky, 2006; Hall, 2011; McLoughlin & Mynard, 2009; Raddawi, 2011; Raddawi & Troudi, 2012; Smith, 2011). Referring to Qatar's new critical thinking-based educational reforms, Romanowski and Nasser (2012) write: "Religion and tradition... govern the political, economic, social, legal and educational aspects of society" (p. 124). The authors highlight religious principles in various MENA (Middle East and North Africa) states which inform civil practices, suggesting that these practices thus become "beyond question" (p. 125). Tertiary level education in the MENA region may still appear incompatible with Freirian pedagogy's focus on the "common good." For

example, Salame (2011) points out that higher education has neglected sustainable development. Some Middle-Eastern states even practice “state censorship” (Romanowski & Nasser, 2012, p.125) and academic freedom is still not seen as a fundamental facet of higher education (Nasser & Abouchedid, 2007).

Cultural Representation

Critical pedagogy is about relating classrooms to social, cultural, political, and ideological concerns (Auerbach, 1995; Benson, 1997). Troudi (2005) points out the necessity of critical knowledge for the TESOL teacher, which requires an awareness of the socio-cultural contexts of the students and “how these shape their approach to learning and attitudes to English as a second or foreign language” (p. 1). Critical pedagogy rejects the distancing of culture from the political and economic life-processes of society; it “cannot be abstracted from the historical and societal context that gave it meaning” (Freire, 1996, p. 39). Culture has become an “object” in Western society repressing its critical elements and even negation of its critical thoughts (Adorno, 1975; Horkheimer, 1972; Lowenthal, 1979, Marcuse, 1978 as cited in Freire 1996, p.40). What these authors mean by western society is probably all communities that are directly derived from and influenced by European cultures.

To some, culture is “a historically transmitted pattern of meanings embodied in symbolic forms by means of which men communicate” (Geertz, 1973, p. 89). To others, it is what people must know to act and make things in a distinctive way (Holland & Quinn, 1987). However, Samovar and McDaniel (2012) argued that culture is not static but dynamic and “transgenerational” at the same time. It is more than ethnicity, where members of a community inherit patterns such as skin color, food, and folklore. It is beyond the four ‘F’ approach advanced by some multiculturalists essentializing culture as Food, Fashion, Festivals, and Folklore (Banks, 2002; Sleeter & Gran, 2008;). Culture is the ‘conscientization’ (Freire, 1996) process in which individuals are aware of their own identities and way of living (one’s self) while at the same time acknowledging variation and the other. We see culture as a set of shared experiences by members of a community in everyday life (Raddawi, 2015), and as “a field of struggle in which the production, legitimation, and circulation of particular forms of knowledge and experience are central areas of conflict” (McLaren, 2009, p. 65). It is this view of culture that will be investigated in the present study.

Global English and World Englishes

The concept of World Englishes has been addressed within the wider framework of critical applied linguistics and a critical stance to teaching English to speakers of other languages (TESOL). As English has become globalized, many varieties of the language have emerged in different parts of the world. Kandiah (1998) explains that even in the earlier stages of its development within the confines of the British Isles, the English language, like any other national language, had remarkably different varieties. However, the language’s spread from its homeland to settle in foreign territories “caused it to become even more differentiated’ because it entered “new and unfamiliar contexts...marked by specific ecological, cultural, linguistic, and other characteristics...radically different from those of England” (Kandiah, 1998, p.2). He aptly summarizes the journey of the language from its homeland to new sociocultural settings as comprising three stages: “transportation, transplantation, and adaptation” (p.12). Kachru (1982) uses the term “acculturation” to describe the adaptation of a given language and adds that it results in “linguistic innovation” as foreign cultural and linguistic elements seep into the borrowed language. Scholars have used a variety of names to group together various ‘Englishes’ across the globe, including “New Englishes,” “Global Englishes,” and “World Englishes” (Jenkins, 2014; Kirkpatrick, 2007, 2012). Kachru and Nelson (2006) state that the unparalleled expansion of the English language has led to the notion of World Englishes and its major varieties include European, North and South American, African, and Asian English. In reference to Kachru’s model of the three concentric circles of English (1982), the term World Englishes (WEs) encompasses all varieties of English—inner, outer, and expanding circle varieties (Sharifian 2009).

The Arabian Gulf region faces some major effects of the global spread of English, which is considered the medium of instruction and the lingua franca in the UAE, Qatar, and Oman. Mahboob (2013) states that despite the significant position of English in Middle Eastern countries, where it is taught as a school subject and used as a medium of communication in social interactions as well as in published materials, research on “the use of English in the region from a World Englishes perspective” (p. 14) is scarce. In fact, there is little research on the nature of the English language in the majority of Middle Eastern countries. Only eight articles published in *World Englishes Journal* focus on the use of English in the Middle East and just four out of these eight articles include a regional author (see Mahboob, 2013). A study by Abdel-Jawad and Abu Radwan (2011) exploring the nature of English used in tertiary institutions in Oman showed that it was used mainly for academic purposes as the medium of instruction and for communication purposes such as internet use, meetings, publications, and advertisements. A recent edited volume by Kirkpatrick (2017) provides insights into English language education in the MENA region. However, there was no mention of which English variety was used in educational institutions. The assumption or “natural position” is that it is one of the two inner circle varieties, British or American English.

Research Questions

The present study aims to examine the situation of EFL teacher education in public secondary schools in the UAE in an attempt to introduce a critical approach to teaching and raise teachers’ awareness about the feasibility of critical pedagogy. The following two research questions informed the design of the study:

1. How familiar is the EFL teacher in the United Arab Emirates with critical pedagogy?
2. What are the possibilities and challenges of introducing three elements of critical pedagogy: cultural representation, World Englishes and local and global issues into the curriculum of the secondary EFL teacher education in the UAE?

Methodology

Research Design

The research design of the study is informed by an action research approach (Zuber-Sklerri, 1996) with a mixed-method design adapted to serve the critical aims of the study (Kemmis & McTaggart, 1988). The action/intervention element of the study allowed us to introduce three elements of critical pedagogy through a workshop to pre-service and in-service EFL teachers in the UAE. These elements are cultural representation, local and global issues, and World Englishes. The study had three phases: pre-action, action, and post-action. In the pre-action phase we explored how the EFL teachers perceived critical pedagogy and its place within the secondary curriculum, and the strategies they were prepared to adopt in their classrooms. Using a questionnaire and an interview at this stage helped us identify the nature of the professional development needs of these teachers. The action phase consisted of the intervention, which was a workshop delivered to EFL secondary school teachers. The post-action phase included a group discussion with the workshop participants to identify the feasibility and challenges of introducing the three selected elements of critical pedagogy to EFL in secondary schools. This marked the evaluation stage of the action research project.

Data Collection

The data collection in the first stage consisted of documentary analysis, questionnaire (see Appendix 1), semi-structured interviews (see Appendix 2), and observations. Documentary analysis was performed on print and online databases in addition to sample documents related to the EFL curriculum in the selected schools such as syllabus, curriculum map, lesson planning, textbooks, sample class handouts, exams, and students’ work to investigate whether the three elements of critical pedagogy under scrutiny were evident. The goal behind the

interviews was to assess teachers' awareness and familiarity with elements of critical pedagogy and whether it is possible to make it part of their training. Upon the request of the participants, a note-taking process with no audio recording was used during the interviews and class observations. All interviews were conducted in formal academic English, with interviewees using a variety of Englishes such as Australian, British, Philippine, Pakistani, Indian, New Zealand, and American Englishes depending on their background. Field notes recording what was said by the participants were later shared with respective participants for verification purposes.

The data collection during the action- and post-action phases revolved around a workshop on critical pedagogy (the 3 elements driving this study) and the researchers' observations/reflections during the workshop. The data collected during the first stage drove this workshop.

Participants

A stratified purposive sampling technique was used in the study (Patton, 2002). The goal was to select teachers from representative public schools of both genders in some of the emirates.

Pre-action Phase

Three female and three male public schools at the secondary level from three emirates, Sharjah, Ajman, and Dubai, were selected to participate in the study. For the pre-action phase, 42 EFL teachers completed the questionnaire and 24 teachers from this group were available and agreed to the semi-structured interviews. Of the 42 teachers, 26 were females teaching in the girls' schools while 16 males were teaching in the boys' schools. These 42 participants and their contributions to the study would also act as the participant pool for the subsequent stages (see action/post-action discussions). Table 1 summarizes the pre-action phase participants' information.

Table 1

Pre-action phase participants' information

Number of Teachers:	42
Gender:	26 Female and 16 Male
Ethnicity:	3 Emiratis 8 European, mainly British 2 Indians 1 Pakistani 1 Pilipino 27 from Palestine, Jordan, Lebanon, Syria, Tunisia, Algeria, and Sudan
Years of Experience:	More than 5 years
Pre-service Training:	5
In-service Training:	42

All teachers had more than five years of teaching experience whether in the UAE or in their home countries. Two of the three Emiratis were relatively new to the teaching profession.—Five out of the 42 participants had pre-service training, and all had in-service compulsory training whether through workshops or seminars. This is called 'professional development' and focused mainly on preparation for the CEPA (Common Entry Proficiency Assessment), a compulsory test that local students need to take prior to entering college.

Action Phase

Following the data collection or pre-action phase, which revealed an absence of critical pedagogical approaches in the English Language teaching process observed in the six public schools, an intervention phase was conducted via a workshop at a university in one of the emirates. An invitation to participate in the four-hour

workshop was sent to all 42 EFL teachers who took the questionnaire. Twenty of them replied positively to the invitation and registered to attend the workshop at a local university. There were 4 Westerners (2 British, 1 Australian, and 1 New Zealander) and the remaining participants were from the MENA region. The 20 teachers varied in terms of number of teaching years.

Ethics

All ethical dimensions and procedures of participant consent, anonymity, confidentiality, and the right to withdraw from the study were observed. Names of schools were kept anonymous and pseudonyms were assigned to participants.

Findings

Pre-Action Phase

Surveys and interviews of teachers revealed an absence of critical pedagogy awareness. None of the participants had heard of the term critical pedagogy, though two out of 42 asked whether it referred to “critical thinking.” However, these two teachers, who came from the MENA region, reported being unable even to provide critical thinking-based activities in the classroom due to time and curriculum constraints; both had more than ten years of teaching experience.

While observing the two classrooms, the researchers could confirm the by-rote and lecture-based teaching methods. The “banking” process was followed in the first classroom. None of the students took the initiative to talk unless asked a particular question. However, in the second-class observation, the teacher tried to deviate from the textbook content to apply the formation of simple and complex sentences to some real-life situations such as the excessive use of mobile phones and their harmful effects on the students.

Cultural Representation

In the curriculum. Upon examining the teaching materials used by the participants in the different schools, the researchers could confirm that all public schools used the same textbook, “On Location” (Bye, 2011) for English language teaching, which is a series of three books that cover grades 10-12. In this context the term “curriculum” is used in a narrow way to refer to a set of handouts and course packets developed locally by the teachers (as is the case of Schools of the Future) as instructed by the Ministry of Education in the UAE or textbooks assigned for development by an international publisher. Teachers needed to cover all the units included in the textbook. In addition to the assigned textbook, they are required to prepare their own teaching materials based on the four language skills. Teachers have no say in designing the curriculum and are restricted in class content and time of delivery. Final exams and midterms are prepared and sent by the Ministry of Education. This information was conveyed to the researchers by some of the teachers during the interviews.

All 42 participants agreed that the “On Location” textbook does not match the students’ needs and culture. Below is an excerpt from our field notes of what Samah, one of the interviewed EFL teachers, said:

The previous book had more practical exercises on the four skills and further cultural references such as the high rate of dowry in the UAE and divorce issues. The actual textbook has one unit in Grade 10 called ‘Proud to be Emirati’ that refers mainly to the Emirati National Day. The rest varies between animals in the world, natural disasters and some ‘know-how’ that does not relate to the local context such as ‘how to assemble a skateboard’ or ‘how to prepare a pizza.’

Teachers expressed their dissatisfaction and frustration with the curriculum by using expressions such as: “we are slaves,” “nobody listens to us,” and “we hope that you can convey our voices to concerned authorities.” According to our field notes, Ahmad stated in an interview:

I wished I could do what I used to do in my home country...I once changed the textbook in the middle of the year [in my home country] as I noticed that the students were not responding to it properly.

While there is clear evidence of frustration and helplessness *vis-a-vis* the contents of the prescribed textbooks, there are also signs of teacher initiative and independence as described by Ahmad above.

In teaching philosophy and class interaction. Thirty-seven (88%) of the surveyed teachers said they delivered the traditional lecture format while 5 (12%) of the teachers said they discussed issues that are related to students’ lives and experiences. For example, Iman, an English teacher with more than 15 years of teaching experience and a mother, stated that female students would tell her about their relationships with their mothers and take her advice on how they can improve them. Sarah described how she shows the picture of her favorite corner at home and asks the students to describe their preferred corners in English. Other examples of teachers’ efforts are from Mustafa who made analogies and references to the local culture every time there was an opportunity, for example by asking them to talk about the heritage in one of the emirates compared to the Australian culture referred to in the book. Suad stated she would discuss issues such as the excessive use of mobile phones and its impact on students’ lives while she is explaining the means of communication in general or discussing segregation in education, and so forth.

Conversely, in the same context, we noted that Maha said in an interview:

I don’t know how to tackle cultural issues. For example, I don’t know how to discuss the issue of high rate of divorce in the UAE or racism in class, I never had courses on Intercultural communication.

Another issue raised by the teachers is the disparity between textbook contents and the exam questions. The latter come ready from the Ministry and the teachers have no knowledge of their content prior to the exam date. Two of the sample grade 12 final exams in English had questions on monuments in Dubai and Abu Dhabi whereas none of the textbook units had these cultural references. Exam questions required students to write a composition about these monuments.

During class observation, the only cultural reference used during the entire session occurred when one teacher, who dominated the speaking in the class, drew on the board a local senior woman wearing an *abaya* (traditional long attire) talking to her granddaughter.

