

The Asian EFL Journal Quarterly

June 2016

Volume 18 Issue 2



Senior Editors:

Paul Robertson and John Adamson

Asian EFL Journal



Published by the English Language Education Publishing

Asian EFL Journal

A Division of TESOL Asia Group

Part of SITE Ltd Australia

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

© Asian EFL Journal 2016

This book is in copyright. Subject to statutory exception no reproduction of any part may take place without the written permission of the Asian EFL Journal Press.

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 the Asian EFL Journal.

editor@asian-efl-journal.com

Publisher: Dr. Paul Robertson

Chief Editor: Dr. John Adamson

Production Editor: Wen-Chi Vivian Wu

ISSN 1738-1460

Table of Contents:

Foreword by Wen-Chi Vivian Wu..... 5-8

1. Pino Cutrone..... 9-88
- Explicit vs. Implicit Instruction: Investigating Backchannel Behavior in the Japanese EFL Classroom

2. Susanna S. Yeung, Mei-lee Ng, & Ronnel B. King..... 89-112
- English vocabulary instruction through storybook reading for Chinese EFL kindergarteners: Comparing rich, embedded, and incidental approaches

3. Ali-Al-Issa..... 113-155
- The Effects and Implications of Implementing Oral Presentations in an Omani ICLHE Classroom

4. Anna Dina L. Joaquin, Stephanie Hyeri Kim, & Sun-Young Shin..... 156-189
- Examining Prewriting Strategies in L2 Writing: Do They Really Work?

5. Soyeon Kim..... 190-215
- No Point in Talking about What I Want to the Teachers: A Call for a Dialogic Needs Assessment

6. Diane Johnson & Hsiu-Chen (Antonia) Lin..... 216-243
- Evaluating genre-based writing instruction: Materials, instructional mode and student learning styles

7. Ratnawati Mohd-Asraf & Hazlina Abdullah 252 - 262
- English Literacy among Malaysian First Graders: Boys' Achievement Relative to Girls

Book Reviews

1. *Multiple Perspectives on the Self in SLA*..... 244-247

S.Mercer & M. Williams (eds). 2014. *Multilingual Matters*: Bristol, ix + 189 pages

Reviewed by **Brian Wadman**, Mahidol University, Thailand

2. *Language, Education and Nation-building- Assimilation and Shift in Southeast Asia*..... 248-251

P. Sercombe and R. Tupas. Palgrave Macmillan, NY. 2014. Pp. XIV, 264

Reviewed by **Raheb Zohrehfard**, University of British Columbia, Vancouver, Canada

June 2016 Foreword

by Wen-Chi Vivian Wu

Welcome to the June issue of Asian EFL Journal in 2016. This issue includes six research articles and two book reviews which explore diverse and vital topics in the field of teaching and learning English as a second/foreign language. Specifically, it addresses backchannel behavior, vocabulary instruction, storybook reading, oral presentation, prewriting strategies, dialogic needs assessment, genre-based writing, and learning style. Such diversity is also embodied by the inclusion of authors from different geographic regions around the globe, well demonstrating the vitality of the Asian EFL community.

This issue opens with a research paper proposed by Pino Cutrone. The study, entitled, “Explicit vs. Implicit Instruction: Investigating Backchannel Behavior in the Japanese EFL Classroom”, examined the effect of explicit and implicit instruction on Japanese EFL learners’ listening behavior over the course of 16 weeks, with three tests (at the beginning, at the 8th week, and at the 16th week) involving the participants’ intercultural conversation, questionnaire responses, and interviews. The participants in the study were divided into three groups, including two experimental groups and one control group. While one experimental group featured the teachers’ explicit explanations of rules, the other experimental group was characterized by implicit instruction evolved from peer group discussions and conversational practice with native English speakers. The findings revealed that while both explicit and implicit instructions yielded a positive effect, the Explicit group outperformed the Implicit group, highlighting the explicit teaching of listening behavior in the Japanese EFL context.

Susanna S. Yeung, Mei-lee Ng, and Ronnel B. King, in their study entitled “English vocabulary instruction through storybook reading for Chinese EFL kindergarteners: Comparing rich, embedded, and incidental approaches”, probed into the effects of vocabulary instruction within a

storybook reading context for young EFL learners in China. The study specifically compared three approaches (i.e., rich, embedded, and incidental) to enhancing vocabulary knowledge among Chinese EFL kindergartners. A within-subject research design was adopted, with all of the 43 participants experiencing the aforementioned three approaches in a storybook reading context. The results indicated that only rich instruction significantly enhanced the participants' receptive and expressive knowledge of word meanings. This article can serve as guidance for English instructors who are interested in integrating storybook reading into Chinese EFL learners' curriculum.

In the following article, Ali Al-Issa focused on examining the role of oral presentations in an Integrating Content and Language in Higher Education (ICLHE) classroom in Oman. A total of 44 students with diverse English proficiencies from the Law Department were subdivided into 11 groups, with each group required to present a Law-related topic in English. The analysis of the participants' responses to the 36-item self-report questionnaire about the learning experience suggested that such learning experience was related to their academic and professional lives, and had positive effects on English language education policy in Oman.

In an attempt to contribute to the insufficiency of prewriting in L2 writing, Anna Dina L. Joaquin, Stephanie Hyeri Kim, and Sun-Young Shin specifically investigated whether prewriting strategies benefited L2 writers. The participants included 513 university level English language learners who have taken a timed essay placement exam. The results revealed that over 50% of the participants chose to prewrite, most of whom chose outlining, listing, or a combination of strategies in the writing process. The participants who elaborated their prewriting scored higher than those who prewrote minimally or in a standard manner. While the findings supported the positive effects of prewriting on L2 writing, the study also yielded conflicting results in that outlining and listening were not as effective as anticipated and that combined strategies did not significantly affect the writing outcome. Pedagogical implications for ESL/EFL researchers and teachers who are interested in understanding the use and effects of prewriting strategies have also been provided by the author.

To avoid problematic needs assessment that teachers and students often face in locating real language needs, Soyeon Kim directed readers' attention to the importance of dialogic needs assessment in EFL writing. Designed to trace the transformations of three Korean university students' needs during a semester using a dialogic frame, the qualitative study included three rounds of in-depth interviews, field notes, and class observation, from which themes and types of needs were analyzed. The author concluded that in the writing process, student needs were transformative from vague to specific needs, and that dialogic needs assessment provided a formative learning experience for students and teachers. However, the author cautioned that limited types of evaluative teacher feedback led to misunderstanding about the teacher's level of interest in students' learning. Pedagogical implications and suggestions for future implementation of dialogic needs assessment in EFL writing are provided.

Diane Johnson and Hsiu-Chen (Antonia) Lin, in their study entitled "Evaluating genre-based writing instruction: Materials, instructional mode and student learning styles", measured the efficacy of a purpose-designed, genre-based writing course and potential impact that the mode of delivery and the students' learning styles might have on engagement and overall writing development. During the 10-day instruction, 28 participants, divided into three groups, were taught using one of three delivery modes—fully online, face-to-face, or blended. Multiple sources of data included pre-test and post-test of a genre-centered writing task, pre-course learning styles inventory, post-course questionnaire, and follow-up focus-group discussion. The participants made improvement in all assessed writing aspects and demonstrated a high level of satisfaction with the course (particularly the participants in the face-to-face group and the blended group). The online group outperformed other groups in terms of writing improvement. Learning styles and learning mode preferences, however, were not related. This article brings new insights for researchers and instructors who are interested in genre-based writing instruction in EFL writing context.

We hope you find the articles in this June 2016 issue to be informative, inspiring, and comprehensive. Bearing in mind the contribution to continuous improvement in English language instruction around the world, we sincerely hope that this issue helps provide new insights into the formulation of future research and innovations for EFL/ESL practitioners in cross-border,

interdisciplinary, and collaborative manners. We would like to express our sincere appreciation to the contributors and reviewers of articles and book reviews who have made this issue possible. Special thanks also go to my research assistant, Mr. Jun Scott Chen Hsieh, for his devotion to careful formatting and overall layout of the issue.

Explicit vs. Implicit Instruction:

Investigating Backchannel Behavior in the Japanese EFL Classroom

Pino Cutrone

Nagasaki University, Japan

Bio data

Pino Cutrone has been teaching in Japan for fifteen years and is currently an Associate Professor at Nagasaki University. He received his PhD in Applied Linguistics from the University of Reading and has published widely in his field. His research interests include intercultural pragmatics, sociolinguistics and EFL pedagogy. k-pino@nagasaki-u.ac.jp

Abstract

This study examined the effect of explicit and implicit instruction on Japanese EFL learners' listening behavior over the course of 16 weeks. 30 university student participants were divided into three groups and given tests at three points in time: at the beginning, at the eight-week mark, and at the 16-week mark of this study. Each of these tests involved participating in an intercultural conversation, completing a questionnaire and being interviewed. In one group, the explanations of rules were given explicitly by a teacher; in the second, implicit instruction evolved mainly from peer group discussions and conversational practice with native English speakers. The third group was a control. The findings demonstrate that both explicit and implicit methods had a positive effect; however, overall, the Explicit group generally outpaced the Implicit group. These findings, thus, provide support for the explicit teaching of listening behavior in the Japanese EFL context.

Address correspondence:

1-14 Bunkyo-machi, Nagasaki
City, Nagasaki Prefecture,
Japan, 852-8521

Key words: EFL pedagogy in the Japanese context,
intercultural pragmatics, explicit versus implicit learning,
listenership, backchannel behavior

Introduction

Since English is generally regarded as the lingua franca for communication used in such fields as international politics, academia, business and science, more people around the world are studying EFL/ESL every day. Few nations have expended greater resources encouraging their citizens to study English than Japan, yet the results to date have been largely unsatisfactory (Nikolova, 2008), particularly concerning oral skills (Alun, 2008). A key aspect of effective oral communication is being able to give effective feedback to one's interlocutor (O'Keeffe, McCarthy & Carter, 2007), and this is a specific area in which the writer contends that Japanese EFL learners (JEFLs hereafter) have experienced problems (Cutrone, 2005). It is becoming increasingly clear that what constitutes effective feedback seems open to interpretation, and there is potential for cross-cultural pragmatic failure and misunderstanding when listening styles differ. In an attempt to inform language pedagogy in the JEFL context, the main aim of this article is to determine how to improve this aspect of JEFLs' English.