Local and Global Issues

Teachers were not really concerned about relating classroom content to issues beyond the classroom, whether related to local or global topics. They were preoccupied with covering the assigned material to which they had no contribution. As Ibrahim stated, ‘There is no way to link the readings in the textbook to the students’ real world, you need to be creative and this requires time’. When asked in an interview whether she would discuss a subject like disabilities, we recorded that Maha’s answer was straight to the point:

NO! It is a taboo. Neither parents nor students admit that their children have disabilities even if they are curable such as dyslexia.

When asked in an interview whether he relates classroom content to students' lives, Samir, an English teacher in one of the male schools was recorded as saying:

I do not see why we should bother about these issues since students are demotivated. They have a ready-made job waiting for them after graduating as long as they pass. They either go to army or to police stations. In both cases, they earn a good income.

A similar response was heard by the teachers in the female schools. Some comments were: "Most of the girls dream of getting married after graduating, that is their only ambition" and "...the same goes for parents in terms of lack of interest in their children education' said other teachers."

In the second-class observation, the teacher provided examples of real life experience such as the excessive use of mobile phones and their impact on the youth. She used these examples while revising a grammar lesson. The teacher asked them to provide examples of simple, compound, and complex sentences using excessive use of mobile phones as a theme. This observed pedagogical behavior contradicts the testimony of the interviewees who claimed that they were incapable of tackling local and/or global issues other than the ones mentioned in the textbook.

Absence of Englishes

The textbook "On Location" refers to British versus US spelling when there is a difference and sometime points out vocabulary variations. The handouts collected by some teachers to supplement their teaching material are selected from the Internet and are either written by British or American authors who would use regional vocabulary and spelling.

None of the interviewed teachers had heard about Englishes or Global English and paid little attention in their teaching to any kind of English apart from British vs. American in terms of existence, usage, and variations if any. In fact, in one of the two class observations, one of the researchers could hear the Emirati teacher impersonating the Indian accent by suddenly changing her accent to imitate the common Indian accent used in the Gulf region. Kubota (2009) states that liberal multiculturalism includes open-mindedness and non-prejudiced attitudes in interacting with people with diverse racial, ethnic, and linguistic backgrounds. She suggests that "liberal multiculturalism promotes tolerance, acceptance and respect toward different cultures and culturally diverse people while supporting equality among them" (p. 30). This was certainly not the case for the teacher who imitated the Indian accent in the first class observed for this study.

Despite the fact that the majority of the participants did not belong to what Kachru (1985) called the "Inner circle" of English speaking countries, they were still compelled to teach according to the English teaching curriculum informed by inner circle countries, mainly the UK and the US.

Implications of Pre-action Phase

The generally qualitative findings reveal a common frustrating and demotivating teaching environment. There is a noticeable disconnect between the three main stakeholders of the studied educational settings: decision makers, teachers, and students. On the whole, teachers seem to be hopeless, disempowered, and demotivated. According to their teachers, students are demotivated because they have a ready-made professional future. This is of course not necessarily the case as there is an issue of unemployment among nationals of the UAE. Decision makers are located at the Ministry and send their correspondents or representatives to check that set policies and curriculum are well implemented. Yet, from the findings, it seems the reality on the ground is different. Policies are in place but the curriculum is disruptive and most of the time, it does not match final exam questions and students' aspirations. This could be confirmed by the low grades and negative outcomes of the English courses as reported by the teachers interviewed for the study.

The fact that, in the observed classes, most of classroom talk was teacher-controlled is a reflection of the dominance of the "banking" model of education (Freire, 1996). Class observation confirmed this view of the

teacher being a “depositor” of information and students a “depository”. According to Freire, the banking approach will never encourage students to consider reality critically. Students need to “domesticate reality” (Freire, 1996, p. 56). They could then perceive reality as a “process,” a constant transformation.

Students’ lack of interest or demotivation can be traced to the disconnection between classroom content and their socio-cultural contexts. Some teachers confirmed that when issues related to students’ heritage are discussed in class, learners show a considerable enthusiasm. Giroux (1988a, 1988b), Kanpol (1994, 1997), McLaren (1989), and others developed a critique of formal education to understand the cultural politics of schooling, addressing the marginalization and exclusions of schooling by encouraging students to develop their own voice. Education is more than “speaking” or “writing,” it is rather another way of articulating reality (Pennycook, 2009, p. 130).

Giroux (1983, 2011), Pennycook (2007a), and Freire (1996) interpret students’ absences, low performance and grades, disinterest in the curriculum, and misbehavior as a form of “resistance.” Kumaravadivelu (1999) observes that sometimes students’ lack of preparation and lack of ability to participate in class discussion is a form of “passive resistance” (p.454). Canagarajah (1993) suggests that this resistance in the classroom may play a role in “larger transformation in the social sphere” (p. 996). Gramsci (as cited in Darder, Marta, & Rodolfo, 2009) believes that the hegemonic spirit dominating in schooling reflects the “hegemonic process that reproduced cultural and economic domination within the society” (p. 7).

The fact that a few teachers have different teaching methods than others showing certain praxis in class means that change is not impossible. It implies that teachers can make a change even in the most restrictive teaching environment. The language of “possibility” and praxis (theory and agency together) are not a myth even in the most hegemonic teaching environments. Explaining a pedagogy of hope, Freire (1992) stresses that “one of the tasks of the progressive educator, through a serious, correct political analysis, is to unveil opportunities for hope, no matter what the obstacles may be” (p. 3).

Most teachers in this study did not have pre-service training which explains the narrow definition they ascribed to curriculum. Some were not familiar with the terminology to describe the components of a language curriculum (Troudi & Alwan, 2010). When asked about their role in the curriculum, they automatically referred to the textbook and supplementary handouts. They were mostly following a traditional pedagogy which according to Moreno-Lopez (as cited in Wachob, 2009) is a name assigned to a period that favored pre-defined syllabi and focused on agreed-upon course materials to be taught.

The limited reference to World Englishes and the emphasis of the studied curriculum on the “inner-circle” English reflect EFL teachers’ lack of awareness of the importance of the three circles of Englishes and the evolving positions of English beyond its original geographical origins. The study also unveiled an absence of knowledge about the world’s “shift of gears” from what was once “supremacy” of inner English as opposed to outer and expanding circles (Aktuna & Hardman, 2008; Kachru & Larry, 2008). It is not about a variety of English used from inner or outer circle in the classroom but rather the awareness that there exist Englishes other than the inner circle English in use in many parts of the world. These World Englishes are, or should be, considered by their users to be just as accurate and legitimate as the English of inner circle to its speakers. The English curriculum and textbooks of the UAE reflect an exonormative native speaker model (Kirkpatrick, 2007). This refers to the deliberate choice of a native speaker model of English as a reference for teaching and learning. Teachers’ training and learning experiences have also been shaped by this model which is automatically reinforced in their classroom practice. In the case of the UAE, an expanding circle country, there is a logical and historical explanation for the exclusive choice of an inner circle model. Its legitimacy and prestige have long been established through media, public institutions and educational policies. Pragmatically, native speaker models, codified curriculum planners and teachers have easy access to dictionaries, grammars, reference tools and materials made available by established and competing publishing industries in the US and Britain. For policy makers in the UAE and similar contexts, codification also “brings with it the notion of acceptance as a standard-learners can be tested and evaluated against codified norms and standards” (Kirkpatrick, 2007, p. 184).

Action Phase

Structure of the Workshop

The workshop consisted of two parts. The first part was a review of the theoretical framework providing some background information on critical pedagogy. The second part of the workshop discussed the mechanism of introducing the three elements of critical pedagogy under study.

The hands-on workshop was based on the Freirian Conscientization concept in which students are viewed as subjects rather than objects in the classroom and in the world. Conscientization about the student teachers' cultural background, the country's linguistic reality and the link of classroom content to the outside real-life context was at the heart of the intervention. Three methods for achieving conscientization (Izadinia, 2009) were applied, first through codification (Heaney, 1995) then generative themes (Shor, 1992), and eventually the problem-posing method (Freire, 1996). An example of the codification approach was initiated by showing the teachers a picture of a child on a wheelchair and asking them what references the picture evoked. There were many responses to the question such as "disability," "special education," "taboo," "accident," "high speed," and "victim."

We divided the 20 student teachers registered for the workshop into five groups and asked each group to discuss one of the suggested references. The first group chose the theme of "disability and its different implications in the UAE," the second group discussed "taboos," the third chose "accidents," the fourth discussed "high speed," and the fifth opted for "victims of road accidents." At the end, each group had to present to their peers their synthesis on the subject thus narrating their own experiences and attitudes towards the topic in real life and sharing their knowledge of the subject matter. An element of Freirean Praxis (action-reflection-action) in this case was attained.

Subsequently, we followed the generative themes approach by writing the word "divorce" on the board. Teachers had different themes proposed upon hearing the word. Some suggested "high rate of divorce in the UAE," others said 'abandoned children in the Arab world' and so forth. The audience was divided into groups, with each group discussing one of the suggested sub-themes then sharing results with the rest of the participants. When asked if this method could be applied in real classrooms, 15 of the 20 volunteer teachers were positive about it. The third application method was the "problem-solution" approach which was introduced to the student teachers by sharing what could be a common and intriguing problem in the Gulf: "high speed." It was left to the audience to suggest solutions with arguments.

For the focus on Global English, the student teachers were given five anonymous texts written by authors from Kachru's three circles and were asked to identify the type of English in which each text was written. Ten (50%) of the teachers could distinguish between British, American, Australian, Kenyan, and Nigerian Englishes. This low rate can be explained in two ways. First, it is possible that the texts did not have significant variation of English as Crystal (2000) shows in their study of a number of newspapers published in various English-speaking countries. Crystal could find various cultural references and names of locations but not real linguistic differences among the studies texts. Another explanation could be that the teachers were not aware of Englishes other than UK or American; we believe it was the latter case.

Reactions to the Action Phase: Enthusiasm and Caution

The hands-on workshop outcomes reflect the reaction and attitudes of 20 EFL teacher volunteers out of the 42 participants who showed enthusiasm about introducing critical pedagogy in their classroom while at the same time expressing some concerns and challenges. The three elements of critical pedagogy: cultural representation, local and global issues, and Global English, were considered during this action phase. Enthusiasm was observed when the researchers introduced the concept of "Conscientization."

Yet, while the 20 EFL teachers showed great interest and motivation towards the introduction of elements of critical pedagogy in their teaching, they expressed some concerns about the applicability of this approach. The

workshop ended with a discussion and recommendations as to how these three core components of critical pedagogy could be introduced in the best way possible into the EFL curriculum.

The four points below are the main findings from the post-action phase:

1. *Constraints on academic freedom.* Educators thought that they should have the freedom to discuss any issue in the classroom provided it is related to classroom content. Lindsay was recorded as stating:

We would love to have some space of freedom to discuss hot global issues in the classroom without being penalized later...

2. *Lack of teachers' participation* in setting up the curriculum and mainly not having a say in the choice of textbooks was raised by participants, we noted Suad's words as follows:

We are like slaves...we do not have any say in choosing the teaching material and especially the textbook we teach.

3. *Lack of cultural competence* to be able to tackle culturally sensitive issues was an interesting finding. Our field notes recorded Fuad's explanation:

To be honest, I don't think I have the necessary skills to discuss culturally sensitive issues in my class but this workshop helped in many ways and wish to see more of Intercultural Communication sessions in our Teacher Development Programs.

4- *Awareness of World Englishes*

This workshop was an eye opener to many of us on the variety of Englishes in the world. There should be more of these texts written by authors of outer and expanded circles in the textbooks we teach but I can always add some into the supplementary teaching material. (Suad's statements as recorded in field notes)

Teachers' suggested solutions to these obstacles were to empower teachers by involving them in the curriculum development process and by adding an element of critical pedagogy to their in-service training programmes. Maha and Ahmad expressed a number of needs:

Teachers need more pre-service training in critical pedagogy and deeper knowledge of Intercultural Communication. (Maha's statements as recorded in field notes)

Individuals who are involved in creating the curriculum should be the same as the ones teaching it to experience the challenges that EFL teachers face in the classroom. (Ahmad's statements as recorded in field notes)

Ahmed also suggested transferring what he learned in the workshop into his classroom practice (according to our field notes):

Look when we brought up the sub-topics to be discussed in this workshop, it worked very well and all participants were excited to take part in the discussion, why not do the same in the actual classroom and let students choose the topics they would like to examine...?

Summative Discussion

Teachers can be empowered through critical pedagogy by following relevant pre- and in-service training programs. Teachers' training could consist of cultural studies component, which will increase their cultural awareness (Samovar & McDaniel, 2012), and avoid "cultural essentialization" (Kubota, 2001 as cited in Aktuna & Hardman 2008, p. 168). Being culturally competent requires an awareness of the self and the other. Cultural competence is more than knowing about the existence of other cultures, it is also understanding how to approach and discuss culturally sensitive issues.

For many EFL teachers, the fact that they are multilingual and multicultural regardless of their ethnic, educational or cultural background can help in integrating critical pedagogy in their teaching. In fact, "their multilingualism will serve to know and share their students' concerns and experiences in learning a language" (Kirkpatrick, 2007, p. 187). He stresses that this multilingualism of English teachers gives them the advantage of understanding their students' difficulties and puts them in a position to empathize with them. In addition, when educators are exposed to new teaching and learning methods where the teacher and students exchange roles, collaborate (Lang & Evans, 2006) and together negotiate the curriculum (Norton & Toohey, 2009), learning becomes a joy and a pleasure to both the teacher and student.

Awareness of World Englishes, if incorporated into the teacher education curricula, can help trainees understand the local uses of English within a global context of communication. Also, such knowledge should counteract negative attitude towards variation and variability of world Englishes, especially since EFL learners usually use their English in outer or expanding circles rather than within inner-circle environments. Furthermore, EFL teacher education could present a broader scope of the ownership of English and show the teachers that English authority and ownership are not limited to those with "native like pronunciation and knowledge of sociocultural norms emerging from inner-circle countries" (Aktuna & Hardman, 2008, p. 167). Equally relevant and important constructs here are English as a Lingua Franca (ELF) (Jenkins, 2007, 2014; Mauranen, 2012) and English as an International Language (EIL) (Holliday, 2005; Jenkins, 2006). The two concepts along with World Englishes have important implications for English language teacher education, English language testing and ELT materials. Jenkins (2008) for example explains how an ELF approach will be increasingly needed to resolve problems of mutual intelligibility. This does not exclude native speakers from Britain or the US who have to "adjust their English for international communication" (Jenkins, 2008, p. 237). It should be noted that EIL does not escape criticism as it has also to be revisited against claims that it is a tool benefiting Western powers in a global race for international markets and resources (Phillipson, 2009). Pennycook (2007b) warns against the myths of EIL stressing that it should not be seen as a natural development. By doing so, he claims, we depoliticize English "making it innocent, giving it a natural and eternal justification" (Pennycook, 2007b, p. 109).

Another component in EFL teacher education is a clear distinction between professionalism in teaching and English proficiency. English proficiency is no longer limited to the mastery of grammar and lexicon (Nelson, 2011). There should be an understanding of the context in which the utterance is said and also an awareness of the "speech community" (Samovar & McDaniel, 2012). EFL teachers need training and deep knowledge of the language and the cultures (s) that shape it.

EFL pre-service training could foster critical approaches in teacher education. When English teachers have an understanding of how education is related to broader social and cultural relations rather than merely attempting to "fulfill predefined curricular goals" (Pennycook, 2009, p. 299), they can contribute to the making of knowledge in their classroom. Within this framework, the role of the language learner is not to imitate a "circumscribed and standardized model" of the native speaker but rather to act as a "border-crosser who negotiates between the universal" (the other/macro) and the "Particular" (the self/micro) "and combines a sense of belonging with a sense of detachment" (Giroux, 1994, p. 68).

The above statements contradict in some way Ellsworth's (1989) outdated claims that Freire (1996) and Shor's (1992) emancipatory authority implies that a teacher knows the object of study "better" than students do. During the workshop, the researchers proposed some local issues for discussion such as divorce and disability, but the students could also propose the topics and teachers facilitate the discussion, as is already the case in some writing courses in universities in the UAE. Ellsworth (1989) also hints at the failure of critical pedagogy to have a balance between teachers and students in terms of "institutionalized power imbalance" (p. 10). We believe that when roles between students and teachers overlap in the classroom, the goal somehow is to design class content and conceive programs that reflect this balance in and outside the classroom.

Furthermore, Johnston (1999) considers critical pedagogy as an exercise of EFL abstraction. However, the hands-on workshop showed the opposite; teachers spent almost two hours touching upon practical topics that relate classroom content to global issues through generative themes (Heaney, 1995) and codification methods (Shor 1992). The methods used allowed these teachers to "conscientize" not only about ethnic diversity in the classroom and hence in the country that hosts more than 200 nationalities (The National, 22 March 2009 as cited in Randall & Samimi, 2010) but also to "conscientize" about the possibility of overcoming some of the challenges that hinder their ability to introduce some elements of critical pedagogy into their teaching. There was no "dictation" to tell students teachers what to do. They led themselves into the different sections of the workshop while identifying some of the challenges and possibilities to apply this critical approach. This was a natural result of the hands-on workshop on critical pedagogy. A similar attitude from the students could be expected in a real critical teaching classroom.

Conclusion

This paper is an attempt to examine the possibility of introducing a critical approach to EFL teacher training in the United Arab Emirates. The ultimate goal is to empower teachers and subsequently their students to become agents of change. Conscientization is the awareness of being a subject rather than an object in the world (Heaney 1995) and according to Izadinia (in Wachob, 2009) 'conscientization' can be reached through codification or the generative theme or the problem-solving method.

The study revealed a gap in teachers' knowledge of critical pedagogy, mainly the three elements under scrutiny in this study: cultural representation, English in the world and local and global issues. A revamp of the EFL teacher pre-service and in-service training programs is required.

In the Arab world, critical pedagogy is still in the infancy stage. Although some teachers are conscientious about their role in promoting freedom of expression in the classroom and making their students social agents, they have not achieved that. If teachers themselves feel they are "slaves" in the educational system, we cannot expect them to teach to liberate their 'oppressed' students. In the Arab world critical pedagogy needs to engage in more than changing the teacher-student relationship (Freire, 1996; Izadinia, 2009; Kanpol, 1994; Kumaravavidelu, 1999). It should go beyond this equation to include the entire chain of command that is involved in the educational system, i.e., the policy maker-administrator-teacher-student relationship. The *culture of silence* that Freire (1996) refers to in which the dominant culture silences the oppressed through marginalizing or undermining any voice that challenges their authority applies as much to teachers as to students. Thus, teachers should "strive not only for educational advancement but also for personal transformation" (Izadinia, 2009, p.11) and therefore become transformative and critically minded intellectuals.

Future Actions

In the future, we intend to include all the Emirates and major cities of the UAE in a plan for a longer period of teacher preparation and initiation into critical pedagogy. We also intend to follow this by a series of classroom observations to see if and how teachers will incorporate any elements of critical pedagogies into their daily

teaching practices. In addition, future research in the UAE and the Gulf needs to investigate the learners' reactions to critical pedagogy and whether it can have any effect on the overall quality of their educational experiences and possibly their language proficiencies.

References

- Abdel-Jawad, H. & Abu-Radwan, A. (2011). The status of English in institutions of higher education in Oman: Sultan Qaboos University as model. In A. Al-Issa and L. S. Dahan (Eds.), *Global English and Arabic: Issues of language, culture and identity*, (pp. 123-151). Oxford: Peter Lang.
- Alrabai, F. (2016). Factors underlying low achievement of Saudi EFL learners. *International Journal of English Linguistics* 6(3), 21-37.
- Al-Suwaidi, J. (2010). Education and human resources development in the Gulf: challenges of the twenty-first century. In The ECSSR (Ed.), *Education and the Arab World: challenges of the next millennium* (pp. 331-333). Abu Dhabi: The Emirates Center for Strategic Studies and Research.
- Akkary, R.K. (2014). Facing the challenges of educational reform in the Arab World. *JEduc Change* 15(2), 179-202.
- Auerbach, E. R. (1995). The politics of the ESL classroom: issues of power in pedagogical choices. In J. Tollefson (Ed.), *Power and inequality in language education* (pp. 9-33). New York: Cambridge University Press.
- Badry, F. & Willoughby, J. (2016). *Higher education revolutions in the Gulf*. New York: Routledge.
- Banks, J. (2002). *Cultural diversity and education: foundations, curriculum, and teaching*. New York: John Wiley & Sons.
- Benson, P. (1997). The philosophy and politics of learner autonomy. In P. Benson and P. Voller (Eds.), *Autonomy and independence in language learning* (pp. 18-34). London: Longman.
- Bruthiaux, P. (2003). Squaring the circles: Issues in modeling English worldwide. *International Journal of Applied Linguistics* 13(2), 159-178.
- Bye, T. (2011). *On Location*. Dubai, UAE: McGraw-Hill.
- Canagarajah, S. A. (1993). Critical ethnography of a Sri Lankan classroom: Ambiguities in student opposition to reproduction through ESOL. *TESOL Quarterly* 27(4), 601-626.
- Cho, S. (2013). *Critical pedagogy and social change: Critical analysis on the language of possibility*. New York: Routledge.
- Clarke, M. & Otaky, D. (2006). Reflection 'on' and 'in' teacher education in the United Arab Emirates. *International Journal of Educational Development* 26, 111-122.
- Crawford, L. (1978). *Paul Freire's philosophy: Derivation of curricular principles and their application to second language curriculum design*. Unpublished doctoral dissertation, University of Minnesota.
- Crystal, D. (2000). The future of English. Where unity and diversity meet. Retrieved from: http://www.davidcrystal.com/DC_articles/English37.pdf
- Darder, A., Marta B., and Rodolfo D. T. (2009). *The critical pedagogy reader*. New York: Routledge Falmer.
- Dogancay-Aktuna, S, & Hardman, J. (Eds.). (2008) *Global English teaching and teacher education: Praxis and possibility*. Alexandria, VA: TESOL, Inc..
- Education system and curriculum in Dubai and UAE schools (2016-2017). Retrieved from: <http://www.dubaifaqs.com/education-dubai.php>
- Ellsworth, E. (1989). Why doesn't this feel empowering? Working through the repressive myths of critical pedagogy. *Harvard Educational Review* 59(3), 297-325.
- Freire, P. (1996). *Pedagogy of the oppressed*. Translated by M.B. Ramos. New York: Continuum.
- Freire, P. (1992). *Pedagogy of hope: Reliving pedagogy of the oppressed*. London: Bloomsbury.
- Geertz, C. (1973). *The Interpretation of cultures: Selected Essays*. New York: Basic Books.
- Giroux, H. A. (1983). *Theory and resistance in education: A pedagogy for the opposition*. South Hadley, MA: Bergin & Garvey.

- Giroux, H. A. (1988a). *Schooling and the struggle for public life: critical pedagogy in the modern age*. Minneapolis: University of Minnesota Press.
- Giroux, H. A. (1988b). *Teachers as intellectuals: Toward a critical pedagogy of learning*. Granby, MA: Bergin & Garvey.
- Giroux, H. A. (1994). *Disturbing pleasures: Learning popular culture*. New York: Routledge.
- Giroux, H. (2009). Teacher education and democratic schooling. In A. Darder, M. Baltodano and D. T. Torres (Eds.) *The Critical pedagogy reader* (pp. 438-459). New York: Routledge.
- Giroux, H. A. (2011). *On Critical pedagogy*. New York: Continuum International Publishing Group.
- Hall, D. (2011). Debate: Innovative teaching to enhance critical thinking and communication skills in healthcare professionals. *The Internet Journal of Allied Health Sciences and Practice* 9(3), Article 7.
- Heaney, T. (1995). Issues in Freirean Pedagogy. *Thresholds in Education*/
<http://nlu.nl.edu/acc/Resources/Documents/FreireIssues.html>.
- Holliday, A. (2005). *The struggle to teach English as an international language*. Oxford: Oxford University Press.
- Holland, D. C. and Quinn, N. (1987). *Cultural models in language and thought*. Cambridge: Cambridge University Press.
- Izadinia, M. (2009) Critical Pedagogy: An Introduction. In P. Wachob (Ed.), *Power in the EFL classroom: Critical pedagogy in the Middle East* (pp. 7-15). Newcastle upon Tyne: Cambridge Scholars.
- Izadinia, M. (2009). Transformative L2 teacher development (TLTD): A tentative proposal. In P. Wachob (Ed.), *Power in the EFL classroom: critical pedagogy in the Middle East* (pp. 263-282). Newcastle upon Tyne: Cambridge Scholars.
- Jenkins, J. (2006). The spread of English as international language: A testing time for testers. *ELT Journal*, 60(1), 42-50.
- Jenkins, J. (2007). *English as a lingua franca: attitude and identity*. Oxford: Oxford University Press.
- Jenkins, J. (2014). *Global Englishes: a resources book for students 3*. London: Routledge.
- Johnston, B. (1999). Putting critical pedagogy in its place: A personal account. *TESOL Quarterly* 33(3), 557-565.
- Kachru, B. (1982). *The other tongue: English across cultures*. Urbana, IL: University of Illinois Press.
- Kachru, B. (1985). Standards, codification and sociolinguistic realism: the English language in the outer circle. In Q. Randolph, H. G. Widdowson, and Y. Cantù (Eds.), *English in the world: Teaching and learning the language and literatures* (pp. 11-30). Cambridge: Cambridge University Press for the British Council.
- Kachru, Y. & Larry E. S. (2008). *Cultures, contexts and world Englishes*. New York: Routledge.
- Kachru, Y. & Nelson, C. (2006). *World Englishes in Asian contexts* (Vol. 1). Hong Kong: Hong Kong University Press.
- Kandiah, T. (1998). Why new Englishes? In J. Foley (Ed.), *English in new cultural contexts: reflections from Singapore*, (pp. 1-40). Singapore: Oxford University Press.
- Kanpol, B. (1994). *Critical pedagogy: An introduction*. Bergin & Garvey.
http://opus.ipfw.edu/educat_facpubs/30
- Kanpol, B. (1997). *Issues and trends in critical pedagogy*. Cresskill, NJ: Hampton Press.
- Kemmis, S. and McTaggart, R. (1988). *The action research planner*. Waurin Ponds, Victoria, Australia: Deakin University.
- Kirkpatrick, A. (2007). *World Englishes: Implications for international communication and English language teaching*. Cambridge: Cambridge University Press.
- Kirkpatrick, A. (2012). *The Routledge handbook of world Englishes*. London: Routledge.
- Kirkpatrick, R. (2017) (Ed.), *English language education policy in the Middle East and North Africa*. Cham, Switzerland: Springer.
- Kubota, R. (2009). Critical multiculturalism and second language education. In B. Norton and K. Toohey (Eds.), *Critical pedagogies and language learning* (pp. 30-52). Cambridge: Cambridge Applied Linguistics.
- Kumaravadivelu, B. (1999). Critical classroom discourse analysis. *TESOL Quarterly* 33(3), 453-484.
- Lang, H. & Evans, D. (2006). *Models, strategies, and methods for effective teaching*. Boston: Pearson.