Examining Listening Behavior from a Research Perspective

There exist several terms in the research literature to describe what is meant by listening behavior. For instance, McCarthy (2002, 2003) uses the term "good listenership" to describe the active responses that listeners produce in conversations. Other researchers, as Fujimoto (2007) notes, have used one of the 24 terms on her extensive list to describe various elements of listening behavior. The term, backchannel, coined by Yngve (1970) is perhaps the one linguists seem to associate most often with listening behavior and will, thus, be the focus of the analysis in this paper. Yngve (1970) describes a backchannel as follows:

When two people are engaged in conversation, they generally take turns...In fact, both the person who has the turn and his partner is simultaneously engaged in both speaking and listening. This is because of the existence of what I call the backchannel, over which the person who has the turn receives short messages such as yes and un-huh without relinquishing the turn. (p. 568)

Although some researchers such as Oreström (1983) continue to follow Yngve's (1970) original definition, others researchers such as Duncan (1974) and Duncan and Fisk (1977) extend what is meant by the term backchannel to include sentence completions, requests for clarification, brief

statements, and non-verbal responses such as head nods and headshakes. In the studies most closely resembling this current project (in terms of research design and in the use of Japanese participants), Cutrone (2005), Maynard (1997) and White (1989) differed slightly in their identification of backchannels. White (1989), focusing solely on non-word vocalizations such as mhm, yeah, uh-huh, oh, and hmm, limited her analysis to audio recording and thus did not include nonverbal behavior. Cutrone (2005) and Maynard (1986, 1987, 1989, 1990, 1997), on the other hand, used a broader identification of backchannels as proposed by Duncan and Fiske (1977) in that they too include sentence completions, requests for clarification, brief statements, and nonverbal items such as head nods and laughing. As the researcher agrees that brief utterances and nonverbal behavior by the listener are indeed backchannels in that they serve as messages to the primary speaker, this broader definition is used to identify backchannels in this current research project.

Listener Backchannels vis-à-vis Speaking Turns

A major issue in identifying a backchannel is determining whether a particular behavior constitutes a backchannel or a separate turn. According to Maynard (1986, p. 1084), much of the confusion may stem from distinguishing between “having a turn” and “having the floor” and can be attributed to self-contradictions in Yngve’s (1970) definition. Yngve’s definition of a backchannel is given in terms of “not relinquishing a turn”; however, he cites the following example as backchannel behavior:

In one case, what looked like backchannel activity consisted of filling in needed personal background so that the person having the floor could continue. This went on for about thirty seconds and involved a number of sentences. It is interesting to note that this extensive backchannel activity was in turn provided with back-back channel activity of the ‘uh-huh’ variety. (p. 568)

In this quote, Yngve appears to be identifying backchannel behavior on the basis of holding the floor, rather than having the turn. Thus, longer utterances such as *You’re ready to go then* in response to the primary speaker’s talk can cause confusion because this utterance may allow the primary speaker to continue holding the floor, yet it appears to be a speaking turn in itself. Consequently, what starts as a backchannel can actually end up as a turn, if the primary speaker

shows no willingness to continue speaking. To differentiate between backchannels and turns, this study identifies backchannels in the context of Markel's (1975) turn-taking system, which he describes as follows:

A speaking turn begins when one interlocutor starts solo talking. For every speaking turn there is a concurrent listening turn, which is the behavior of one or more nontalking interlocutors present. (p. 190)

Thus, following the work of Cutrone (2005) and Maynard (1997), the position taken in this study is to identify a brief statement as a backchannel and not a primary turn when it serves only to react to what the primary speaker is saying (i.e., having a listening function) and not to add any new information to the conversation (i.e., having a speaking function). Therefore, brief questions such as *Is that right?* or *Oh really?*, which are formed in terms of requests for clarification, are classified as backchannels. However, a question such as *Why did she do that?* would be identified as a full speaking turn because it serves a speaking function in terms of driving the conversation in a new direction. Further, responses to questions are not regarded as backchannels because, as Ward and Tsukuhara (2000) have pointed out, backchannels are unlike responses to questions in that they are optional and not required. In addition, answers to questions, which are sometimes quite brief and include ellipsis, would also seem to offer new information that pushes the conversation forward constituting a change of primary speakership. Finally, researchers have to make decisions regarding how to deal with pauses and utterances found between turns at talk. Following the aforementioned studies, this study identifies an utterance as a backchannel only when it occurs immediately after the primary speaker stops talking (within one second) and is followed by a substantial pause before the next turn at talk starts (exceeding one second). This decision was made because it was felt that such backchannels are produced in response to the primary speaker's speech, and they occur before a substantial turn transitional period starts.

Functions of Backchannels

Classifying the functions of backchannels in an organized list is a difficult task for several reasons. First, as Cutrone (2010) has shown, the overlap between form and function is considerable, and the function that any given backchannel serves is highly dependent on the

context of the conversation as well as the listener's personality. Although researchers have proposed a range of functions, there has been little consensus to date, as there appear to be reliability issues in measuring this aspect of conversation. In many studies, it is not always clear how analysts reach their conclusions concerning backchannel functions, and it appears that many based their findings on conversational analyses focusing on the primary speakers' interpretations of their interlocutors' backchannel functions, as shown by how the conversation unfolded. While the post-hoc examination of conversational transcripts may offer important clues as to what went on in the conversation, it does not necessarily take into account what the non-primary speaker meant to convey with their backchannel utterance. Therefore, this study aims to take into account the observed backchannel behavior of the participants, the non-primary speakers' stated backchannel intentions, and their interlocutors' perceptions of these backchannels. Classifying the various functions of backchannels, Maynard (1997) has attempted to sum up the previous work in this area by identifying the following six categories: continuer, understanding, agreement, support and empathy, emotive and minor additions. Detailed explanations and hypothetical examples to demonstrate which forms are used to convey each function are described in Cutrone's (2010) article. In this current study however, what is of particular interest to the researcher is the use of unconventional backchannel behavior by the JEFs, i.e., forms and functions that do not seem to correspond. As the section on functions below explains, this includes the JEF tendency to send supportive backchannel forms such as those to allow the speaker to continue, to show understanding, agreement, and/or support and empathy when they do not understand or agree with what their interlocutor is saying.

Why is Backchannel Behavior Important in the Japanese EFL context?

The primary reason backchannel behavior is becoming such an important topic in language learning is that more and more people are becoming acutely aware of the great impact it can have on intercultural communication (IC). As O'Keeffe, Clancy and Adolphs (2011, p. 100) point out, attempts to move between L1 and L2 pragmatic norms can feel like "a minefield for learners of a language". Various studies in this area have shown that JEFs' listening behavior differs to that of native English speakers (NESs) in many respects (Clancy, Thompson, Suzuki & Tao, 1996; Maynard, 1986, 1987, 1990, 1997; White 1989) and these differences sometimes lead to miscommunication and negative perceptions across cultures (Blanche, 1987; Boxer, 1993;

Cutrone, 2005). This section, thus, outlines some of these differences in listening behavior and discusses their potential impact on IC. To this end, this section also serves to provide a general barometer for JEFs to produce effective listening behavior.

Frequency of Backchannels

Several intercultural analyses (Clancy et al. 1996; Crawford, 2003; Cutrone, 2005; Maynard, 1986, 1990, 1997; White, 1989) have observed JEFs producing backchannels far more frequently than their NES interlocutors (i.e., Britons, Americans and Australians) in both their L1 and L2 English. As discussed below, such frequent interjections may be taken as a sign of impatience and demand for a quick completion of the statement (Cutrone, 2005; Lebra, 1976; Mizutani, 1982).

Variability of Backchannels

In their intercultural analyses, Cutrone (2005) and Maynard (1997) reported that the JEFs' backchannels consisted mainly of non-word vocalizations and head nods, whereas the NESs (American and British respectively) exhibited greater variability in the types of backchannels they sent. That is, the NESs employed a far greater range of content words in their backchannels and were also able to produce far more extended and complex backchannel responses. Lack of variability, as well as repetition, in backchannel form may be interpreted as a sign of boredom and inattention (Cutrone, 2005; McCarthy, 2002, 2003).

Discourse Contexts Attracting Backchannels

A common finding in the studies to date has been that Japanese discourse contexts attracting backchannels varied considerably (in both the L1 and L2 English), while grammatical completion points (i.e., clausal boundaries), especially those coinciding with a pause, were the single most important discourse contexts for NESs (Cutrone, 2005; Maynard, 1986, 1990, 1997; White, 1989).

Backchannels Creating Simultaneous Speech

A general finding in the research literature is that Japanese people, regardless of whether they are speaking English or Japanese, tend to backchannel more frequently than NESs, and a great portion of these backchannels are produced during the primary speakers' speech, creating simultaneous speech (Cutrone, 2005; Hayashi, 1988; Maynard, 1997). As Lebra (1976) and

Mizutani (1982) have hypothesized, some listeners may take these frequent interjections as a sign of the listener's impatience and demand for a quick completion of the statement. Hence, in attempt to dig deeper, White (1989) and Cutrone (2005) conducted correlational analyses to find out what effect Japanese people's frequent backchannels might have on their cross-cultural interlocutors perceptions of them. Reporting quite different results, White (1989) found that the ten Americans perceived more frequent backchannels by their Japanese interlocutors as a positive trait (i.e., showing signs of comprehension, encouragement, and interest and concern), whereas Cutrone (2005) reported that the eight Britons in his intercultural analysis perceived more frequent backchannels by their JEFL interlocutors as interruptions and signs of impatience.

Backchannel Functions

Although research in this subcategory of listenership is scant, there is some evidence beginning to emerge suggesting some key functional differences across cultures. For instance, Blanche (1987) and Cutrone (2005) have provided anecdotal evidence of Japanese people providing unconventional backchannels in English, such as by employing continuer, understanding, agreement, and/or empathy/support type backchannels in situations when they did not understand or disagreed with what their interlocutor was saying at the time. For instance, in the classroom scenario described by Blanche (2005) and Cutrone (2005), NES teachers sometimes misinterpret students' classrooms responses (such as nods coupled with vocalisations of *yeah* and *uhuh*) as displays of understanding, rather than the mere polite expressions of attending that the students mean to express. When teachers realize much later on that students have not understood them, they may occasionally feel confused and/or even slightly annoyed by what they perceive to be mixed signals, or in extreme cases, deceptive messages, resulting in the squandering of valuable class time.

In situations outside the classroom where there may be more at stake, these misunderstandings can have dire consequences, as was the case in the Hitachi-Mitsubishi trial (The Japan Times 1983, p. 2). One of the defendants in the case, Mr. Takaya Ishida of Mitsubishi, maintained that he had not agreed with the undercover FBI agents when they told him he had to steal some information/documents. His defense counselor argued that Mr. Ishida's responses of *yeah* and *uhuh* were not to show agreement, but rather to show he was listening and to allow the other person to continue.

Involvement in the Conversation

A great many intercultural analyses comparing the communicative behaviors of Japanese L2 English speakers and NESs have shown that the Japanese L2 English speakers in these studies spoke less than NESs, did not elaborate as much, and were less likely to engage in small talk (Cutrone, 2005; Hill, 1990; Sato, 2008). Undoubtedly, this is potentially a source of misunderstanding in an English conversation as the importance of making small talk, taking the initiative to speak, and elaboration towards making a positive impression have been documented by several sources (Cutrone, 2005; McCarthy, 2002, 2003; McCroskey, 1992; Sato, 2008; Stubbe, 1998).