- Mauranen, A. (2012). *Exploring ELF: Academic English shaped by non-native speakers*. Cambridge: Cambridge University Press.
- McLaren, P. (1989). *Life in schools: An introduction to critical pedagogy in the foundations of education*. New York: Longman.
- McLaren, P. (2009). Critical pedagogy: A look at the major concepts. In A. Darder, M. Baltodano, and R. D. Torres (Eds.), *The Critical pedagogy reader* (pp. 61-83). New York: Routledge.
- McLoughlin, D. & Mynard J. (2009). An analysis of higher order thinking in online discussions. *Innovations in Education and Teaching International* 46(2), 147-60.
- Mahboob, A. (2013). Englishes of the Middle East: A focus on the Kingdom of Saudi Arabia. In R. Akbari, & C. Coombe (Eds.), *Middle East handbook of applied linguistics* (pp. 14-27), Dubai: TESOL Arabia Publications.
- Nasser, R. N. & Abouhedid, K. (2007). The academic 'Patras' of the Arab World: Creating a climate of academic apartheid. *PORTAL Journal of Multidisciplinary International Studies* 4(1).
- Nelson, C. L. (2011). *Intelligibility in World Englishes: Theory and application*. New York: Routledge.
- Norton, B. & Toohey, K. (2009). *Critical pedagogies and language learning*. Cambridge: Cambridge University Press.
- Patton, M. Q. (2002). *Qualitative research and evaluation methods*. Thousand Oaks, CA: Sage Publications.
- Pennycook, A. (1994). *The cultural politics of English as an international language*. London: Longman.
- Pennycook A. (2009): Global Englishes and Transcultural Flows. *Applied Linguistics* 30(2), 305-307.
- Pennycook A. (2007a). *Global Englishes and transcultural flows*. London: Routledge.
- Pennycook, A. (2007b). The myth of English as an international language. In S. Makoni & A. Pennycook (Eds.), *Disinvesting and reconstituting languages* (pp. 90-115). Clevedon: Multilingual Matters.
- Pennycook, A. (2001). *Critical applied linguistics: A critical introduction*. New York: Routledge.
- Phillipson, R. (2009). *Linguistic imperialism continued*. Oxford: Oxford University Press.
- Raddawi, R. (2011). Teaching critical thinking skills to Arab university students. In C. Gitsaki (Ed.), *Teaching and learning in the Arab World* (pp. 71-91). London: Peter Lang Publishing.
- Raddawi, R. (2015) (Ed.). *Intercultural communication with Arabs: Studies in educational, professional and societal contexts*, New York, Singapore: Springer
- Raddawi, R. & Troudi, S. (2012). Three Elements of Critical Pedagogy in ELT. *Achieving excellence through life skills education*, 73-82. Proceedings of the 18th International TESOL Arabia Conference Dubai: TESOL Arabia Publications.
- Randall, M. and Samimi, M. (2010). The status of English in Dubai. *English Today*, 26(1), 43-50.
- Romanowski, M. H. & Nasser, R. (2012). Critical thinking and Qatar's education for a new era: negotiating possibilities. *International Journal of Critical Pedagogy*, 4(1), 118-134.
- Salame, R. (2011). Higher education commitment to sustainability in the Arab states. In Granados Sanchez, Jesús & Escrigas, Cristina & Ferrer, Didac & Barcelo, Miquel & Cruz, Yazmin. (Eds). *Higher education's commitment to sustainability: From understanding to action*. GUNi Series on the social commitment of universities, 4, 69-74. Basingstoke: Palgrave Macmillan.
- Samovar, L. P. R. & McDaniel, E. (2012). *Intercultural communication: A reader* (13th edition). Boston: Cengage Learning.
- Sharifian, F. (2009). *English as an international language: Perspectives and pedagogical issues*. Bristol, UK: Multilingual Matters.
- Shor, I. (1992). *Critical teaching and everyday life*. Boston: South End Press.
- Sleeter, C. E. & Gran, C.A. (2008). *Making choices for multicultural education: Five approaches to RACE, CLASS*. John Wiley & Sons.
- Smith, E. (2011). Teaching critical reflection. *Teaching in higher education* 16(2), 211-223.
- Toohey, K. and Waterstone, B. (2009). Negotiating expertise in an action research community. In B. Norton and K. Toohey (Eds.), *Critical pedagogies and language learning* (pp. 291-310). UK: Cambridge Applied Linguistics.

- Troudi, S. (2005). Critical content and cultural knowledge for TESOL teachers. *Teacher Development* 9(1), 115-129.
- Troudi, S. & Alwan, F. (2010). Teachers' feelings during curriculum change in the United Arab Emirates: Opening Pandora's Box. *Teacher Development* 14(1), 107-121.
- UNESCO (2005). *Literacy and adult education in the Arab World*. Hamburg: UNESCO Institute for Education.
- Wachob, P. (2009) (Ed.). *Power in the EFL classroom: critical pedagogy in the Middle East*. Newcastle upon Tyne: Cambridge Scholars.
- Zuber S. O. (1996). *Action learning and action research: Song lines through interviews*. Rotterdam, Netherlands: Sense Publishers.

APPENDIX 1

EFL Teachers Questionnaire

The purpose of this questionnaire is to inquire about schools' and teachers' background and ways to introduce Critical Pedagogy in the curriculum

1. Information about the school

- a) Private
- b) Public

- a) Primary
- b) Secondary
- c) Both

Number of students in the school

Average number of students per class

2. Teacher's gender

- a) M
- b) F

3. Years of teaching experience

4. Have you ever heard the term critical pedagogy?

- a) Yes
- b) No

If yes describe briefly what you know about it in 2 lines

If no how do you wish to be informed about it? Through (circle that applies) :

- a) Readings
- b) Workshops
- c) Lectures
- d) Courses
- e) If other, please state

5. Do you contribute to the preparation of your course(s) syllabus?

- a) Yes
- b) No

If yes how?

6. Do you contribute in the course teaching material?

- a) Yes
- b) No

If yes, state how in 2 lines

If no, who is in charge?

- a) Ministry
 - b) Principal
 - c) Supervisor
 - d) Other, please state
-

7. Do the students have a say in the teaching material?

- a) Yes
- How?
-

- b) No

8. Do you accept criticism from students?

Yes

Such as?

No

Why?

9. What is your most frequent teaching methodology?

- a) Traditional by rote (memorizing) method
 - b) Experiential
 - c) Lecture format
 - d) Collaborative
 - e) Problem-solving
 - f) Others, what?
-

10. How much time do you allot for interaction with students in class?

- a) 5 mn
- b) 15mn
- c) most of the class

11. How would you define diversity in the classroom in no more than 2 lines?

12. Did you wish to integrate CP in the curriculum such as critical thinking, more interaction with students, change of roles, change of methodology, democracy in class, (this could be an interview question)

APPENDIX 2

EFL Teachers interview questions

- 1) What do you think of the textbook especially that it is new?
- 2) Do you have leeway in adding material to the set curriculum?
- 3) If yes, tell me about your role in the “hidden curriculum”? How do you fill the gap(s) if any whether it is related to content in general, culture or skills.
- 4) Do you tackle culturally sensitive issues in the classroom? Do you let the students talk about their personal lives and problems? How much time do you allot for students’ talk in class?
- 5) Do you relate classroom discussion to the outside world by tackling global issues for example.....? If yes how? If no why?

About the Authors

Rana Raddawi, PhD, is a Faculty member in the Department of Modern Languages at DePaul University, Chicago. She holds a PhD in Translation Studies from Sorbonne University in Paris, France. She is the editor of the book *Intercultural Communication with Arabs* (2015). Her research interests relate to cross-cultural studies, TESOL, translation, and critical pedagogy. She was a keynote speaker at several international conferences in the US and Middle East. She is a member on the Editorial Board of *Arab World English Journal* (AWEJ) and *International Journal of Bilingual & Multilingual Teachers of English*.

Salah Troudi a faculty member at the Graduate School of Education of the University of Exeter where he is the director the Doctorate in TESOL in Dubai and the supervisory coordinator of the PhD in TESOL. His teaching and research interests include language teacher education, critical issues in language education, language policy, curriculum development and evaluation, and classroom-based research. He has published articles in several language education journals and edited a number of books.

Teachers' Needs in the Advancement of Communicative Language Teaching (CLT) in Taiwan

Shu-Hsiu Huang*

Kaohsiung Medical University, Taiwan

Li-Chin Yang

Kaohsiung Medical University, Taiwan

Abstract

English language education in Taiwan has experienced a number of modifications over the past decade. The Ministry of Education (MOE) has initiated several reforms since 1994 to change the historical grammar-translation pedagogy into one emphasizing more communication. The purpose of these alterations was to increase learners' communication proficiency, so they could meet the new demands resulting from increased internationalization and globalization. Previous studies regarding Communicative Language Teaching (CLT) in many other EFL (English as a foreign language) settings and Taiwan have shown promising results, yet the process of implementing CLT has often been challenging. Many factors, such as large class sizes and parents' negative attitudes toward CLT, have been found to negatively influence teachers' willingness to implement CLT. Using a multi-methodological approach of quantitative surveys and qualitative interviews, this study aims to examine teachers' needs for better communication-oriented practices in the classrooms of Taiwan. In total, 75 teachers were surveyed. Based on their degree of willingness to participate, 15 of them were further contacted for more in-depth interviews. The findings of the study revealed that to make CLT more applicable, teachers demanded in-service training and assistance from native English-speaking teachers (NESTs). In addition, students' and parents' re-education to value the development of communication proficiency, and to gain more knowledge about CLT, and the support from school authorities, were considered crucial.

Keywords: communicative language teaching, CLT, multi-methodological approach, learning motivation, communicative competence

Introduction

In order to meet the demands of the fast-growing global economy, to recognize the status of English as an international communication tool, and to increase the number of people who can communicate effectively in English, many countries in Asia have reformed their English language education in the past two decades (Littlewood, 2007). New English syllabi aimed at "teaching English for effective and appropriate communication" have been released regularly in Singapore (Zhang, 2006). The MOE in Hong Kong introduced a policy of trilingualism (English, Cantonese, and Chinese), which emphasizes the development of oral proficiency (Law, 2003). English has been used as the medium for instruction at schools led by native English-speaking teachers (NESTs). Communicative Language Teaching (CLT) pedagogy was first introduced in the sixth curriculum (Yoon, 2004), and further reformed in the seventh curriculum in South Korea in 1997 to initiate new English education. It was promoted as a compulsory subject in every primary school. CLT was utilized in order to

* E-mail: huang4@umail.iu.edu; Tel: (7) 312 1101 ex 2116*311; Address: 100, Shih-Chuan 1st Road, Kaohsiung, 80708, Taiwan.

enhance students' interest in English communication (Jung & Norton, 2002). The CLT movement started in Japan in 1985, and has since then been maintained. Educational movements highlighting "practical English" were generated to develop learners' oral-aural communication ability in order to make them the functional users rather than knowledge accumulators (Aliponga et al., 2013; Bulter & Lino, 2005; Ohashi, 2015). Globalization has also been the driving force for the education reforms in China (Chang, 2006; Cheng, 1988; Garbe & Mahon, 1981; Hu, 2005; Hui, 2001; Zhu, 2003). Despite the resistance after the implementation of CLT at undergraduate and high schools, the State Education Development Commission in China authorized and issued three major English-reformed syllabi in 1992, 1993, and 1996 respectively to enforce the cultivation of communicative competence (CC) starting at the secondary level (Liao, 2004). The ambition was advanced in 2001 through the application of task-based instruction starting from the third grade (Hu, 2005; Hui, 2001). These communication reforms reached Taiwan from 1994 to 1995 when a new curriculum for junior and high schools was published with a clear objective asserting that "communication-orientedness was the principle of high school textbook compilation and classroom instruction" (Wang, 2002, p. 135). New textbooks featuring communicative activities have been used since 1999. English language learning was further lowered to the third grade in 2004 with the suggestion of an English-only policy and the adoption of "active and interactive" models via various teaching genres, realia, and other materials from diverse topics (Ministry of Education, 2014). With the pervasive implementation of CLT in the EFL context, future difficulties and alienation were first anticipated, especially among teachers. Therefore, a thorough and critical investigation on their current needs and interests will help alleviate this often chaotic atmosphere and better their CLT practices in the future. It is precisely this research gap that this study aims to fulfill.

It Matters to Communicate

The great debate of the constituents of communicative competence has been ongoing in the literature regarding second/foreign language (L2) education (Berns, 1990; Canale & Swain, 1980; Omaggio, 2001, Savignon, 1983). Such debate reveals the significance of communicative competence and its development using authentic CLT activities, which encourage learners' maximum communication in many different contexts (Wu, 2008). When engaging in CLT activities, learners learn by doing and testing each other's perceptions through interaction in a positive and non-threatening environment. They acquire the meaning and knowledge on their own (Hendrickson, 1991), and gain grammatical/sociolinguist/discourse/strategic competencies (Pokoma & Vasylieva, 2014). Ideas or concepts of the activities that matter to learners increase their motivation and of involvement. Simply having the knowledge of a language is not enough. Only through meaningful negotiations can students become efficient learners and administer what they have learned (Allwright, 1984; Antón, 1999; Englander, 2002; Oxford, 1997; Rao, 1996). Zhang (2006) confirmed that the ultimately successful language learning experiences were created through interactive and meaningful communication. Including communicative competence as one of his teaching principles, Brown (2007) has also argued that it is the "goal" of language classrooms and should be achieved by constant and extensive language use. Likewise, Littlewood (1981) summarized several contributions that CLT activities make. They provide learners with whole-task practices, which are structured to suit learners' ability levels and to help maintain (or enhance) learning motivation. Learners' motivation is more likely to continue (or even increase) if seeing how and what they have learned is successfully employed in communication with others. The more effective in communicating with others, the higher the motivation will be maintained, or it can even be enhanced. In addition, CLT activities allow natural learning. Much language learning takes place through natural processes when real communication is achieved, thus making either inside- or outside-communicative activities a key portion of the total learning process. Positive relationships are fostered when completing CLT activities, thus humanizing the classroom by turning it into a learning-supportive context (Chang, 2011a). Joyful atmospheres among teachers and students are produced, which consequently sustains students' efforts to learn.

CLT Practices and Resistance in Classrooms

The emergence of CLT in the 1970s, and the prosperity of western countries (for example the USA and the UK) in general, made more innovative teaching techniques available, such as content-based instruction, task-based teaching, and problem-based learning (Richards, 2006; Spada, 2007). Despite its wide acceptance, its sequential introduction into an eastern context has led to widespread dissatisfaction and resistance in many EFL (English as a foreign language) contexts (Ahmad & Rao, 2012; Barkhuizen, 1998; Celce-Murcia, Dörnyei, & Thurrell, 1997; Hu, 2002; Kumar & Kainth, 2015; Li, 1998; Lo, 2001; Shamim, 1996; Yu, 2001). In this context, cultural differences were often mentioned as problematic. As Ellis (1996) pointed out, process and meaning are what CLT emphasizes while content and forms are highly valued in EFL classrooms. The distinctive learning motivation held by learners is another reason. ESL (English as a second language) learners have an urgent need to communicate because of the existence of an English-speaking community beyond the classroom; in contrast, EFL learners lack such urgency. Often, English is merely a compulsory school subject or a “maybe” useful tool for job-hunting in the future (Sreehari, 2012). Echoing Ellis’s viewpoint, Lo (2001) asserts that many EFL practitioners, despite receiving a master’s or doctoral degree in the fields of language instruction in English-speaking countries, found it difficult to carry out ESL-based theories (CLT included) after returning to their home countries due to sociocultural variances. Littlewood’s (2007) review of several published papers has also revealed that factors associating with classroom management, students’ avoidance of English, minimal demands on English competence, and conflicts with educational values and traditions, have possibly constrained CLT in many Asian countries.

In Vietnam, Pham’s (2005, 2007) interviews with teachers found that there exist contextual conflicts, such as large class sizes, traditional examinations, personal beliefs of teachers’ and students’ roles (Iwashita & Ngoc, 2012), and students’ low motivation. Moreover, Zhang (2006) notes that the consequence of applying CLT marginalized grammar teaching, leading to failure in achieving the target outcome of teaching reading and writing in Singapore. In Thailand, Saengboon (2002) has confirmed that school administrative policy might sometimes impede CLT by grouping a large number of students (up to 100) with heterogeneous levels of English proficiency in class. Teachers were forced to use non-CLT-based textbooks to teach for tests.

Regarding South Korea, Jung and Norton’s (2002) observation suggests that many teachers complained that materials development and large class sizes functioned as a hindrance to CLT activities. Similarly, Li’s (1998) survey reveals that many constraints, caused by teachers themselves (deficiency in spoken English, low strategic and sociolinguistic competence, lack of training in CLT, few CLT re-training opportunities, misconceptions about CLT, insufficient time, and no expertise in CLT material development), by students (low English proficiency and motivation to advance communicative competence), by the educational system (large class sizes, grammar-based examinations, insufficient funding, and a lack of support from schools), and by CLT per se (CLT’s inadequate account of EFL teaching and a lack of effective and efficient evaluating instruments), negatively impact CLT (Kleinsasser & Sato, 1999; Sato, 2002). With regards to Japan, Kubota (2002) has revealed an unwelcome attitude held by teachers at public secondary schools toward NESTs. They, in effect, regarded CLT as a virus impeding students from intellectual and cultural virtue development.

In China, Burnaby and Sun’s (2007) study has suggested that many teachers believed that CLT is mainly applicable to those students that major in English. Zhu (2003) observed that Chinese students were strongly influenced by Confucian concepts. They were trained to be obedient, but not to challenge authority. Consequently, they tended to keep their opinions to themselves, and passively hid their ability as knowledge-receivers. Hu (2002) also noted that CLT tenets contradict Chinese culture in terms of their embodiment of opposite teaching philosophies. It advocated interactiveness, learner-centeredness, verbal activeness, independence, and individuality; whereas Chinese learning cultures asserted ancient epistemology, teacher dominance, mental activeness, receptiveness, and conformity. Likewise, Aldred and Miller’s (2000) investigation has pointed out that the active roles that learners play in CLT classrooms contradict the socio-cultural traditions of Hong Kong where students are supposed to be silent and avoid making mistakes by not raising their hands to ask or answer questions.

English is a compulsory subject and the only foreign language that is tested for all kinds of entrance examinations in Taiwan (Ho, 1998). Despite the importance of English and the popularity of CLT in classrooms,

due to insufficient and inadequate channels for CLT-related training (Chang, 2011b), and access to authentic materials (Kuo, 1995), low support from school administrators, and parents' demands for good test results and standardized answers to check their children's learning outcomes (Su, 2006), Wang (2002) observes that many Taiwanese EFL teachers tended to neglect the communicative activities compiled in the reformed-CLT textbooks. Instead, they tended to favor the traditional grammar translation method (GTM) for convenient and immediate learning outcomes, leading to learners' inclination toward memorization, grammar, reading, composition, and translation (Chen, 2001; Chung & Huang, 2009). Some parents even believed that the best teachers were the ones who had taught their children to score high on tests, instead of helping them gain valuable communication skills. Instilled education values and beliefs also play an important part. Many Taiwanese students are educated to maintain a more listener-centered standpoint within communication (Liu, 2005). A call-upon for opinion in class was beyond their expectation, and this could often result in a communication shut-down between teachers and learners (Babcock, 1993). Learning assessment was another problem. Much emphasis was still placed on the evaluation of written skills via standardized discrete-point tests (Wang, 2010). Thus, the development of new assessing techniques such as a portfolio or teacher's observation and recording of a learner's performance was obviously crucial.

The discussion regarding the practical difficulties associated with CLT in diverse EFL settings has revealed not only its significance, but also concerns over the consequent controversies and cultural appropriateness (Tanaka, 2009). Notwithstanding the extensive investigation on logistical problems, little attention has been paid to teachers' needs and interests. This study therefore aims to fill this obvious research gap.

Research Questions

Employing multiple data sources, this study attempts to understand EFL teachers' requests in order to advance their CLT practices in classrooms after its prevalence in Taiwan. The particular research questions addressed here were: (1) How have the teachers in Taiwan perceived CLT?, (2) How has CLT reflected in their teaching of English?, (3) To facilitate CLT, what support have they needed from school authorities, parents, and students?, and (4) What improvements could be made to better their CLT practice now and in the future?

Methods

Questionnaire and Survey

The methodological assumptions utilized in the study were derived primarily from Kleinsaser and Sato's (1990) work. With the objective of understanding teachers to better their CLT practice, we used a mixed-method approach since a quantitative approach provides a general (broad) view that controls statistically the bias, and external factors of the phenomena studied. Conversely, the multiple sources collected via qualitative approach provided a more complex (deep) understanding of the issue. A "Teachers' Needs for Better Communicative Language Teaching" questionnaire adapted from Li's (1998) and Rao's (1996) studies was administered. It was divided into three main parts. The first part contained questions regarding the interviewee's background, including age, educational degree, and years and grade level of teaching. The second part listed questions regarding teachers' general views on CLT and their actual classroom teaching. The final part included statements about the support and resources that the teachers anticipated. The surveys were conducted on a one-to-one basis either in Chinese or English, through either telephone or in person, to encourage cooperation and rapport creation for a potential interview to take place later (Dörnyei, 2003).

Interviews and Interview Questions

Interview is the best way of learning about people's interior experiences and how they perceive and interpret their perceptions, which in turn was affected by their thoughts and feelings about a particular event (Weiss, 1994). The narration and viewpoints elaborated by the interviewees are a convenient, yet faithful channel beyond any possible substantial boundary to get a glance of their world (English teaching in this study). This, again, provides reliable scientific explanations to understand the meaning of the particular phenomena described above. In this

study, almost all the interviews were conducted in Chinese through telephone, online, or in person to avoid possible limitations and miscommunication. Each interview lasted approximately from 30 to 60 minutes.

The whole process was audio-recorded and further translated and transcribed verbatim to ensure accuracy and richness of data for later analysis. The interview questions mainly consisted of three sections: (a) the major questions: eight open-ended descriptive questions from the themes that had emerged from the questionnaire (Brinkmann and Kavle, 2009; Karner & Warren, 2010), which were the teachers' general views on CLT, their actual classroom teaching, and the support and resources they anticipated (see Figure 1: the transformation of questionnaire themes into interview questions); (b) additional questions used to "explore particular themes, concepts, and ideals introduced by the conversational partner" (Rubin & Rubin, 2005, p. 136), such as "how do you usually teach vocabulary / sentence patterns / main text?; how do you think of MOE's supplementary policy for CLT?; do you think it applicable at the school where you teach?; how many students do you usually have in one class?; have you ever thought of ways to overcome difficulties?"; and (c) probing questions: used to clarify missing information or ambiguous concepts while keeping the discussion going, for example "what do you mean by....?; would you explain....?; please give me an example of how you....; can you say something more about...." (Biklen & Bogdan, 2003).

Following the "tree and branch model,"¹ the researcher asked all the main questions and then follow-up questions drawn from each interviewee's response to each main question. Occasionally, appropriate probes (repetition of particular words with questioning intonation, asking questions for more details, showing attention to encourage elaboration, and asking for an explanation or clarification) were used to ensure that we would obtain vivid, thick, deep, and detailed descriptions of the investigated phenomena (Rubin & Rubin, 2005).

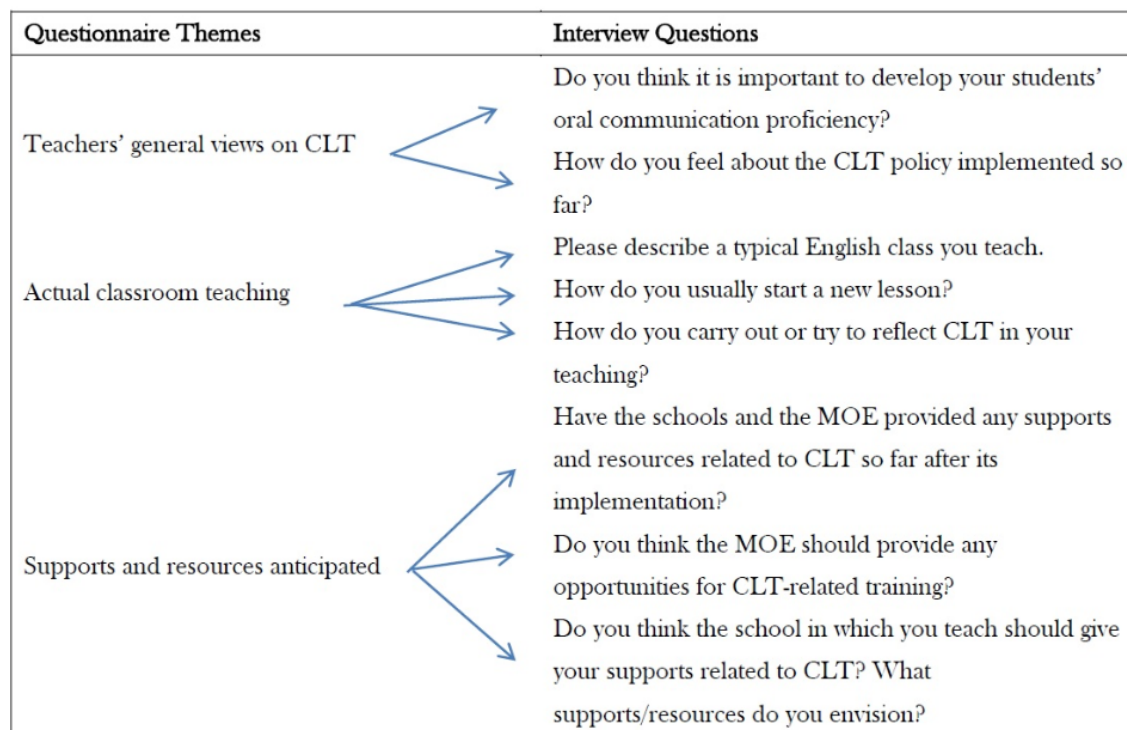


Figure 1. The Transformation of Questionnaire Themes into Interview Questions

Participants

The participants were selected randomly to “minimize the effects of any extraneous or subjective variables that might affect the outcome of the survey study” (Dörnyei, 2003, p. 73). In total, 75 English language teachers, which had been selected randomly from school websites,² were contacted and surveyed. The age of the participants ranged from 24 to 60, with the majority being in their 30’s. Altogether, 62% had less than 10 years’ of teaching experience, while 38% had more. Thirty-five (46.6%) had a bachelor’s degree and 40 (53.4%) held a master’s degree. Based on their willingness to participate, 15 were further interviewed (see Table 1). For confidentiality, pseudonyms were used throughout the paper.

Table 1
Teachers’ Background Information

Teacher	Name	Gender	Age	Teaching Years	Field of Study	Degree / Country
1	Fanny	F	52	30	TESOL	MA/TW
2	Jimmy	M	28	1.5	English Literature	MA/AU
3	Kelly	F	38	13	English	MA/USA
4	Lucy	F	34	11	English Ed.	MA/TW
5	Maggie	F	25	2	TESOL	BA/TW
6	Mandy	F	40	10	TESOL	MA/USA/UK
7	Mary	F	37	13	Social Study	BA/TW
8	Penny	F	31	5.5	Language Ed.	MA/UK
9	Sophie	F	41	17	Counseling	MA/TW
10	Sunny	F	30	6	Art and Media	BA/TW
11	Tiffany	F	37	15	English	BA/TW
12	Yuki	F	39	10	TESOL	MA/TW
13	Yvonne	F	36	11	Business Management	MA/UK
14	Wendy	F	40	10	English	MA/USA
15	Zoe	F	51	29	TESOL	BA/TW

Data Analysis

The collected survey data were analyzed using SPSS 20 to calculate the means and a total number of the participants’ responses, frequency, and percentage marking in each response for each statement. These quantitative data were sorted into three major categories based on the three themes found in the formation of interview questions and several sub-topics to integrate with qualitative data. The translation and transcription of interview data were the second phase of the data analysis. After translating and transcribing, concepts,³ themes,⁴ events,⁵ and topical marker⁶ were first identified, followed by a further and more thorough examination for clarification of unclear concepts and themes, and the synthesis of different events for better comprehension of the overall narratives. The final stage of the data analysis was coding. As Weiss (1994) states, “the idea in coding is to link what the respondents says in his or her interview to the concepts and categories that will appear in the report” (p. 154). Appropriate codes were derived primarily from the identified concepts, themes, events, and topical markers, or from the reviewed literature related to the issues under examination. The categories of codes that were applied include (1) activity codes: behavior occurring regularly (teachers’ description of their teaching situation); (2) event codes: specific activities that had occurred in the setting or the lives of the interviewees (the

mention of the difficulties encountered); and (3) strategy codes: methods, techniques, or other ways interviewees used to accomplish something (support and needs teachers asked to better their CLT implementation in schools).