Explicit and Implicit Learning

Considering the importance of listenership in IC, and the fact that it is largely neglected in EFL pedagogy in Japan (Capper, 2000), the primary aim of this study is to determine how this aspect of pragmatic competence can best be learned in the EFL classroom. The context of instruction underpinning the examination of JEFLL's acquisition of L2 listenership is a central issue in this study and touches upon the more general, and oft-debated, issue of how languages are best learned: explicitly or implicitly (Rose, 2005). The main feature distinguishing these two instructional approaches is the general provision of metapragmatic information designed to make the target features more salient in the explicit approach. Explicit instruction commonly involves providing students with explicit metapragmatic information about L2 rules through explanations, metacognitive discussions and corrective feedback. Conversely, implicit instruction generally involves presenting learners with prototypical uses of the target language in meaningful contexts with or without input enhancement. The underlying assumption here is that the models of language given to learners should help raise their awareness so that they will be able to induce the rules for appropriate L2 use on their own.

Following the well-known skills' acquisition theory known as the Adaptive Control of Thought (ACT) theory (Anderson & Lebiere, 1998), supporters of an explicit approach view language learning as progression from declarative/explicit knowledge through proceduralization to final automatization. Conversely, in line with the tenets of Krashen's (1982) Input Hypothesis theory, proponents of a non-interface position believe that explicit knowledge cannot be converted to implicit knowledge. Drawing on the differences between L1 and L2 acquisition processes, Ellis (2005, 2006a,b) has made attempts to explain the dichotomy between explicit and

implicit learning. Cognitively, unlike L1 learners who are thought to acquire language implicitly, L2 learners come to the learning environment with minds already endowed with knowledge and experience of a prior linguistic system (i.e., the L1). This cognitive state seems to act as a constraint and filter in L2 input reception and output generation, two critical processes of SLA. Goldschneider and DeKeyser (2001) and Ellis (2005, 2006a,b) have identified factors that enable features of language to be noticed and hence guide practitioners to what might need explicit instruction. Therefore, explicit knowledge is thought to contribute to implicit learning. According to Ellis (2006a, p. 19), “the linguistic forms that L2 learners fail to adopt and to use routinely thereafter in their L2 processing are those which, whether available as a result of frequency, recency, or context, fall short of intake because of one of the following associative learning factors”: (1) unreliable predictors of outcome, (2) not attended because of low cue salience, (3) not attended because of low functional outcome in the overall interpretation of the message, (4) not attended because they are redundant in the immediate understanding of an utterance, being overshadowed or blocked by higher salience cues which have previously been selected, and (5) ignored because L1 experience of form→meaning contingencies affects the cues and dimensions that an L2 learner’s language input systems can best distinguish (perceptual learning), and L1 experience of meaning→form contingencies affects the way a L2 learner routinely expresses their meanings in language.

While all five of Ellis’ (2006a) factors listed above have the potential to hinder the acquisition of L2 listenership, factors four and five likely pose the greatest threats. As previous studies have shown (Crawford, 2003; Maynard, 1997; White, 1989), there is a great deal of evidence showing the negative transfer of backchannel behavior from L1 Japanese to L2 English, which suggests the possibility of L1-influenced blocking. As Brozyna (2007) and Takimoto (2009) have pointed out, despite their pervasiveness, conversational discourse features such as backchannels are largely non-salient and often pass unnoticed unless attention is drawn to them intentionally and explicitly. Ellis (2006a, b), thus, advocates explicit and conscious L2 learning as a way of supplementing or directing frequency-driven learning of patterns. Ellis believes that consciousness is required to change behavior in SLA, as L2 learners do not seem to be equipped to notice low-salient cues at a subconscious level, particularly when proficiency levels using the more obvious cues are already sufficient for everyday communicative survival. These beliefs are in line with Schmidt’s (1993) Noticing Hypothesis, which stipulates that learners must

consciously notice linguistic input in order for it to become intake.

Can L2 English Listenership Be Learned?

Research has shown that the acquisition of L1 backchannels is a fairly slow developmental process, stretching into adolescence (Hess & Johnston, 1988). As several studies have shown evidence of L1 Japanese to L2 English negative transfer and fossilization where backchannels are concerned (Crawford, 2003; Maynard, 1997; White, 1989), L2 listenership is thought to be an especially difficult skill-set to learn. This seemingly uphill battle may be one of the reasons why it has been rarely taught and/or studied (Thonus, 2007). While evidence showing the benefits of form-focused-instruction (FFI) on various aspects of SLA such as grammatical development is indeed mounting (Ellis, Rod, 2008), the effects of FFI on pragmatic development remains unclear. Despite the many calls for studies examining the teachability of listenership (Cutrone, 2005; O’Keeffe & Adolphs, 2008; McCarthy, 2002, 2003), only a handful of studies in this area have been published to date, most of which seemed to have been brief and/or focused on other aspects of SLA. For instance, concerning the latter, Schmidt (1983), in his longitudinal case study reporting on the English development of a native speaker (NS) of Japanese (whom he called Wes in the reporting of his study), in an immersion setting, inadvertently noticed a great improvement in Wes’ listenership after three years abroad, which would seem to provide support for the implicit learning of this skill-set. Schmidt (1983), however, did not analyze listening behaviour in any systematic way; rather, his inferences seemed to be based mainly on personal observations and anecdotal evidence.

Ward and Tsukuhara (2000) and Ward, Escalante, Yaffa and Solorio (2007) have experimented with CALL methods as a means of teaching the timing of backchannels in English and Arabic respectively. A major issue involving this type of instruction is the seemingly narrow focus of teaching the appropriate timing of English backchannels without giving attention to other crucial elements of backchannel behavior such as form, function and contextual variables. In fact, mastering the skills to produce backchannels at appropriate times, whether they understand what the primary speaker is saying or not, may actually reduce learners’ overall L2 communicative ability. That is, there exists evidence of IC being negatively affected by instances of JEFLL speakers producing continuer, understanding, or agreement type backchannels even when they disagree or did not understand what the primary speaker was saying (Blanche, 1987;

Cutrone, 2005; LoCastro, 1987). A backchannel trainer that focuses solely on timing would, in essence, be helping learners to fake understanding, which has previously been identified as a potential problem area that warrants instructional attention to remedy.

In her unpublished Master's thesis, Brozyna (2007) attempted to assess the teachability of backchannels in a study that also sought to investigate the effects of instruction on discourse markers functioning as hedges and fillers. Basing her study on the awareness raising methodology of Illustration-Interaction-Induction (Carter & McCarthy, 1995, 2004) and on the "noticing the gap" (Schmidt & Frota, 1986, p. 311) potential for learning activation, Brozyna compared the performance of a control group (attending regular English classes four hours per week), with two treatment groups receiving an additional two hours of experimental instruction over 12 weeks. One of the experimental groups was given rich exposure to target features as well as explicit instruction in terms of metalinguistic information regarding backchannels and discourse markers, whereas the other experimental group received the same treatment with the addition of opportunities for intensive and focused practice. Brozyna reported a significant improvement in the group that received practice opportunities, but not in the group that did not receive practice opportunities or the control group. Although Brozyna intended this project to be a pilot study for future research, and the practical constraints of her teaching context certainly played a role, there are a number of issues associated with the design as well as findings of her study. First, the six participants in each of the treatment groups consisted of male and female high school students, whereas the four participants in the control group were all adult females, aged 24-26, in full-time employment. The fact that age, gender and socioeconomic status were not controlled makes the results difficult to compare between groups.

Second, the brevity of the method of analysis, which was limited to counting the number of backchannels in one two-minute extract pre instruction and another two-minute extract post instruction, would also seem to bring into question the validity of the results. There would also appear to be some validity concerns involving the fact that these two-minute extracts appear to have involved all six participants conversing at once. Considering the complexities of how group dynamics might influence backchannel behavior, it is easy to imagine how one or two members of a particular group, with potentially divergent and/or altered backchannel behavior, could skew the results of such a small scale study. Moreover, concerning studies that measure the effectiveness of instruction on a particular language feature, it is difficult to give findings any

credence without a delayed post-test to demonstrate whether the treatment had any sustained effects.

Finally, Brozyna appears to have based her evaluation of the efficacy of the instruction solely on the quantity of backchannels used and did not include other important aspects of quality such as form and function. That is, the fact that the group that received full treatment increased the amount of backchannels they uttered from 3 to 12 (or from .012 to .04 / number of primary speaker words) after instruction seems to have provided the foundation for stating that these participants had improved. In the researcher's opinion, the assessment of backchannel behavior is not as straightforward as this. The relationship between backchannels and discourse is not a linear one; rather, it is complex and multifaceted, with frequency comprising only one of the many interrelated and overlapping components in this skill-set. In the case of JEFs, whose style of frequent backchannels was found excessive in various intercultural analyses (Cutrone, 2005; LoCastro, 1987), the goal of instruction would be, conversely, to limit the overall frequency of their backchannels (especially the minimal and repetitive ones), while increasing their length and variability at context-specific moments (repair strategies to deal with non-understanding).

Lastly, a study administered by Sardegna and Molle (2010) examined the learnability of English listener responses in five JEFs after a two-hour video conference lesson. Sardegna and Molle claim that this pedagogic intervention had a positive short-term effect on the production of students' listener responses; however, it seems difficult to give too much credence to this finding based on the data presented, as well as the brevity of the treatment and analysis. Sardegna and Molle seem to be basing their conclusion solely on the limited amount of data presented in a short post-treatment conversational excerpt involving the five JEFs and the teacher. Further, besides the problems associated with analyzing all the JEFs together discussed above, the criteria for assessing listening behavior again appears to be oversimplified as the lone determinant of success was the observation that the JEFs used fewer backchannel forms common in Japanese and more that are common in English. Finally, technological limitations associated with the videoconferencing set-up forced the researchers to exclude nonverbal backchannels in their analyses.

Research Questions (RQs)

While there is evidence to suggest that pedagogical intervention can facilitate some elements of pragmatic development (Rose, 2005), little is known about how learners acquire L2 listener responses over time, as well as what effect different types of instruction might have on development in this area. RQ 1, thus, seeks to answer the more general question of whether pedagogical intervention addressing conversational discourse features of language can have a positive effect, while RQ 2, more specifically, seeks to shed light on the Explicit-Implicit debate where listener responses are concerned.

RQ 1: Will instructional treatment help facilitate the JEFLL learners' backchannel behavior?

RQ 2: If so, will explicit treatments be more effective than implicit ones (a) in the short term and (b) in the long term?