Results

Teachers' Perceptions and Practice of CLT

Table 2

Teacher's General Views on Current CLT Policy, Report of Teaching Methods Used, and Needs to Improve Their CLT Practice

	Frequency (n = 75)
<u>Teacher's general view on current CLT policy</u>	
The development of students' English communication is important	100%
CLT policy is applicable	49.3%
CLT is impossible now or in the future	8%
<u>Teachers' report of teaching methods used</u>	
Grammar Translation Method (GTM)	48%
Communicative Language Teaching (CLT)	54.6%
Both GTM and CLT based on the actual class situation	87%
<u>Types of need</u>	
Making English our second language	49.3%
More funding for teaching equipment	58%
Chances of overseas studying programs in English-speaking countries	88%
Regular CLT-related training	94.6%
The establishment of more English Villages	49.3%
Modification of English textbooks	74.6%
The promotion of General English Proficiency Test (GEPT)	39%
More English-proficient training led by native English-speaking teachers (NESTs)	96%
CLT-related teaching demonstration conducted by "seed teachers"	89.3
Class-size reduction	94.6%
Students' re-grouping based on their English proficiency	70.6%
The addition of class meeting time	66.6%
Authority for English assessment	68%
The inclusion of oral proficiency into term exams	81%
Parents' re-education to better understand CLT and new possible assessment tools	88%
Students' education of the importance of communication-proficient development	94%
Co-teaching with NESTs	99%
The hiring of more teachers of overseas-studying experience	50.6%
Good interaction and regular conferences among teachers, parents and school administrators.	44.6%

Table 2 presents the percentages counted on items regarding teacher's general views on current CLT policy, actual classroom teaching, and needs for better CLT practice. It is clear that all of the surveyed and interviewed teachers asserted that the cultivation of students' oral communication ability was important and necessary, although almost only half of them (49.3%) believed that CLT policy was applicable. They agreed on CLT's tenets and practiced CLT activities if there was enough time. To fulfill CLT, most teachers (87%) had conducted

a semi-traditional teaching approach, or a “reconciling communicative approach” (Rao, 1996, p. 456), a mixture of the two main teaching approaches, CLT and GTM, regardless of various implemental conflicts.

Teachers’ Expressed Needs to Better their CLT Practice

To efficiently practice CLT, they indicated a need for (1) training to improve English proficiency, (2) opportunities for CLT-related workshops, (3) opportunities to work with native English-speaking teachers (NESTs), (4) raising students’ awareness of the importance of communication-proficient development, (5) increasing parents’ recognition of CLT, (6) class re-organization, and (7) textbook modification and testing-format alteration.

Training to improve English proficiency

English language proficiency has always been a concern to many non-native EFL teachers (Bulter, 2004). This concern has taken several forms. Some teachers were worried that their language skills were inadequate, unbalanced among four skills (Bulter, 2004; Chang, 2006; Li, 1998) and deficient in oral English (Brutt-Griffler & Samimy, 1999). Some thought that they lacked strategic and sociolinguistic competences (Li, 1998). Many felt that they were not equipped to teach in communicative manner (Anderson, 1993) or that they were unable to deal with students’ unforeseen needs (Littlewood, 2007). Consequently, they struggled to introduce communicative activities, or avoided English communication courses per se (Wada, 2002). Having similar concerns, 3 out of the 15 interviewed teachers argued that it was necessary to improve their English proficiency, as Mary noted, “I think in-service training is insufficient. I personally wish that every year we were given the chance to review or improve our English skills, just like in those classes we had when we were undergraduates.” The quality of training mattered to these teachers. Opportunities of more knowledge and better language enhancement were the most welcomed and motivated, as Yuki asserted, “If more advanced training is provided, I believe we will be happy to cooperate.” To improve their English proficiency, 88% of the surveyed teachers embraced the chances of overseas study programs in English-speaking countries, and 96% of them welcomed the training classes led by NESTs.

Wherever the location, an ideal language training workshop should not mainly be lecture-based and teacher-centered, especially if it is for the improvement of oral proficiency. According to Fanny, well-organized and highly-proficient training sessions should be structured and centered on “communication.” They are not just language classes per se. Indeed, they offer an opportunity for teachers to gain further knowledge in their subject area. They are also channels that allow teachers to critically analyze not only each other’s language proficiency, but also the identity of English speakers and teachers:

It doesn’t need to be long... maybe one or two hours a day and six days a week during the summer or winter break. We get together to study with the native speakers. It can be in the format of a teacher study group (TSG). Just let us get together and chat with the native speakers. Through activities, we not only learn the content-area knowledge, but also get the chance to communicate. The teachers with stronger communication skills can be the stimulus to bring positive impetus to those who are weaker. We learn from each other. By so doing, I believe what is achieved is not just the main function of the classes. Teachers will also be empowered. (Fanny)

These teachers’ demands reflected Cullen’s (1994) suggestions about non-native English-speaking teachers (NNESTs), indicating a fact that NNESTs “need to improve their own command of the language so that they can use it more fluently, and above all, more confidently in the classroom” (p. 164). Indeed, in-service language training is essential, especially for veteran teachers, like Fanny and Mary who have been in the field for more than 10 years, but have lost their skills owing to the lack of utilization and practice.

Opportunities for CLT-related workshops

To promote CLT policy, the MOE and many textbook publishers in Taiwan offer numerous on-the-job training workshops, although according to several of the interviewed teachers (3 out of 15), many of these occasions have been fruitless, simply because they are usually the announcement and advertisement of either testing plan. “I feel the workshops for high school teachers are not CLT-related at all. The topics discussed are usually about the

General English Proficiency Test (GEPT),” Lucy commented. Or, they are about a certain education policy, as Mary stated, “Teacher training, I think, does not focus on teachers’ needs. Often I attend training that is focused only on the promotion of education policy.” Echoing such dissatisfaction, 94.6% of the surveyed teachers requested more appropriate CLT-related training opportunities and 89.3% of them asked for the teaching demonstration conducted by CLT seed teachers, as Meggie noted, “I want to know if any organizations or schools carry out CLT successfully, how they make CLT happen, and what their students’ learning outcomes are. Their visiting is highly welcomed, but I do believe their suggestions are valuable.”

Despite the considerable amount of knowledge about language instruction they obtained before their on-site service, teachers have inevitably encountered problems during their actual classroom practices, especially regarding new methods, like CLT, which was part of university methods courses, but was never put in practice. It is natural that teachers doubt their abilities, and the feasibility of practicing CLT without sustainable training. Regular and constructive workshops, even if they are in the form of lectures about CLT or other seminars, allow teachers to share their teaching experiences or encountered problems, thus helping them to deal with new innovations and changes in methodology (Koosha & Yakhabi, 2013), as Fanny suggested, “No matter what the topic is... about teaching materials, games, chants, or songs. I feel that each workshop is an opportunity to improve myself.”

Opportunities to work with native English-speaking teachers (NESTs)

Facing the situation of insufficient English-improving opportunities, and the urgency to teach communicatively, these teachers (93%) have sought the assistance of NESTs through co-teaching. They value co-teaching chances because NESTs compensate for their oral-deficient shortcomings and help them solve this pressing need. “It is important to enhance students’ oral abilities. That’s why I have applied for a NEST. The MOE will offer me one and he/she is coming next semester. I hope he/she can teach the entire class in English,” said Kelly. Considering the communication deficiency among her colleagues and herself, Mandy expressed a similar opinion about the invitation of NESTs. She believed NESTs’ presence is one of the elements for feasible creation of total English immersion for herself, her colleagues, and students. All of them will benefit from such situations because of the instant and continuous English practices and reinforcement accompanied, as she pointed out, “Not all English teachers at my school have spontaneous conversation abilities. Having a full-time NEST here, students and I will be forced to speak English to him/her. A more spontaneous English learning environment will be created then.”

The teachers’ opinions regarding the construction of a more spontaneous language learning environment went beyond the classroom. In their words, if the use of English stayed at schools only, CLT’s efficacy would be limited, and learning would be constrained. For example, to extend this argument, according to Zoe, a new language policy legislating English as one of our official languages is needed: “If they treat English as a second language, it won’t be learned merely in schools. When it is needed at each corner in Taiwan, students will be forced to learn and to use it communicatively.” Notwithstanding the good intention embedded, not many teachers supported the idea, for only 49.3% of the surveyed teachers agreed that making English the second language would make CLT more feasible, too.

Raising students’ awareness of the importance of communication-proficient development

As mentioned previously, for most school learners in Taiwan (as in many Asian countries), English is an essential subject and the only foreign language in both junior and high schools. It may be an important subject that is tested regularly (even daily) at schools, but not an imminent one that significantly affects the students’ daily lives (Ellis, 1996). After all, the communities they live do not depend on English as their main medium of communication. Therefore, students lack integrative motivation for improvement. As a result, English learning becomes only a “need-to-do” routine, instead of a “want-to-do” passion, not to mention the cultivation of communication proficiency.

In the study, 94% of the surveyed teachers agreed that learners’ cultivation on the importance of learning English and developing communication ability could positively boost CLT. In the interview, Wendy further argued for the necessity of an uprooted instillation, an action to alter learners’ belief, and their attitudes toward English learning:

We should do something to let our students know that English is really important, especially in the development of communication. Otherwise, they will always think that English is only a school subject. They must know that the English they learn is a usable tool for international communication, but not just for tests.

Increasing parents' recognition of CLT

The side-effect of viewing English as a school subject is the parents' over-emphasis on their children's testing outcomes, that is, the higher the test score, the stronger the satisfaction. This potentially influences teachers' selection of a particular teaching method. Namely, teachers would simply choose the methods (usually analytical, but not communicative ones) best fit to pursue high scores to ease and please the parents. To win the parents' cooperation, 88% of the teachers suggested that parents should be re-educated to better understand CLT, and to know new possible assessment tools for CLT. Thus, regular parent-teacher meeting should be compulsory since they are occasions that permit teachers to decode and explain to the parents the essence and benefit of learning a language through CLT, as Yuki confirmed, "Parent-teacher meetings...one or two times, especially in the beginning of each semester is to let parents know at least what CLT is and that the activities conducted are not just games. They are meaningful and functional."

Language learning is a lifelong journey requiring the assistance of skilled and beneficial guides (good learning and teaching methods, such as CLT). Teachers pointed toward the right direction. Students (the travelers) decide the depth and width of their trip. Once in a while, parents join in, take the lead, and even redirect without the awareness of a possible detour. Therefore, without a doubt, a pre-traveling education and clear instructions are necessary in order to maintain and ensure the joy of the journey.

Class re-organization

Large class sizes (ranging from 30 to 100) and grouping of students in the same class based on heterogeneous English proficiency (Jung & Norton, 2002; Littlewood, 2007; Pham, 2005, 2007; Saengboon, 2002; Yu, 2001) have been an issue since the CLT's prevalence in many EFL settings. Such conditions resulted in classroom-management problems, for example, unbalance oral practice opportunities or teachers' failure to pay attention to low-performing students. Having encountered similar problems, most surveyed teachers (94.6%) favored a reduction in class size. Brown (2007) writes that a class of 12-15 students is ideal because it is "large enough to provide diversity and student interaction and small enough to give students plenty of opportunity to participate and to give individual attention" (p. 245). The figure for many of our interviewee teachers was 20. The number was perfect enough to facilitate CLT activities, as Yvonne asserted, "I think 20 students should be good. This makes direct communication among us possible and easier."

Taking only the quantity into consideration was not good enough. Quality counts, too. In the survey, teachers (70.6%) claimed that it was necessary to group students based on their English proficiency. Ideally, students of equivalent English proficiency should be grouped in the same class. They learn better and advance to a higher level quicker. Doubtless, such an arrangement alleviates the teacher's role and promotes instant communication:

Students should be grouped as basic, intermediate, or advanced level. I remember when I learned English in cram schools. My classmates were students who had the same proficiency as I did. I think teaching the class with students of similar level of proficiency is easier. When I communicate with one student, I communicate with the rest simultaneously. (Jimmy)

Textbook modification and testing-format alteration

Teaching materials are the best and the most convenient resources that allow teachers access to various activities and ideas to facilitate instruction and promote communication among students (Pan, 2013). Unfortunately, according to Jung & Norton (2002), they are usually the most difficult part for CLT's realization. In Taiwan and many Asian countries, textbooks compiled after CLT policy usually followed a similar pattern, starting with a set of CLT-featured guidelines posted by MOE. Publishers needed to produce, review, and revise the textbooks accordingly. Teachers were usually put at the end of this assembly line. If they were lucky enough, they were

permitted to make a “personal” choice among these licensed and published textbooks. If not, they were just forced to use the ones that were selected by school authorities who were usually not English teachers (Richards, 1993). In fact, teachers’ needs and voices are seldom taken into consideration during the compilation process. Compared to the previous non-CLT textbooks, these so-called CLT-featured editions despite being improved and localized to include the local culture (Nguyen, 2005) still did not meet many EFL teachers’ requirements. In short, they were simply not CLT enough (Wu, 2001). Being the first-hand users of these materials, 74.6% of the surveyed teachers asserted that there is a need for a textbook modification. According to Jimmy, a good-enough recompilation must be integrated and conducted by a state-owned organization, but not by a private publisher. This is a huge project. It should be monitored and supervised closely and unanimously. The re-compiled textbooks must be classified. The books of each level should list information suitable specifically to the students of that level. New information and topics should be added level by level. The degree of difficulty and complexity increase with the advancement.