Methodology

Participants

Constituting an opportunistic sample, this study involved 46 participants known to the researcher. First, in the intercultural conversation phase, this study used 30 JEFLLs born and raised in Japan (24 females and six males) and six L1 American English speakers from the United States (three females and three males). All conversational participants lived in Nagasaki Prefecture in Japan and were university students ranging in age from 18-20 at the time of the study. The Americans participating in the conversations (called the NES interlocutors in this study) of the study were visiting Japan as exchange students. In addition, ten American participants (called NES observers in this study) who did not participate in the conversations and/or have any particular affiliation to Japan were used to provide unbiased assessments of the JEFLLs in the conversations (i.e., see Procedural Step 6 below). The NES observers ranged in age from 22-48 and were from various parts of the United States. Participating of their free will and understanding the nature of the study, all 46 participants read and signed a Participation Consent Form and were given explicit instructions regarding this study and their role in it. All forms were typed in English with Japanese translations provided to the native Japanese speakers in this study to ensure these participants had a full understanding of the contents in each form. In referring to participants in

this study, pseudonyms have been given to protect participants' privacy.

Procedures in Collecting Data

As the RQs were complex and multifaceted in nature, it was thought that using a mixed-methods approach would not only serve to strengthen the reliability and validity of the data but also provide for a broader view with multiple perspectives of the researched topic. To this end, the following methods were used to collect data in this study: observations, questionnaires and interviews.

Observations

The observation phase involved the video recording of 30 intercultural dyadic conversations in English between JEFL and NES participants. To control for sociolinguistic variables, all dyads were paired according to gender, age (within two years) and social status (i.e., all university students), and participants were not well-acquainted prior to each conversation. The conversations took place in the researcher's office in Nagasaki Prefecture, Japan, and, once brief instructions were given, only the participants were present in the room. Initial conversational prompts (i.e., involving peer-mentoring) were offered to help stimulate conversation; however, it was made clear to participants that they were free to talk about anything they liked. Each conversation was video recorded for a period of thirty minutes, of which only the middle three minutes of each conversation were included as data to be transcribed (see Appendix A for sample transcription). It was thought that the participants would become less conscious of the camera as the conversation progressed, and that the middle part of the conversation would be the most natural as it avoids the awkwardness which often occurs at the beginning and end of conversations between people who do not know each other well. Further, as discussed below, one of the purposes of the interviews was to examine the extent to which the conversational participants might have changed their behavior due to the presence of the video camera in the room (known as Observer's Paradox). Consistent with similar studies (Cutrone, 2005; Maynard, 1986) that have involved video recorded conversations, Observer's Paradox did not seem to be a major issue in this study. Only three of the participants reported to being aware of the video camera initially, as they looked towards the camera and made metamessage comments such as *we can begin now* and *I shouldn't say that in front of the camera*; however, they, along with the rest of the participants, indicated that they were not conscious of the camera at all once the

conversation developed and felt that it did not influence their behavior in any way.

Lastly, to strengthen the internal reliability of the transcriptions, the researcher analyzed the video recorded conversation with the assistance of a colleague. The colleague, who was trained by the researcher to recognize the transcription conventions used in this study (outlined in Appendix A), which were adopted from the pioneering Conversation Analysis (CA) research of Sacks, Schegloff and Jefferson (1974), and the revised CA conventions of Jefferson (2002), assisted in two ways: first, regarding some of the words and/or gestures that were not clear to the researcher in the conversations, the researcher sought out the colleague's opinions so as to make a more informed choice in transcribing the word(s) or behavior in question; second, the colleague double-checked the researcher's initial transcriptions to ensure that the conventions showing the features of language were accurately presented in the transcriptions. The conversations were conducted in the researcher's office. Video recording equipment consisted of a Sony digital video camera, which was placed unobtrusively in the corner of the room. At the time of the video recording, only the participants were present in the room when the conversation was taking place.

Questionnaires

There were three types of questionnaires used in this study. The first two, inventories to measure WTC (McCroskey, 1992; see Appendix B) and the Extraversion dimension of personality (Oshio, Abe & Cutrone, 2012; see Appendix C), were used to create equally balanced groups prior to assessment in this study (as Procedural Step 2 below explains). The third type of questionnaire, Hecht's (1978) conversational satisfaction inventory, was used more extensively throughout the study as part of the assessment of the JEFs' performance. To this end, slightly modified versions of Hecht's questionnaire were given to the participant groups of this study, as shown in Appendix D (i.e., the NES and JEF conversational participants and the NES observers). Questionnaires for the conversational participants consisted of a fifteen-item inventory and one open-response question at the end in case the participant wanted to add something that had not been addressed in the other questions. The questions on the fifteen-item inventory were closed-ended, consisting of statements on a Likert-scale ranging from one to seven. The researcher modified the questionnaire from Hecht's (1978) original and the one White (1989) used because pilot studies revealed that some vocabulary and some of the statements, which contained double negatives, confused participants. For instance, participants of the pilot study had problems rating yes or no on the Likert-scale because of negatively worded statements such as *He/she didn't seem to care*

and *He/she did not interrupt me*. Consequently, the negatively worded statement *He/she didn't seem to care* was excluded from the questionnaire because its positive counterpart *He expressed a lot of interest in what I had to say* already existed in the questionnaire, and similarly, the negatively worded statement *He/she did not interrupt me* was replaced with the positive statement *He/she interrupted me*. Moreover, item 15 *My conversation partner seems to want to avoid speaking* was added to Hecht's (1978) original questionnaire. The JEFL participants' questionnaires included Japanese translations underneath each of the items typed in English to ensure comprehension.

Other modifications were made specific to the questionnaire objectives for each of the two groups. For instance, item 2 on Hecht's (1978) original questionnaire *I felt I was able to present myself favorably during the conversation* was altered for both groups. On the JEFL participants' questionnaire, the adverb *favorably* was changed to *fairly* as the focus here was to determine whether these participants felt that they represented themselves adequately, sincerely, and as they had intended (i.e., this need not necessarily be favorably). To avoid confusion, this was emphasized in the Japanese translation. On the NESs' questionnaire, the original item 2 was changed in its entirety to *The feelings that my partner expressed by means of listening feedback during the conversation seemed authentic and sincere*. The focus here was to examine whether the NESs could sense their JEFL interlocutor faking understanding and/or agreement. The post-conversation questionnaires were given to each participant in the dyad directly after their video recorded conversation and simultaneously completed in separate rooms.

Subsequently, a slightly modified version of Hecht's questionnaire was given to a group of American observers to assess their perceptions of the JEFLs' performances from watching the video recorded conversations. First, since the members of the NES observer group did not participate in the conversations, referent nouns, subject pronouns, object pronouns and possessive adjectives were changed accordingly. Second, NES observers' questionnaires included two supplemental items (items 16 and 17) on the Likert-scale designed to examine specific areas of relevance to the study of JEFLs' backchannel behavior that were not covered in the original questionnaire. These consisted of item 16 *When the Japanese person did not understand something, they were able to clearly convey this to their conversational partner with their listening feedback* and item 17 *The Japanese person's listening behavior seemed inadequate in some ways*. For participants expressing disagreement (i.e., 5, 6, or 7 on the Likert-scale) to item

17, an open-item response was subsequently solicited asking them to explain their answer.

Interviews

The interviews involved the two members of each dyad being interviewed separately and in succession, with the NES first and the JEFL second. No one else was present at the time of the interview, as the researcher asked the participant not being interviewed to wait in another room while the interview was being conducted. None of the interviews took longer than twenty minutes. The interviews consisted of the researcher playing back a portion of the video recorded conversation and asking each participant a few questions pertaining to the listening behavior displayed in the conversation. The researcher took field notes and audio recorded all interviews to refer to in the data analysis. The interviews were semi-structured in that the researcher had a general plan but did not enter with a predetermined set of questions, as some of the questions were guided by the circumstances in the video recorded conversations and the responses of the interviewee. In an effort to help participants feel more comfortable, questions were sequenced to begin with general inquiries and gradually move towards more specific and potentially sensitive questions. As mentioned above, the interviewer began by assessing the extent to which the conversational participants might have been affected by Observer's Paradox.

The primary aim in the interviews with the American participants was to learn how they perceived their Japanese interlocutors' backchannel behavior. A major part of the interview involved playing back the video recording and asking the American interviewee to comment on the listening behavior of their Japanese interlocutor. In instances singled out for analysis (where some of the JEFLs' backchannels occur), the researcher stopped the video recording and asked specific questions such as *What function do you think that head nod serves?*, *Do you think s/he understands what you are saying here?*, and follow up questions such as *Why do you think so?* The researcher made a note of any data which was thought to be useful in the subsequent interview with the Japanese participant of the dyad.

Subsequently interviewing the Japanese participant, the researcher had two main objectives: to gain insights into why JEFLs use backchannels the way they do and to determine if there were any misunderstandings or miscommunications caused by their use of backchannels in the video recorded conversations. Regarding the latter, this involved the interviewer asking the JEFLs to comment on what they were feeling at certain points in the conversation, as well as the intended functions of their backchannel responses. In particular, the researcher investigated what the

JEFLs were doing when they did not understand and if they were indeed feigning understanding as various researchers claim is common for Japanese backchannel behavior. In the successive interviews, the researcher documented any notable differences between the Japanese participants' backchannel intentions and their NES interlocutors' perceptions on a data record sheet (see Appendix E). In the cases where the JEFL's backchannel explanation differed greatly from their NES interlocutor's interpretation, the interviewer asked the JEFL potentially sensitive follow-up questions such as *Why did you nod if you did not understand what he/she was saying here?* and *Why did you say yeah if you disagreed with what he/she was saying?* In cases where the interviewee seemed uncomfortable in answering, the interviewer did not persist with this line of questioning and instead shifted to a less sensitive area.

Procedural Steps

Step 1: Pre-test

Each JEFL received one pre-test which contained three parts: observations, questionnaires and interviews (as described above in the Procedures of Collecting Data section). The first part, observations, refers to the JEFLs participating in a conversation with a NES (which was video recorded and subsequently watched, transcribed and assessed). The second part refers to a conversational satisfaction questionnaire filled out by both JEFLs and their NES interlocutors. The third part refers to retrospective interviews with both JEFLs and their NES interlocutors regarding their intercultural conversations.

Step 2: Group Formation

The 30 JEFLs in this study were divided into three groups of ten (see member characteristics of each group in Appendix F). These groups comprised the two experimental groups (Groups A and B) and one control group (Group Z). The main objective in grouping the participants was to attempt to create comparable and thus equally balanced groups in terms of gender, EFL proficiency as based on the Test of English for International Communication (TOEIC) scores, WTC (see McCroskey, 1992), personality (i.e. based on the Extraversion dimension, see Oshio, Abe & Cutrone, 2012) and age.

Step 3: Treatment

Group A: Receiving Explicit Instruction

Members of Group A received explicit instruction over the course of two months, i.e., one 90-minute lesson a week for eight weeks. This began with a focus of explicitly and consciously raising learners' awareness of the use and dimensions of listener responses in English as compared to Japanese. The following teaching methods were used in the classroom: (1) the form, function and perception of listener responses, as well as other communicative behavior thought to be relevant, were compared across languages and cultures, (2) with input enhancement from the teacher, members of Group A closely examined intracultural and intercultural conversations and subsequently discussed how differences in form, frequency, placement and function of backchannels in varying contexts affect IC, (3) learners were provided with overt descriptions and concrete examples of how NESs employ backchannels and conversational micro-skills in varying contexts, (4) the JEFs in Group A completed exercises and tasks to reinforce the newly learned conventions of backchannel behavior, and, (5) with a focus on fluency, learners practiced using listener responses in role-plays. By video recording role-play performances and watching them afterwards, students were able to analyse their performances and receive feedback from the teacher and other students.