MOE should appoint our National Institute for Compilation and Translation (NICT)⁷ to carry out textbook recompilation. I don’t think private textbook publishers can accomplish this work. The content of each level should be proposed clearly. For example, for the basic level, learners should learn the 26 alphabets. The acknowledgement of vocabulary related to daily things and colors is necessary. At level two, grammar rules should be taught. (Jimmy)

The discussion on textbook recompilation was followed by the request for testing-format alteration. As revealed in the survey, 81% of the teachers called for a shift of test format from paper-baseness (usually the test of reading and writing skills) to the inclusion of oral proficiency. Without appropriate alteration, whatever effort put in the educational reforms or the improvement of students’ communicative competence would be fruitless, as Penny commented, “Unless they change the direction of big examinations by including the assessment of oral proficiency. I mean, I believe teachers spend time teaching communicatively, but I believe they spend more time teaching reading, writing, and grammar for good test results.”

Discussion

Many of the CLT-related constraints that occur in Taiwan are also common in other EFL contexts. Thus, teachers’ needs may be close to identical there. The following discussion applies to the issues proposed by the teachers in Taiwan. This also extends to other contexts where CLT prevails.

Implications for Teacher Training Programs

Cullen (1994) points out that a NNEST training session without taking English language into consideration fails to meet the EFL practitioners’ needs and expectations. Many scholars (Brinton, Kamhi-Stein, & Snow, 2006; Edge, 1988; Ellis, 1986) have argued for the necessity of putting a language component into training workshops. To do so, Edge (1998) has suggested that teachers (as trainers) and trainees (as teacher students) should both be viewed as “language users,” because besides methodology, a “language improvement” component should always be added into EFL teacher training sessions. To develop a CLT and an English proficiency improvement course for teachers, the principles of maximizing English exposure, offering opportunities to share knowledge and ideas, and basing training on task-based and inductive arrangements (for more details, see Brinton, Kamhi-Stein, & Snow, 2006; Britten 1988; Hayes, 1995) are essential. The training techniques utilized to achieve the above principles include (1) lectures / demonstration: trainers’ provision of straight input, raw materials and demonstration of particular techniques; (2) elicitation: using question-and-answer technique to try to draw out teacher trainees’ opinions on specific topics; (3) workshops: trainees’ individual or group work to prepare materials, teaching aids and lesson plans; (4) whole discussion: a general discussion of any topics with all trainees together; (5) group/pair discussion: teacher trainees’ work in group or pairs using an activity sheet; (6) panel discussion: asking a group of trainees to form a panel. The rest should prepare and ask a number of questions relevant to chosen issues. The trainer will act as chairperson of the panel or the facilitator of panel discussion

(based on Ellis, 1986, p. 95). Workshops and training adhering to the above-mentioned principles and techniques are not just about training per se. In fact, they are CLT customized learning experiences themselves. They gratify teachers' eagerness to gain appropriate language training. Through discussion activities, teachers practice English by sharing their opinions and knowledge with one another. Simultaneously, they can critically analyze and make a contribution to each other's language learning techniques and teaching progress.

Tips for Working with a Large Class of Mixed-level Learners

Although teachers' wishes reflect their urgent needs to advance their CLT practice, in reality, it is the school authorities who make administrative decisions (such as those regarding the curriculum, class size, and schedule, students' performance tests, budgets, and even some teaching materials). Unfortunately, they are also usually the ones who jeopardize CLT (Saengboon, 2002). Under the circumstance when class-size reduction and re-arrangement is impossible, the tips on working with large classes of heterogeneous-proficient students with limited time will possibly ease teachers' anxiety. Teachers should make each student feel that he/she is important by remembering their names. By so doing, teachers show their awareness of and respect for students as individuals, and put a value on each student's presence and contribution in class. In this sense, the maximization of English practice opportunities using pair- or group-activities is vital. Teachers may feel chaotic during the practice, but this may be the only way to give students time to practice their English. To balance students' diverse English proficiencies and skills when pairing or grouping students, teachers can occasionally place students of similar proficiency and skills in the same pair/group, or sometimes place them of different levels within the same pair/group. To ensure equal and selective monitoring and feedback, teachers can spend the bulk of their time on just a small number of groups during students' collaborative work, and the groups that are not monitored are invited to report results to the rest of the class. It is the teachers' obligation to make sure that each student has a chance to talk. A constant and updated record of who has or has not been called to talk during a lesson or a whole semester would be the best way to ensure a random and equal roll call (Brown, 2007; Wharton & Race, 1999).

Alternative Assessment: Assessing Students' Communicative Competence

To develop a suitable tool for oral-proficient assessments can be a challenge for many EFL teachers. Integrated Performance Assessment (IPA), a tool designed to meet the American Council on the Teaching of Foreign Language (ACTFL) proficiency guidelines,⁸ provides a good model to meet the needs for valid and reliable assessment to determine students' competences (Adair-Hauck et al., 2006). Under IPA's framework, language performances can be divided into three types of task, and each can further be tailored to fit the learner at specific levels:

- I. Interpretive communication task: at this phase, students will be required to read or listen to an authentic text (weather forecast, commercial, letter, short story, or film) and reiterate the text or answer questions relevant to the text, either in a spoken or written form.
- II. Interpersonal communication task: performing task at this phase requires dual interpretation and negotiation between two learners. They may be given information that the other person may not have about a particular topic. They, therefore, need to exchange and negotiate to obtain the missing information.
- III. Presentational task: the activity used at this phase is one-way. It requires learners' to give presentation on a given topic to a specific group of audience (teachers, classmates, or parents), such as giving a speech on an event or introducing things that they have created.

If the IPA model is relevant to oral proficiency assessment, portfolio assessment would be an excellent tool to assess other language skills (mostly reading and writing). Portfolio assessment, in the words of Moore (1994), is a purposeful "collection of evidence used by the teacher and students to monitor the growth of the students' knowledge of content, use of strategies, and attitudes toward the accomplishment of goals in an organized and

systematic way” (p. 170). What goes into students’ portfolios is determined by teachers or students themselves, depending on the learning goals and achievement presentation (Pierce & O’Malley, 1992). Farr & Tone (1998) provided some general guidelines for this type of collection, and Chen (2000) suggested that the items should include: (1) sets of papers reflecting students’ cognitive learning processes, such as rough drafts from different genres (letters, essays reports, and personal narratives), polished products, and a learning log; (2) reaction and reflection papers showing feelings, problem solving and critical thinking, and a dialog journal reflecting numerous purposes for writing and reading; (3) books or other reading materials, selected classroom tests, audiotapes of students’ reading that display the examples of what students have accomplished and read; and (4) art, audio/video recordings, and photographs that exhibit the skills that the students can master.

These two types of assessment (performance and portfolio assessment), according to Pierce & O’Malley (1992), complement each other in terms of evaluating students’ overall performance and bring a “washback effect” on teaching. They emphasize that performance and portfolio assessment:

Together represent authentic assessment, continuous assessment of student progress, possibilities for integrating assessment with instruction, assessment of learning process and higher-order thinking skills, and a collaborative approach to assessment that enables teachers and students to interact in the teaching/learning process. (Pierce & O’Malley, p. 2, 1992)

Conclusion

English has been recognized as an international language of communication for a long time now. In an attempt to increase the number of people who can communicate efficiently in English for the reason of fast-growing economic globalization, many Asian countries have launched a series of English education reforms (Littlewood, 2007). New syllabi featuring CLT to enhance students’ interests in English communication, and to develop learners effective and appropriate communicative competencies (proficiency) have been introduced (Hu, 2005; Jung & Norton, 2002; Yoon, 2004; Zhang, 2006). English was suggested as the major instrument for instruction. Communication-oriented classes that had been implemented at the undergraduate and senior high schools were lowered and practiced starting from the junior high schools, even the third grade (Hu, 2005; Hui, 2001; Liao, 2004; Wang, 2002). The publication of a “fresh” curriculum emphasizing CLT-related characteristics initiated a new era for English language education in Taiwan. The textbooks featuring CLT were subsequently compiled and used, which was accompanied by the advocacy of English-only policy (Ministry of Education, 2014; Wang, 2002). Doubtless, many of the actions have further cemented the dominance of English in classrooms, and stressed the importance of acquiring English language skills through CLT. Despite these ambitions and good intentions, without sufficient support and thorough preparation, the results have been somewhat disappointing.

In fact, prior to its prevalence in Asian classrooms, CLT has long been questioned due to its failure in fitting into the social-cultures of Asia in practice. Problems and doubts have been reported regarding its applicability (Barkhuizen, 1998; Celce-Murcia, Dörnyei, & Thurrell, 1997; Hu, 2002; Li, 1998; Lo, 2001; Shamim, 1996; Yu, 2001). Simple factors, such as learning motivation, values, and beliefs about students’ and teachers’ roles in classroom, as well as teaching philosophies potentially impaired CLT’s implementation efficacy (Aldred & Miller, 2000; Ellis, 1996; Lo, 2001; Zhu, 2003). The policies by school authorities aggravated the “alien problems” (Pham, 2005, 2007; Zhang, 2006). Despite the stated goals to develop communicative skills and the top-down adoption of CLT, administrators have tended to put a large number of students with various-degree of English proficiency in one class (Li, 1998; Saengboon, 2002), and ask teachers to use non-CLT-oriented textbooks to teach for paper-based examination only. Most importantly, parents’ high expectations of good test results have forced teachers to stick to traditional GTM methods (Su, 2006; Wang, 2002).

The English-education reforms in Asia, the discussion on the impact of CLT-related activities in language acquisition, and many of the previous descriptions about practical resistance reveal that CLT is appreciably questioned by teachers. The cultivation of students’ communicative competencies was undeniably necessary. Hence, it is worthwhile to further critically examine the issue in order to understand more about the teachers’ needs so as to more efficiently practice their CLT teaching. The results of this study reflect many of the problems

that teachers meet in EFL settings. In this study, the participants confirmed that to better their CLT practice, they needed training to improve their communicative proficiencies and workshop to gain more knowledge about CLT (Chang, 2011b; Li, 1998). For teachers without access to sufficient training, the chance to co-teach with NESTs became valuable. According to the interviewees, student-parent cooperation was also crucial. Teachers contended that there was a need to recognize English as a communication tool (but not just a school subject), and the development of communicative competence (Chen, 2001; Chung & Huang, 2009; Li, 1998; Su, 2006; Wang, 2010). They also expressed that classes should be re-organized by reducing student numbers (to 20 at most), and that students should be re-grouped according to their level of proficiency (Jung & Norton, 2002; Li, 1998; Pham, 2005, 2007; Saengboon, 2002). It was also suggested that textbooks should be recompiled and stratified to fit each level of study (Kuo, 1995; Li, 1998; Saengboon, 2002), while oral-proficient assessment was believed to be necessary to facilitate successful CLT (Li, 1998; Wang, 2010). Scholars who are interested in the continuous exploration of similar topics may consider EFL classroom observation to examine teachers' actual issues in using CLT.

Notes

- ¹ In the tree and branch model, "the interview is likened to tree with the trunk as the research problems and the branches as the main research questions. Each deals with a separate but more or less equal concern. In the interviews, the researcher would try to ask all the main questions and then the follow-ups to obtain the same degree of depth, detail, richness and nuance" (Rubin & Rubin, 2005, p. 145).
- ² For parents' better understanding of each school's rationale and policy, teachers' teaching and students' performances, every school in Taiwan has established its own website. The website (in both Chinese and English) lists the school's phone number, history, syllabus for each class, teachers' names, and the subjects taught. The following is an example of this type of website in Chinese: <http://www.jges.mlc.edu.tw/woops/html/> and in English: <http://163.19.160.248/~eng/ex1/index.html/>.
- ³ A concept is a word or term that represents an idea important to the research problem.
- ⁴ Themes: are summary statements and explanations of what is going on.
- ⁵ Events: are occurrences that have taken place.
- ⁶ Topical markers: are names of places, people, organizations, pets, numbers...and so on (Rubin & Rubin, 2005, p. 207)
- ⁷ National Institute for Compilation and Translation (NICT): is the highest agency in Taiwan for the compiling and translating of textbooks for various subjects and grade levels. Their compilation and translation also include works of academia and the culture of Taiwan. For more details, see <http://www.nict.gov.tw/en/>.
- ⁸ The American Council on the Teaching of Foreign Language (ACTFL) proficiency guidelines: are standards developed by ACTFL to serve as a direct reference when deciding learners' target language proficiency for teachers of the foreign languages taught in the United States (USA) (for more information, please visit <http://www.actfl.org/i4a/pages/index.cfm?pageid=1>).

References

- Adair-Hauck, B., Glisan, E. W., Koda, K., Sandrock, P., & Swender, E. B. (2006). The integrated performance assessment (IPA): connecting assessment to instruction and learning. *Foreign Language Annals*, 39(3), 359-382.
- Ahmad, S., & Rao, C. (2012). Does it work? Implementing communicative language teaching approach in EFL context. *Journal of Education and Policies*, 3(2), 28-35.
- Allwright, R. L. (1984). The importance of interaction in classroom language learning. *Applied Linguistics*, 5(2), 156-171.
- Aldred, D., & Miller, L. (2000). Student teachers' perceptions about communicative language methods. *RELC Journal*, 31(1), 1-20.
- Aliponga, J., Ando, S., Gamble, C., Koshiyama, Y., Wilkins, M., & Yoshida, K., (2013). University students' beliefs, perceptions, attitudes toward communicative language teaching. *MEXTESOL Journal*, 37(2), 1-9.