Group B: Receiving Implicit Instruction

Members of Group B received implicit instruction over the course of two months. While the overall amount of instruction (and the eight-week time-line) was identical to that received by Group A, the lesson parameters were altered. The lessons involving classroom interaction and discussion amongst peers, which took place in Weeks 1, 4 and 8, were each administered in one 90-minute sitting; however, the sessions in which the JEFs conversed with a NES, which occurred in Weeks 2, 3, 5, 6 and 7, took place twice a week and were 45 minutes in duration each. Similar to the aims of the explicit instruction above, raising learners' consciousness regarding the use and dimensions of listener responses across cultures was among the main goals of implicit instruction in this study; however, different methods were used to achieve this. Unlike the instruction Group A received, learners in Group B were not given any explicit metapragmatic information about listener responses through explanations and concrete examples. While the members of Group B were asked initially to consider the general qualities belonging to good conversationalists and listeners alike, they were not instructed to focus on listening behavior as

part of this treatment. The teacher's role was limited to that of facilitator in that the JEFs in Group B were required to induce rules and meaning on their own based on exposure to prototypical uses of the target language in meaningful contexts. Members of Group B received the following pedagogical interventions. First, learners were provided with models of conversations in three categories: (1) NS-NS discourse in Japanese, (2) NS-NS discourse in English and (3) NES-JEFL conversations in English. Second, with limited input enhancement from the teacher, the JEFs in Group B were asked to reflect upon and discuss their observations of communicative behavior across languages and cultures in journal-writing entries and small group discussions. Third, making up the greater part of the treatment Group B received, learners were given the opportunity to communicate with NESs face-to-face in authentic contexts on 12 separate occasions. Each conversational session was followed by a brief period of reflection in which the JEFs in Group B recorded their thoughts in their journals. In addition, after every three conversational sessions, learner-led group discussions took place, affording them the opportunities to further reflect and share their general observations regarding their own as well as their cross-cultural interlocutor's communicative style(s), what they might have learned, and what they hope to change or modify about their own behavior in future conversations in English.

Group Z: No Treatment

The JEFs in the control group received no formal treatment regarding listenership.

Step 4: Post-test 1

Post-test 1 was given to all JEFs within three days of their last treatment, and followed the same procedure as the Pre-test outlined in Step 1 above.

Step 5: Post-test 2

To assess whether the treatment had any sustained effects, a delayed post-test was given. Applying identical procedures as those used in the Pre-test and Post-test 1, Post-test 2 was administered approximately eight weeks after Post-test 1.

Step 6: NESs' Observations

Ten NES observers were instructed to watch each three-minute video recorded conversation and subsequently provide their impression as to the adequacy of each JEF's conversational

performance by completing a modified version of Hecht's (1978) conversational satisfaction questionnaire.

Data Analysis and Assessment Criteria

As mentioned above, three data collection methods were used in this study: observations, questionnaires and interviews. The data produced by these methods combine to inform judgements within the following assessment categories of backchannel behavior: the approximation of NES backchannel behavior, Willingness to Communicate (WTC), the ability to use conversational micro-skills and repair strategies, and Intercultural Communicative Competence (ICC). This section, thus, presents the methods of data analysis in each of these assessment categories.

Inferential statistics were included at various points to determine the probability that an observed difference between means (such as the means involving the same group at two points in time, as from the Pre-test to Post-test 2) was a significant one or one that might have happened by chance. Both parametric and non-parametric statistical tests were used depending on the type of data analyzed. Regarding the observational data (i.e., measuring the approximation of NES backchannel behavior) and the WTC questionnaires, Paired-Samples t-tests were used because the data in these areas fulfilled the following four requirements of parametric testing:

- (1) The pairs of scores must be related to each other.
- (2) The scores must be of at least interval status.
- (3) The scores in each group must be normally distributed.
- (4) The two sets of scores have equal variances. (Clegg, 1982, p. 167)

Conversely, concerning the data collected from the NES observers' questionnaires, the non-parametric Wilcoxon Signed Rank test was used because these data were not drawn from a normal distribution (as shown by an exploratory data analysis that included Q-Q plot graphs, a normal curve superimposed over histograms of the data and the Shapiro-Wilk normality test). To analyze the data in this study, the Statistical Package for the Social Sciences (SPSS), version 14.0, was used. In all inferential statistical tests used in this study, two-tailed tests were used, and alpha levels (α) were set at 0.05 ($p < .05$ is marked by a single asterisk *). Since the groups in this study were small, probability statistics have to be viewed with caution. Hence, considering the possibility of Type 1 errors (i.e., the false rejection of the null hypothesis) occurring, probabilities less than the more stringent 0.01 level were also emphasized ($p < .01$ level =**).

The Approximation of NES Backchannel Behavior

To determine how well the JEFs were able to approximate NES norms (as presented in Cutrone, 2010), the researcher carefully analyzed the JEFs' performances in the intercultural conversations. As mentioned above, all of these conversations were transcribed and, thence, formed the *observation* part of the tests in this study (sample transcription shown in Appendix A). To analyze the data of the observation phase of this study, the transcriptions were closely examined for patterns and tendencies as well as how frequently certain behaviors occurred in the data.

Frequency

Following the precise definitions of backchannels given above, the researcher determined the frequency of backchannels produced by a person or group in this study by counting the number of backchannels in the conversational transcripts. Further, to provide a more accurate representation of how frequently participants sent backchannels according to the opportunities they were given (i.e., how much the other conversational participant spoke), the number of a participant's backchannels was divided by the number of their interlocutor's words.

Variability

Similar to the overall Frequency category above, the frequency of various subtypes of backchannels is based on how frequently the subtype occurred in the data of the transcripts. While numerous verbal and nonverbal backchannel subtypes have been explored, this paper presents only the findings of the superordinate groups, minimal vis-à-vis extended backchannels, because they are thought to have the greatest influence on IC (see Cutrone, 2011; Stubbe, 1998). Minimal responses can be defined here as any brief (non-lexical) and/or nonverbal backchannel occurring in isolation, which include head nods and/or non-word vocalizations such as *uhuh* and *mhm*, whereas extended responses, conversely, refer to the lengthier verbal listener feedback consisting of multiple and varied words found in sentences (irrespective of nonverbal backchannel accompaniment) such as *How great that is* or *Do you think so?*

Discourse Contexts Favoring Backchannels

Upon exploring several subcategories, the primary discourse context favoring backchannels in English has consistently proven to be the final clausal boundary (see Cutrone, 2005; Maynard,

1997; White, 1989). In simple terms, a final clause boundary refers to the points in the primary speaker's speech where a clause (i.e., containing a subject and predicate) is completed. It is important to note, however, the distinction between a final clause boundary and an internal clause boundary. A final clause boundary is one that makes complete sense (i.e., fully meaningful) and could end the utterance there, whereas an internal clause is one in which the meaning is not complete, and there is a requirement for the utterance to go on in order for the meaning to be complete. As shown in Appendix A, two slashes side by side (//) mark the points at which final clause boundaries occur in the primary speakers' speech. Since backchannels in the transcripts are presented in italics below the primary speakers' talk at the point they occurred in the talk, the backchannels that occurred at final clause boundaries in the primary speaker's speech were clearly discernible. In the findings, the two main categories relating to this area that will be presented are as follows: (1) the mean percentage of opportunities (Opps) that final clause boundaries attracted backchannels (with SDs), and (2) the mean percentage of backchannels constituted by final clause boundaries (with SDs).

Simultaneous Speech

Since Simultaneous Speech Backchannels (SSBs) have the danger of being misconstrued as interruptions (Cutrone, 2011), this paper includes a separate category for them. SSBs are recorded when a backchannel is uttered during the primary speaker's speech, and, hence, backchannels which occur during pause periods in the primary speaker's speech are not considered SSBs. SSBs include laughter; however, as SSB laughter is not thought to influence communication negatively (i.e., it was doubtful that SSB laughter would be seen as interruptive), this study limits its reporting in this area to the non-laughter SSB category.

Increase Willingness to Communicate (WTC)

The examination of participants' WTC is based on two main sources: McCroskey's (1992) self-report WTC questionnaires and behavioral observations. Regarding the former, analyzing JEFLL scores in the interpersonal communication sub-category of McCroskey's (1992) WTC inventory (see Appendix B) offers a sense of how willing the JEFLLs were to communicate. To determine if the WTC self-ratings were borne out in the conversational data, the amount that each person spoke (i.e., in terms of how many words they uttered in the transcriptions) was also included in this analysis. Lastly, since the importance of asking questions has been widely stated, the

researcher included this as another area for analysis.

The Development of Conversational Micro-skills and Repair Strategies

Analysis of this category was delimited to situations when JEFs did not understand or agree with their interlocutor. In addition to documenting the number of times JEFs use potentially unconventional and/or inappropriate backchannels in these situations, this analysis also examines the instances in which they use conversational repair strategies. Conversational repair strategies are divided into two types of listener response: minimal and extended responses. Minimal responses refer to any brief (non-lexical) backchannel considered to be a request for clarification and/or a demonstration of non-understanding. The intent of such backchannels was judged in two ways: by asking the participants in the playback interviews what they meant to convey with their backchannel response, and by examining the intonation patterns of minimal backchannels. Concerning the latter method, backchannel repetitions with a rising intonation commonly serve to request clarification; thus, using the well-known Phonetics software called Praat (Version 5.0.18), minimal backchannels with pitch contours containing rapid rises of 600 Hertz or more were marked as conversational repair strategies (i.e., requests for clarification). Extended responses refer to a specific set of typical conversational phrases and routines such as *I beg your pardon* or *what does that mean?* As a basis for identifying expressions that make up these repair strategies, this study recognizes text strings that correspond to and are similar to the models set forth in two language teacher's resource books that have been entirely based on these structures: *Function in English* (Blundell, Higgins, & Middlemiss, 1982) and *Conversation and Dialogues in Action* (Dörnyei & Thurrell, 1992).

The Development of Intercultural Communicative Competence (ICC)

A basic requirement of ICC is for conversational participants to be seen as competent by members of the target culture (Byram, 1997; Spitzberg, 2000). To assess this in this study, the NES interlocutors (i.e., conversational participants) were given post-conversation questionnaires and interviews (as described above). Further, the NES observer group completed a modified version of Hecht's (1978) Interpersonal Communication Satisfaction Inventory (see third questionnaire in Appendix D). Since the NES observers were based abroad, did not participate in the conversations and did not have any affiliation to Japan (and were, thus, unbiased and offered perceptions that were potentially closer to what JEFs would encounter abroad), the researcher

examined their questionnaire responses in greater detail.