- Anderson, J. (1993). Is a communicative approach practical for teaching English in China? Pros and cons. *System*, 21(4), 471-480.
- Antón, M. (1999). The discourse of a learner-centered classroom: Sociocultural perspectives on teacher-learner interaction in the second-language classroom. *The Modern Language Journal*, 83(iii), 303-318.
- Babcock, S. F. (1993). The significance of cultural influences within the ESL/EFL classroom: A Taiwan experience. *ERIC, ED375681*, 1-29.
- Barkhuizen, G. P. (1998). Discovering learners' perceptions of ESL classroom teaching/learning activities in a South African context. *TESOL Quarterly*, 32(1), 85-108.
- Berns, M. S. (1990). Contexts of competence: Social and cultural considerations in communicative language teaching. New York: Springer
- Biklen, R. C., & Bogdan, S. K. (2003). *Qualitative research for education: An introduction to theories and methods* (4th ed.). Pearson Education Group, Inc.
- Britten, D. (1988). Three stages in teacher training. *ELT Journal*, 42(1), 3-8.
- Brinkmann, S., & Kavle, S. (2009). *Interviews: Learning the craft of qualitative research interviewing*. SAGE Publications: London.
- Brinton, D. M., Kamhi-Stein, L. D., & Snow, A. M. (2006). Teacher training for English as a lingua Franca. *Annual Review of Applied Linguistics*, 26, 161-281.
- Brown, H. D. (2007). *Teaching by principles: An interactive approach to language pedagogy* (3rd ed.). New York: Pearson Education, Inc.
- Brutt-Griffler, J., & Samimy, K. K. (1999). Revisiting the colonial in the postcolonial: critical praxis for nonnative-English-speaking teachers in a TESOL program. *TESOL Quarterly*, 33(3), 413-431.
- Bulter, Y. G. (2004). What level of English proficiency do elementary school teachers need to attain to teach EFL? case studies from Korea, Taiwan, and Japan. *TESOL Quarterly*, 38(2), 245-277.
- Bulter, Y. G., & Lino, M. (2005). Current Japanese reforms in English language education: The 2003 "action plan". *Language Policy*, 4, 25-45.
- Burnaby, B., & Sun, Y. (2007). Chinese teachers' views of western language teaching: context informs paradigms. *TESOL Quarterly*, 23(2), 219-238.
- Canale, M., & Swain, M. (1980). Theoretical bases of communicative approaches to second language teaching and testing. *Applied Linguistic*, 1(1), 1-47.
- Celce-Murcia, M., Dörnyei, Z., & Thurrell, S. (1997). Direct approaches in L2 instruction: a turning point in communicative language teaching? *TESOL Quarterly*, 31(1), 141-152.
- Chang, J. (2006). Globalization and English in Chinese higher education. *World Englishes*, 25(3/4), 513-525.
- Chang, M. (2011a). EFL teachers' attitudes toward communicative language teaching in Taiwanese college. *Asian EFL Journal Professional Teaching Articles*, 53, 17-34.
- Chang, M. (2011b). *Factors affecting the implementation of communicative language teaching in Taiwanese college English classes*. *English Language Teaching*, 4(2), 3-12.
- Chen, A. H. (2001). Balancing communicative language teaching and grammar in Taiwan. *Hsiuping Journal of Humanities and social sciences*, 1, 13-24.
- Chen, L. M. (2000). Portfolios in second and foreign language classroom assessment. *English Teaching and Learning*, 25, 28-41.
- Cheng, H. P. (1988). Teaching English as a foreign language in China. *TESL Canada Journal*, 5(2), 88-93.
- Chung, I. F., & Huang, Y. C. (2009). The implementation of communicative language teaching: An investigation of students' viewpoints. *Asia-Pacific Education Researcher*, 18(1), 67-78.
- Cullen, R. (1994). Incorporating a language improvement component in teacher training programs. *ELT Journal*, 48(2), 162-172.
- Dörnyei, Z. (2003). *Questionnaires in second language research: Construction, administration and processing*. New Jersey: Lawrence Erlbaum Associates, Publishers.
- Edge, J. (1988). Applying linguistics in English language teacher training for speakers of other languages. *ELT Journal*, 42(1), 9-13.
- Ellis, R. (1986). Activities and procedures for teacher training. *ELT Journal*, 40(2), pp. 91-99.

- Ellis, G. (1996). How culturally appropriate is the communicative approach? *ELT Journal*, 50(3), 213-219.
- Englander, K. (2002). Real life problem solving: A collaborative learning activity. *English Teaching Forum*, 8-11.
- Farr, R., & Tone, B. (1998). *Portfolio and performance assessment: Helping students evaluate their progress as readers and writers*. (2nd ed.). New York: Harcourt Brace College Publishers.
- Garbe, W., & Mahon, D. (1981). Comments on methodology-oriented teacher training programs in China, *TESOL Quarterly*, 15(2), 207-209.
- Hayes, D. (1995). In-service teacher development: Some basic principles. *ELT Journal*, 49(3), 252-261.
- Hendrickson, J. M. (1991). On communicative language teaching. *Hispania*, 74(1), 197-198.
- Ho, M. C. (1998). Culture studies and motivation in foreign and second language learning in Taiwan, *Language, Culture and Curriculum*, 11(2), 165-182.
- Hu, G. (2002). Potential cultural resistance to pedagogical imports: the case of communicative language teaching in China. *Language Culture and Curriculum*, 15(2), 93-105.
- Hu, G. (2005). English language education in China: policies, progress and problems. *Language Policy*, 4, 5-24.
- Hui, D. (2001). The globalization of the English language: reflections on the teaching of English in China. *International Education Journal*, 2(4), 126-133.
- Iwashita, N., & Ngoc, K. M. (2012). A comparison of learners' and teachers' attitudes toward communicative language teaching at two universities in Vietnam. *University of Sydney Papers in TESOL*, 7, 25-49.
- Jung, S. K., & Norton, B. (2002). Language planning in Korea: the new elementary English program. In Tollefson, J.W. (Ed.), *Language policies in education: Critical issues* (pp. 245-265). New Jersey: Lawrence Erlbaum Associates, Publishers.
- Karner, T. X., & Warren, C. A. B. (2010). *Discovering qualitative methods: Field research, interviews, and analysis* (2nd ed.). New York: Oxford University Press.
- Kleinsasser, R. C., & Sato, K. (1999). Communicative language teaching (CLT): practical understanding. *The Modern Language Journal*, 83, 494-517.
- Kosha, M., & Yakhabi, M. (2013). Problems associated with the use of communicative language teaching in EFL contexts and possible solutions. *International Journal of Foreign Language Teaching and Research*, 1(2), 63-76.
- Kubota, R. (2002). The impact of globalization on language teaching in Japan. In Block, D. & Cameron, D. (Ed.). *Globalization and language teaching* (pp. 13-28). London and New York: Routledge.
- Kumar, M., & Kainth, M. K. (2015). Learners' perception of communicative language teaching in the engineering institute in Punjab. *International Journal of Technical Research and Applications*, 18, 20-25.
- Kuo, H. J. (1995). The (in)appropriateness and (in)effectiveness of importing communicative language teaching to Taiwan, *University of Hawai'i Working Papers in ESL*, 13(2), 21-47.
- Law, W. W. (2003). Globalization as both threat and opportunity for the Hong Kong teaching profession. *Journal of Educational Change*, 4, 147-179.
- Li, D. (1998). "It's always more difficult than you plan and imagine": teachers' perceived difficulties in introducing the communicative approach in South Korea. *TESOL Quarterly*, 32(4), 667-703.
- Liao, X. (2004). The need of communicative language teaching in China. *ELT Journal*, 58(3), 270-273.
- Littlewood, W. (1981). *Communicative language teaching: An introduction*. New York: Cambridge University.
- Littlewood, W. (2007). Communicative and task-based language teaching in East Asian classrooms. *Language Teaching*, 40, 243-249.
- Liu, G. Z. (2005). The trend and challenge for teaching EFL at Taiwanese Universities. *RELC Journal*, 36(2), 211-221.
- Lo, Y. G. (2001). From Western Language Teacher Education Programs to Eastern EFL Classrooms: Alienation of Theory and Practice (Publication no. ED 462004).
- Ministry of Education, Taiwan. (2014). *The outline of English curriculum for secondary and primary schools*. Taipei, Taiwan: Author.
- Moore, Z. T. (1994). The portfolio and testing culture. In Hancock, C. (Ed.), *Teaching, testing, and assessment: Making the connection* (pp. 163-176). Illinois: National Textbook Company.
- Nguyen, L. T. (2005). From passive participant to active thinker: a learner-centered approach to materials development. *English Teaching Forum*, 43(3), 2-27.

- Ohashi, Y. (2015). Communicative language teaching for young learners of English: A critical perspective. *Fukuoka Jo Gakuin University Bulletin Faculty of International Career Development*, 1, 17-30.
- Omaggio, A. H. (2001). *Teaching language in context* (3rd ed.). United States: Thomson Heinle.
- Oxford, R. (1997). Cooperative learning, collaborative learning, and interaction: Three communicative strands in the language classroom. *The Modern English Journal*, 81(iv), 443-456.
- Pan, Y. C. (2013). A closer examination of communicative language teaching. *Studies in Literature and Language*, 6(2), 40-44.
- Pham, H. H. (2005). "Imported" communicative language teaching: implications for local teachers. *English Teaching Forum*, 43(4), 2-9.
- Pham, H. H. (2007). Communicative language teaching: unity within diversity. *ELT Journal*, 61(3), 193-201.
- Pierce, L. V., & O'Malley, J. M. (1992). Performance and portfolio assessment for language minority students. *NCBE Program Information Guide Series*, 9, 1-33.
- Pokoma, L. N., & Vasylieva, O. V. (2014). The pedagogical approach to communicative language teaching in primary school. *GISAP Educational Sciences*, 4, 47-49.
- Rao, Z. (1996). Reconciling communicative approaches to the teaching of English with traditional Chinese methods. *Research in the Teaching of English*, 30(4), 458-471.
- Richards, J. C. (1993). Beyond the textbook: the role of commercial materials in language teaching. *ERLC Journal*, 24(1), 1-14.
- Richards, J. C. (2006). *Communicative language teaching today*. New York: Cambridge University Press.
- Rubin, H. J., & Rubin, I. S. (2005). *Qualitative interviewing: The art of hearing data* (2nd ed.). London: SAGE Publications.
- Saengboon, S. (2002). *Beliefs of Thai EFL teachers about communicative language teaching* (Doctoral dissertation, Indiana University Bloomington, 2002).
- Sato, K. (2002). Practical understandings of communicative language teaching and teacher development. In Savignon, S. J. (Ed.), *interpreting communicative language teaching: Contexts and concerns in teacher education* (pp. 40-81). New Haven, London: Yale University Press.
- Savignon, S. J. (1983). *Communicative competence: Theory and classroom practice*. Menlo Park, California: Addison-Wesley Publishing Company.
- Shamim, F. (1996). Learner resistance to innovation in classroom methodology. In H. Coleman (Ed.), *Society and the Language Classroom*, 105-121. Cambridge: Cambridge University Press.
- Spada, N. (2007). "Communicative language teaching: Current status and future prospects." In J. Cummins & C. Davison (Eds.), *International Handbook of English Language Teaching* (Vol.1, pp. 271-288). New York: Springer.
- Sreehari, P. (2012). Communicative language teaching: Possibilities and problems. *English Language Teaching*, 5(12), 87-93.
- Su, Y. C. (2006). EFL teachers' perceptions of English language policy at the elementary level in Taiwan. *Educational Studies*, 32(3), 265-283.
- Tanaka, T. (2009). Communicative language teaching and its cultural appropriateness in Japan. *Doshisha Studies in English*, 107-123.
- Wada, M. (2002). Teacher education for curricular innovation in Japan. In Savignon, S. J. (Ed.), *Interpreting communicative language teaching: Contexts and concerns in teacher education* (pp. 31-39). New Haven, London: Yale University Press.
- Wang, C. (2002). Innovative teaching in foreign language contexts: the case of Taiwan. In Savignon, S. J. (Ed.), *Interpreting communicative language teaching: Contexts and concerns in teacher education* (pp. 131-153). New Haven, London: Yale University Press.
- Wang, Y. H. (2010). Using communicative language games in teaching and learning English in Taiwanese primary schools. *Journal of Engineering Technology and Education*, 7(1), 126-142.
- Weiss, S. R. (1994). *Learning from strangers: the art and method of qualitative interview studies*. New York: The Free Press.
- Wharton, S., & Race, P. (1999). *500 tips for TESOL: teaching English to speakers of other languages*. Sterling: Stylus Publishing Inc.
- Wu, W. (2008). Misunderstandings of communicative language teaching. *English Language Teaching*, 1(1), 50-53.

- Wu, Y. (2001). English language teaching in China: trends and challenge. *TESOL Quarterly*, 35(1), 191-194.
- Yoon, K. E. (2004). CLT theories and practices in EFL curricula: a case study of Korea. *Asian EFL Journal*. 1-16.
- Yu, L. (2001). Communicative language teaching in China: progress and resistance. *TESOL Quarterly*, 35(1), 194-198.
- Zhang, J. L. (2006). The ecology of communicative language teaching: Reflecting on the Singapore experience. *Paper presented at the Annual CELEA International Conference: Innovating English Teaching: Communicative Language Teaching (CLT) and Other Approaches, 11-13 November 2006 China English Language Education Association (CELEA) and Guangdong University of Foreign Studies, Guangzhou, China.*
- Zhu, H. (2003). Globalization and new ELT challenges in China. *English Today* 76, 19(4), 36-41.

About the Authors

Shu-Hsiu (Sofie) Huang attended the Indiana University Bloomington (IUB), where she completed her Master's and PhD degrees in language education. She has experience in teaching Mandarin Chinese (at IUB) and English (in Taiwan). Her research interests focus on English Language Teaching, English as a Foreign Language, Language Learning Motivation, and Bilingual Education.

Li-Chin (Jennifer) Yang received her PhD degree in Linguistics at the Department of English, National Taiwan Normal University. She is currently an assistant professor in the Center for Language and Culture at the Kaohsiung Medical University. Her research interests include language acquisition, psycholinguistics, OT syntax, and interface between cognitive linguistics & SLA.