Results

The results of the JEFLL groups' performances at each of the three points of measurement (i.e., the Pre-test, Post-test 1 and Post-test 2) are presented quantitatively according to each group (in the tables below) and collectively (in the figures below) in the following areas of backchannel behavior: frequency, involvement in the conversation, variability, discourse contexts favoring backchannels, simultaneous speech backchannels, and perceptions of NES observers. In addition, the interview data, which are presented both quantitatively and qualitatively, report on how members of each group dealt with situations of non-understanding and how they perceived their improvements or lack thereof throughout this study.

Frequency

Tables 1, 2 and 3 report the backchannel frequencies of Group A (Explicit), Group B (Implicit) and Group Z (Control) respectively at the three points of measurement in this study.

Table 1

Group A's backchannel frequencies

| N = 10 | Total Backchannels | | | Interlocutor's Words | | | Average number of interlocutor's words between backchannels | | |
|--------------------|--------------------|--------|--------|----------------------|--------|--------|---|--------------------|--------------------|
| | Pre | Post 1 | Post 2 | Pre | Post 1 | Post 2 | Pre | Post 1 | Post 2 |
| Total | 450 | 290 | 371 | 2700 | 2874 | 3217 | — | — | — |
| Mean (\bar{x}) | 45 | 29 | 37.1 | 270 | 287.4 | 321.7 | 6 | 9.91** (p<.001) | 8.67** (p<.001) |
| SD | 15.23 | 11.12 | 12.33 | 21.83 | 12.39 | 18.14 | 1.19 | 2.45 | 1.94 |

(\bar{x} difference of Pre-test → Post-test 1, and Pre-test → Post-test 2 significant at p<.05 level = *; significant at p<.01 level = **)

Table 2

Group B's backchannel frequencies

| N = 10 | Backchannels | | | Interlocutor's Words | | | Average number of interlocutor's words between backchannels | | |
|--------------------|--------------|--------|--------|----------------------|--------|--------|---|-------------------------|--------|
| | Pre | Post 1 | Post 2 | Pre | Post 1 | Post 2 | Pre | Post 1 | Post 2 |
| Total | 370 | 268 | 391 | 2530 | 2355 | 3120 | — | — | — |
| Mean (\bar{x}) | 37 | 26.8* | 39.1 | 253 | 235.5 | 312 | 6.84 | 8.79* ($p < .046$) | 7.98 |
| SD | 10.67 | 4.83 | 9.78 | 15.25 | 20.16 | 15.79 | 1.52 | 1.87 | 1.29 |

(\bar{x} difference of Pre-test \rightarrow Post-test 1, and Pre-test \rightarrow Post-test 2 significant at $p < .05$ level = *; significant at $p < .01$ level = **)

Table 3

Group Z's backchannel frequencies

| N = 10 | Backchannels | | | Interlocutor's Words | | | Average number of interlocutor's words between backchannels | | |
|--------------------|--------------|--------|--------|----------------------|--------|--------|---|--------|--------|
| | Pre | Post 1 | Post 2 | Pre | Post 1 | Post 2 | Pre | Post 1 | Post 2 |
| Total | 445 | 317 | 474 | 2914 | 2381 | 3408 | — | — | — |
| Mean (\bar{x}) | 44.5 | 31.7 | 47.4 | 291.4 | 238.1 | 340.8 | 6.55 | 7.51 | 7.19 |
| SD | 13.01 | 9.15 | 12.93 | 18.23 | 10.47 | 11.24 | 1.37 | 1.26 | .89 |

(\bar{x} difference of Pre-test \rightarrow Post-test 1, and Pre-test \rightarrow Post-test 2 significant at $p < .05$ level = *; significant at $p < .01$ level = **)

Examining the three groups' data collectively, Figure 1 shows that members of the experimental groups (Groups A and B) sent backchannels (BCs) much less frequently in Post-test 1 compared to the Pre-test. Subsequently, in Post-test 2, the number of backchannels the experimental groups sent returned to a level closer to their original Pre-test level. In comparison, the Control group (Group Z) only showed a slight decrease in the number of backchannels they sent from the Pre-test to Post-test 1. From the Pre-test to Post-test 1, the 3.91 backchannel per interlocutor word decrease shown by the Explicit group was found to be strongly significant ($p < .001$), and the 1.95 decrease for the Implicit group was significant at the .05 level ($p < .046$). Further, while all three groups had reverted back to providing more frequent backchannels in Post-test 2, there was some variability between the frequencies in each group. For example, the Control group sent

backchannels only slightly less frequently in the Pre-test than in Post-test 2 (i.e., a difference of only .64 backchannels per interlocutor word), whereas the Implicit group sent backchannels noticeably less frequently in the Pre-test than Post-test 2 (difference of 1.14 backchannels per interlocutor word). The Explicit group, however, was the only one to maintain statistically significant changes ($p < .001$) in frequency from the Pre-test to Post-test 2 (difference of 2.67 backchannels per interlocutor word).

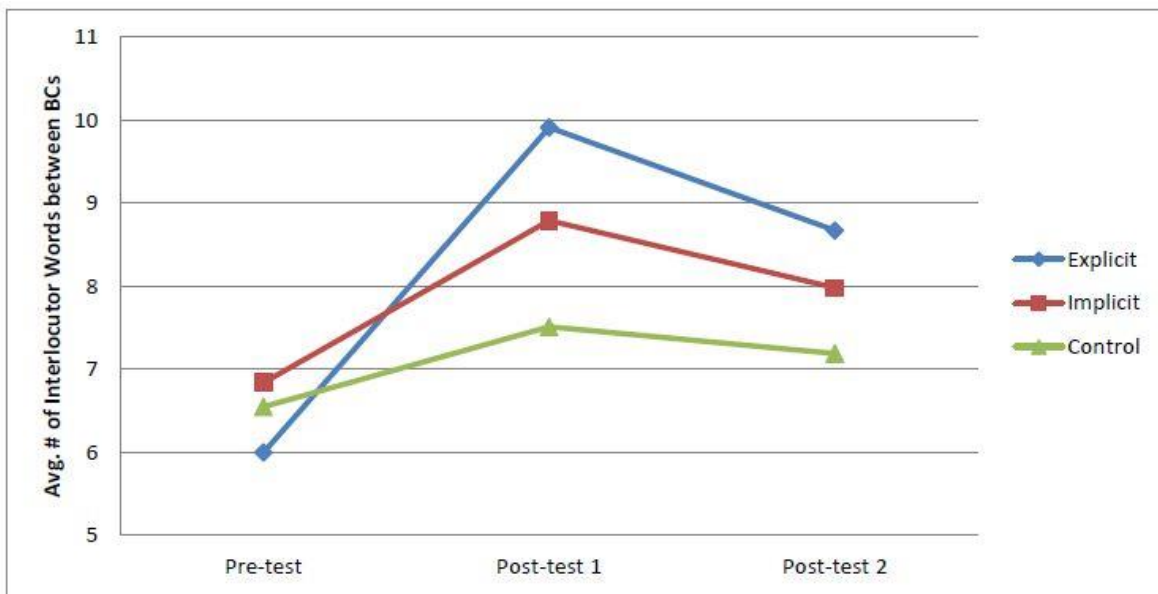


Figure 1. Backchannel frequencies of the three groups over time

Willingness to Communicate

As mentioned above, the JEFs' willingness to communicate takes into account their WTC scores, how much they spoke in the conversations and the number of questions they asked their interlocutor. Tables 4, 5 and 6 present the results of these features of conversations for Group A (Explicit), Group B (Implicit) and Group Z (Control) respectively at the three points of measurement in this study.

Table 4

Group A's willingness to communicate

| N = 10 | WTC | | Words | | | Questions | | |
|--------|--------------------|-------|-------|--------------------|-------|-----------|-------------------|-----|
| | \bar{x} | SD | Total | \bar{x} | SD | Total | \bar{x} | SD |
| Pre | 48.4 | 10.74 | 583 | 58.3 | 32.82 | 8 | .08 | .79 |
| Post 1 | 70** (p<.001) | 6.79 | 973 | 97.3** (p<.007) | 35.87 | 21 | 2.1** (p<.001) | .88 |
| Post 2 | 57.7** (p<.001) | 9.12 | 567 | 56.7 | 26.37 | 14 | 1.6* (p<.011) | .70 |

(\bar{x} difference of Pre-test \rightarrow Post-test 1, and Pre-test \rightarrow Post-test 2 significant at p<.05 level = *; significant at p<.01 level = **)

Table 5

Group B's willingness to communicate

| N = 10 | WTC | | Words | | | Questions | | |
|--------|--------------------|------|-------|-----------|-------|-----------|-----------|------|
| | \bar{x} | SD | Total | \bar{x} | SD | Total | \bar{x} | SD |
| Pre | 50.4 | 8.14 | 650 | 65 | 32.33 | 12 | 1.2 | .63 |
| Post 1 | 60.6** (p<.002) | 6.08 | 901 | 90.1 | 31.19 | 16 | 1.6 | 1.17 |
| Post 2 | 56.6* (p<.022) | 8.21 | 705 | 70.5 | 23.01 | 11 | 1.1 | 1.29 |

(\bar{x} difference of Pre-test \rightarrow Post-test 1, and Pre-test \rightarrow Post-test 2 significant at p<.05 level = *; significant at p<.01 level = **)

Table 6

Group Z's willingness to communicate

| N = 10 | WTC | | Words | | | Questions | | |
|--------|-----------|------|-------|-----------|-------|-----------|-----------------|-----|
| | \bar{x} | SD | Total | \bar{x} | SD | Total | \bar{x} | SD |
| Pre | 56.2 | 7.27 | 640 | 64 | 32 | 18 | 1.8 | .92 |
| Post 1 | 57.3 | 7.42 | 963 | 96.3 | 30.63 | 14 | 1.4 | .84 |
| Post 2 | 58.5 | 7.37 | 641 | 64.1 | 33 | 9 | .9* (p<.041) | .74 |

(\bar{x} difference of Pre-test \rightarrow Post-test 1, and Pre-test \rightarrow Post-test 2 significant at p<.05 level = *; significant at p<.01 level = **)

Presenting the three groups' data collectively, WTC scores, number of words and number of questions asked at the three tests are compared in Figures 2, 3 and 4 respectively. First, as Figure 2 demonstrates, the Explicit group showed the greatest increase in average WTC score from the Pre-test to Post-test 1 (+21.6), as well as the greatest sustained increase from the Pre-test to Post-test 2 (+9.3). These increases were both found to be strongly significant (at the .01 level). The path of the Implicit group generally mirrored that of the Explicit group but without the same range in scores. The 10.2 increase in this group's average WTC score from the Pre-test to Post-test 1 was found to be strongly significant (at the .01 level), and the 6.2 increase from the Pre-test to Post-test 2 was significant at the .05 level or below. For the Control group, the average WTC scores remained fairly constant over time, only showing minor increases of 1.1 and 2.3 from the Pre-test to Post-tests 1 and 2 respectively.

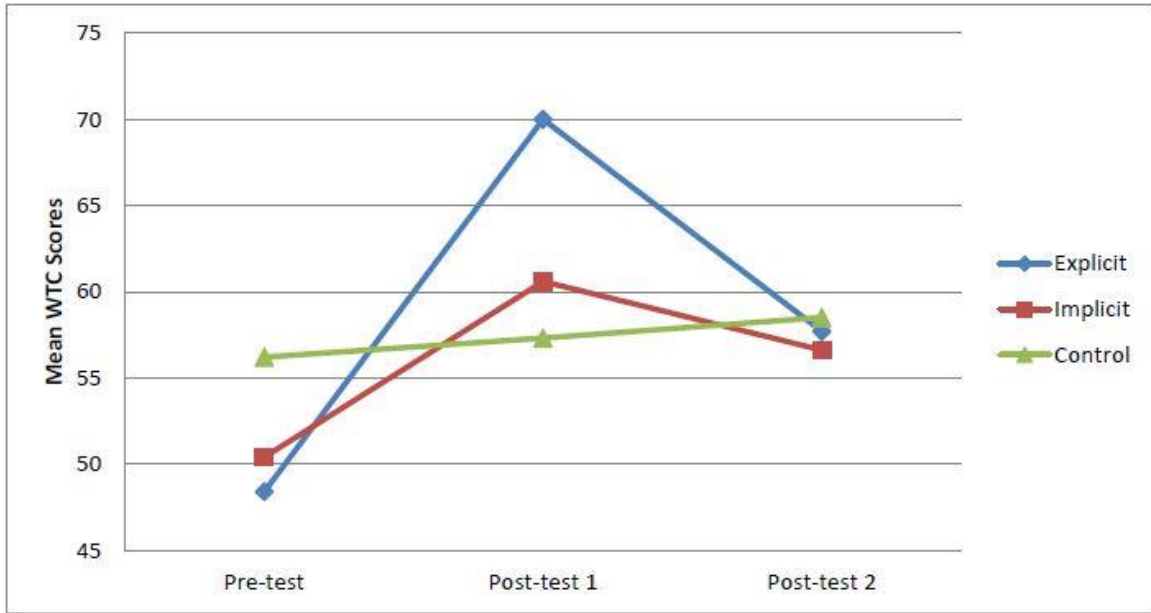


Figure 2. WTC questionnaire scores of the three groups over time

As Figure 3 illustrates, the three groups followed a similar path in terms of word output. Similar to the frequency category (see Figure 1), the general trend, regardless of which group they belonged to, was for the JEFs to speak much more frequently in Post-test 1 compared to the Pre-test and then, in Post-test 2, revert to a level closer to their original Pre-test level. The mean increase exhibited by the Explicit group (39 words) was the only one found to be statistically significant from the Pre-test to Post-test 1 ($p < .007$); however, the large differences for the Implicit (25.1) and Control groups (32.3) were also noticeable.

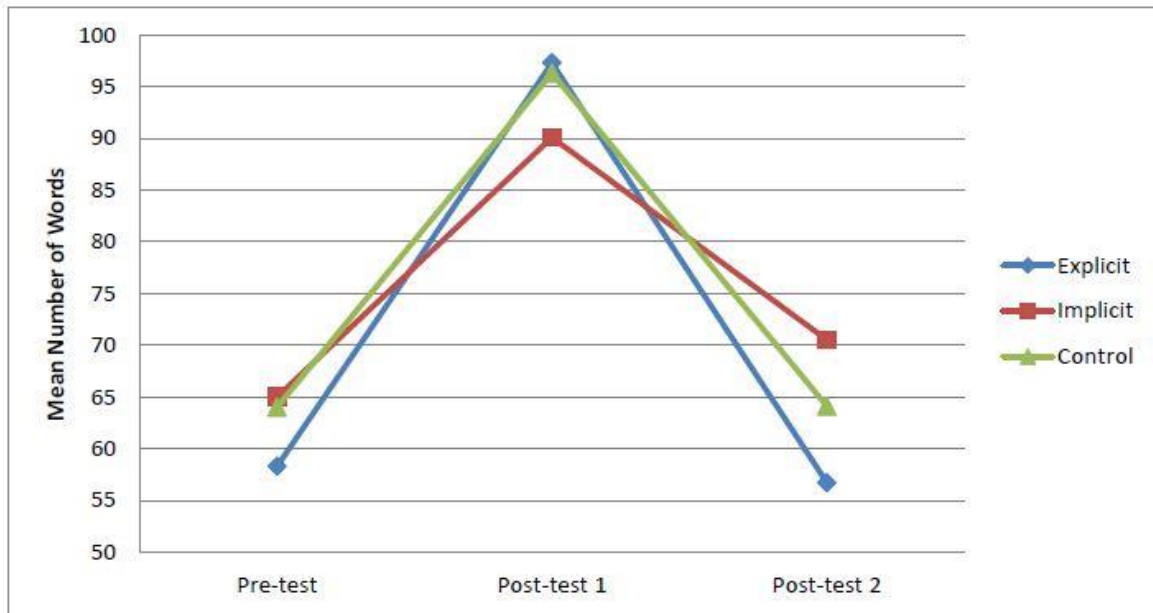


Figure 3. Number of words of the three groups over time

As shown in Figure 4, the path that each group followed in terms of the average number of questions posed was different. The Explicit group displayed the greatest initial increase in questions from the Pre-test to Post-test 1 (+1.3), as well as a sustained increase from the Pre-test to Post-test 2 (+.8). Paired-samples t-tests found both of these increases to be statistically significant ($p < .001$ from Pre-test to Post-test 1, and $p < .011$ from Pre-test to Post-test 2). In comparison, the Implicit group showed only a modest initial increase from the Pre-test to Post-test 1 (+.4), and ultimately a slight decrease overall from the Pre-test to Post-test 2 (-.1). The Control group was the only group to decrease in both measurements after the Pre-test, i.e., (-.4) from the Pre-test to Post-test 1 and (-.9) from the Pre-test to Post-test 2. The latter decrease was found to be statistically significant at the .05 level or below.

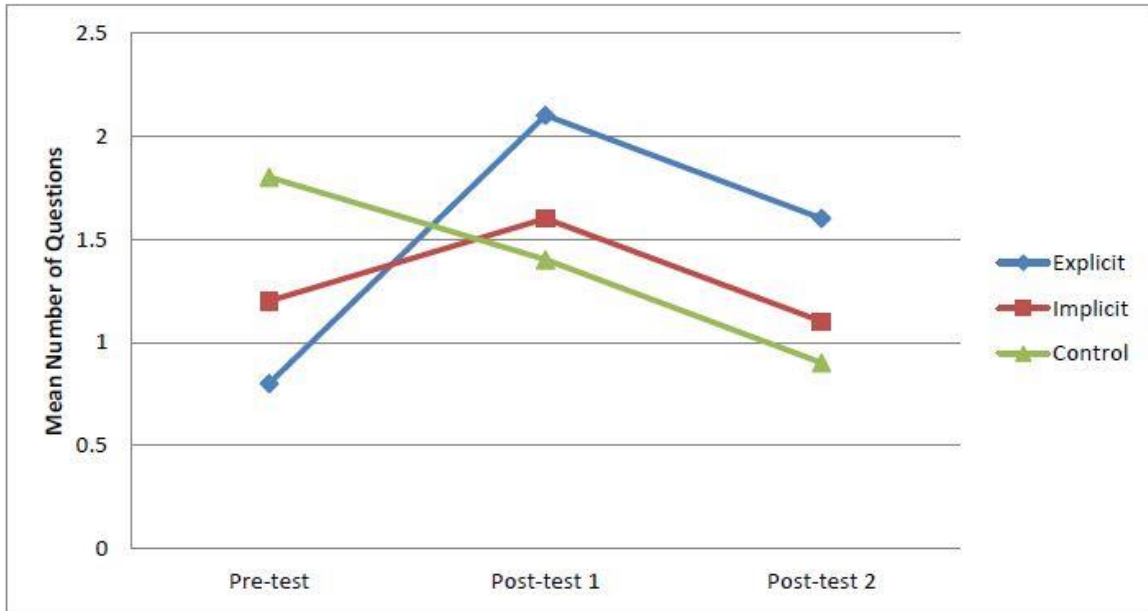


Figure 4. Number of questions of the three groups over time

Variability

As discussed above, numerous verbal and nonverbal backchannel subtypes have been explored; however, this paper presents only the findings of the superordinate groups, minimal and extended backchannels, because they are thought to have the greatest influence on IC. Tables 7, 8 and 9 report the results of these features of conversations for Group A (Explicit), Group B (Implicit) and Group Z (Control) respectively at the three points of measurement in this study.

Table 7

Group A's use of minimal versus extended backchannels over time

| N = 10 | Pre-test | | | Post-test 1 | | | Post-test 2 | | |
|-------------------|----------|--------------------------|------|-------------|--------------------------|------|-------------|--------------------------|-------|
| | Total | \bar{x} % of Total BCs | SD | Total | \bar{x} % of Total BCs | SD | Total | \bar{x} % of Total BCs | SD |
| Minimal Response | 372 | 82.1 | 10.6 | 182 | 62.1** (p<.003) | 8.49 | 232 | 60.4** (p<.002) | 13.02 |
| Extended Response | 29 | 7 | 4.37 | 62 | 18** (p<.009) | 9.04 | 78 | 23** (p<.004) | 13.49 |

(\bar{x} difference of Pre-test → Post-test 1, and Pre-test → Post-test 2 significant at p<.05 level = *; significant at p<.01 level = **)

Table 8

Group B's use of minimal versus extended backchannels over time

| N = 10 | Pre-test | | | Post-test 1 | | | Post-test 2 | | |
|-------------------|---------------------|-------|--------------------------|-------------|--------------------|--------------------------|-------------|-------------------|--------------------------|
| | Type of Backchannel | Total | \bar{x} % of Total BCs | SD | Total | \bar{x} % of Total BCs | SD | Total | \bar{x} % of Total BCs |
| Minimal Response | 302 | 79.7 | 10.87 | 157 | 59.1** (p<.001) | 9.09 | 269 | 66.9* (p<.022) | 12.87 |
| Extended Response | 23 | 7.1 | 4.43 | 47 | 17.1** (p<.001) | 7.23 | 51 | 14.5* (p<.015) | 10.01 |

(\bar{x} difference of Pre-test \rightarrow Post-test 1, and Pre-test \rightarrow Post-test 2 significant at p<.05 level = *; significant at p<.01 level = **)

Table 9

Group Z's use of minimal versus extended backchannels over time

| N = 10 | Pre-test | | | Post-test 1 | | | Post-test 2 | | |
|-------------------|---------------------|-------|--------------------------|-------------|-------|--------------------------|-------------|-----------------|--------------------------|
| | Type of Backchannel | Total | \bar{x} % of Total BCs | SD | Total | \bar{x} % of Total BCs | SD | Total | \bar{x} % of Total BCs |
| Minimal Response | 321 | 74.4 | 11.3 | 241 | 75.5 | 10.79 | 398 | 82.4 | 5.6 |
| Extended Response | 48 | 11.2 | 6.55 | 27 | 10.8 | 5.41 | 36 | 7.4* (p<.04) | 3.31 |

(\bar{x} difference of Pre-test \rightarrow Post-test 1, and Pre-test \rightarrow Post-test 2 significant at p<.05 level = *; significant at p<.01 level = **)

In terms of the frequency by which minimal and extended backchannel types were used in this study, Figures 5 and 6 demonstrate respectively the differences between three groups. As seen in Figure 5, both the Explicit and Implicit groups showed a substantial decrease in their mean percentage of backchannels constituted by minimal backchannels from the Pre-test to Post-tests 1 and 2. The Explicit group decreased by 20% from the Pre-test to Post-test 1 and by 21.7% from the Pre-test to Post-test 2, while the Implicit Group decreased by 20.06% from the Pre-test to Post-test 1 and by 12.8% from the Pre-test to Post-test 2. Paired-samples t-tests showed that the

differences between the means of the Pre-test and Post-test 1 were strongly significant for both experimental groups: the Explicit group ($p < .003$) and Implicit group ($p < .001$). The differences between the means of the Pre-test and Post-test 2 were also found to be statistically significant for both experimental groups; however, the difference for the Explicit group, at the .01 level, was again strongly significant ($p < .002$), whereas the difference for the Implicit group was significant at the .05 level ($p < .022$). In stark contrast, the Control group increased in the mean percentage of total backchannels constituted by minimal responses from the Pre-test to Post-test 1 (+1.1%) and the Pre-test to Post-test 2 (+8%); however, these increases were not statistically significant.

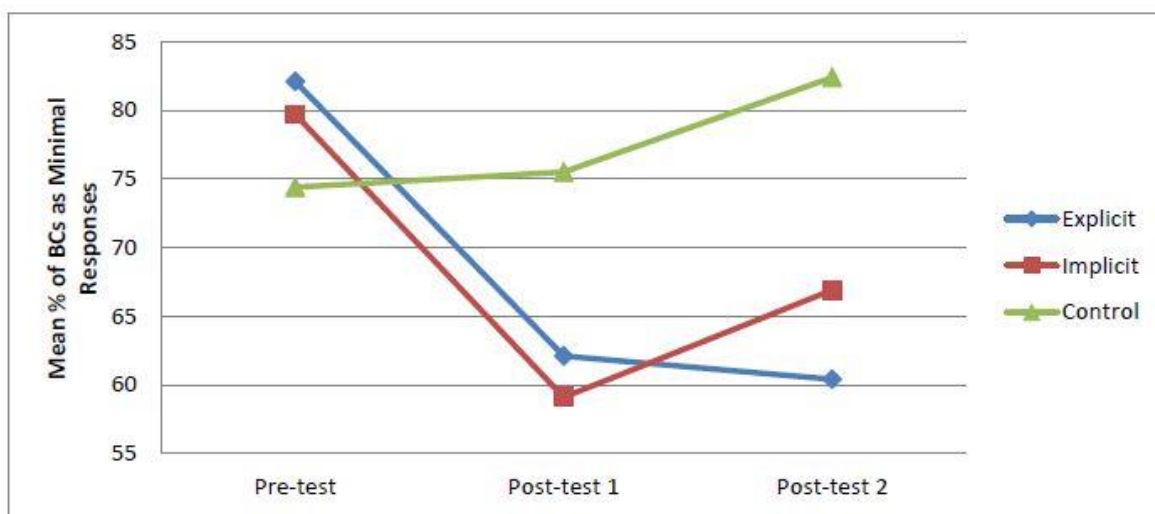


Figure 5. Proportions of minimal backchannels of the three groups over time

As Figure 6 reports, the Explicit and Implicit groups exhibited a considerable increase in the mean percentage of backchannels constituted by extended responses from the Pre-test to Post-tests 1 and 2. The mean percentage of the Explicit Group increased by 11% from the Pre-test to Post-test 1 and by 16% from the Pre-test to Post-test 2, whereas the Implicit Group increased by 10% from the Pre-test to Post-test 1 and by 7.4% from the Pre-test to Post-test 2. Paired-samples t-tests showed that the differences between the means of the Pre-test and Post-test 1 were strongly significant for the Explicit group ($p < .009$) and Implicit group ($p < .001$), and, once again, the difference was strongly significant for the Explicit group from the Pre-test to Post-test 2 ($p < .004$), while it was significant at the .05 level for the Implicit group ($p < .015$). Contrasting the path of the experimental groups, the Control group exhibited decreases in the mean percentage of

backchannels constituted by extended responses from the Pre-test to Post-test 1 (-.4%) and the Pre-test to Post-test 2 (-3.8%). The latter of which was statistically significant at the .05 level ($p < .04$).

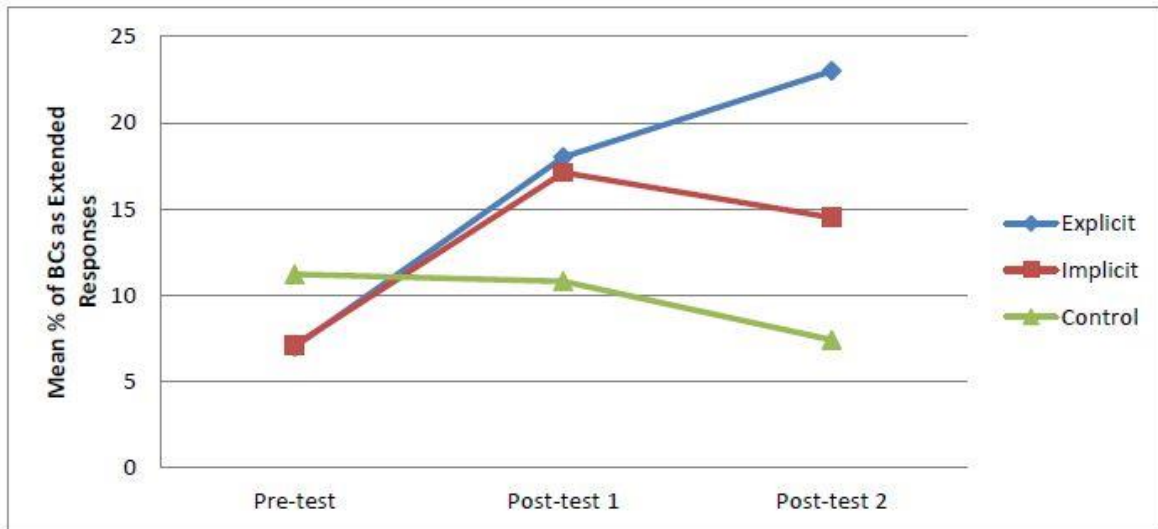


Figure 6. Proportions of extended backchannels of three groups over time

Discourse Contexts Favoring Backchannels

As the primary discourse context attracting backchannel is the final clausal boundary, the results in this section focus on the following areas: the mean percentages of final clause boundaries (CBs) eliciting backchannels (BCs), and the mean percentages of total backchannels constituted by backchannels occurring at final clause boundaries. Accordingly, Tables 10, 11 and 12 present the results of these features of conversations for Group A (Explicit), Group B (Implicit) and Group Z (Control) respectively at the three points of measurement in this study.

Table 10

Group A's backchannels at final clausal boundaries over time

| N = 10 | Pre-test | | Post-test 1 | | Post-test 2 | |
|---------------------------------|--------------------------------|-------------------------------|--------------------------------|-------------------------------|--------------------------------|-------------------------------|
| | \bar{x} % of Opps (SD) | \bar{x} % of BCs (SD) | \bar{x} % of Opps (SD) | \bar{x} % of BCs (SD) | \bar{x} % of Opps (SD) | \bar{x} % of BCs (SD) |
| Final Clausal Boundaries | 57.8 (17.78) | 52.68 (9.15) | 45.8* (17.73) | 66.27* (20) | 52.5 (9.47) | 58.82 (19.08) |

(\bar{x} difference of Pre-test \rightarrow Post-test 1, and Pre-test \rightarrow Post-test 2 significant at $p < .05$ level = *; significant at $p < .01$ level = **)**Table 11**

Group B's backchannels at final clausal boundaries over time

| N = 10 | Pre-test | | Post-test 1 | | Post-test 2 | |
|---------------------------------|--------------------------------|-------------------------------|--------------------------------|-------------------------------|--------------------------------|-------------------------------|
| | \bar{x} % of Opps (SD) | \bar{x} % of BCs (SD) | \bar{x} % of Opps (SD) | \bar{x} % of BCs (SD) | \bar{x} % of Opps (SD) | \bar{x} % of BCs (SD) |
| Final Clausal Boundaries | 62.1 (12.85) | 62.46 (8.58) | 42.1* (8.91) | 57.24 (12.44) | 50.3* (10.14) | 53.06 (14.69) |

(\bar{x} difference of Pre-test \rightarrow Post-test 1, and Pre-test \rightarrow Post-test 2 significant at $p < .05$ level = *; significant at $p < .01$ level = **)**Table 12**

Group Z's backchannels at final clausal boundaries over time

| N = 10 | Pre-test | | Post-test 1 | | Post-test 2 | |
|---------------------------------|--------------------------------|-------------------------------|--------------------------------|-------------------------------|--------------------------------|-------------------------------|
| | \bar{x} % of Opps (SD) | \bar{x} % of BCs (SD) | \bar{x} % of Opps (SD) | \bar{x} % of BCs (SD) | \bar{x} % of Opps (SD) | \bar{x} % of BCs (SD) |
| Final Clausal Boundaries | 65.6 (14.57) | 59.78 (9.8) | 44.4** (14.44) | 53.68 (17.47) | 55.7 (9.33) | 52.24 (8.87) |

(\bar{x} difference of Pre-test \rightarrow Post-test 1, and Pre-test \rightarrow Post-test 2 significant at $p < .05$ level = *; significant at $p < .01$ level = **)

Examining the three groups collectively, Figure 7 compares the three groups' mean percentages of final clause boundaries (CBs) eliciting backchannels (BCs), and Figure 8 compares the three groups' mean percentages of total backchannels constituted by backchannels occurring at final clause boundaries. While there was some variation in the range of each group, Figure 7 shows

that the general trend for all three groups was to send backchannels much less frequently at final clause boundary opportunities in Post-test 1 as compared to the Pre-test and then, in Post-test 2, to revert to a level closer to their original. The decrease in mean percentage from the Pre-test to the Post-test 1 was statistically significant for all three groups; however, the Implicit group was the only group to record a statistically significant decrease from the Pre-test to the Post-test 2.

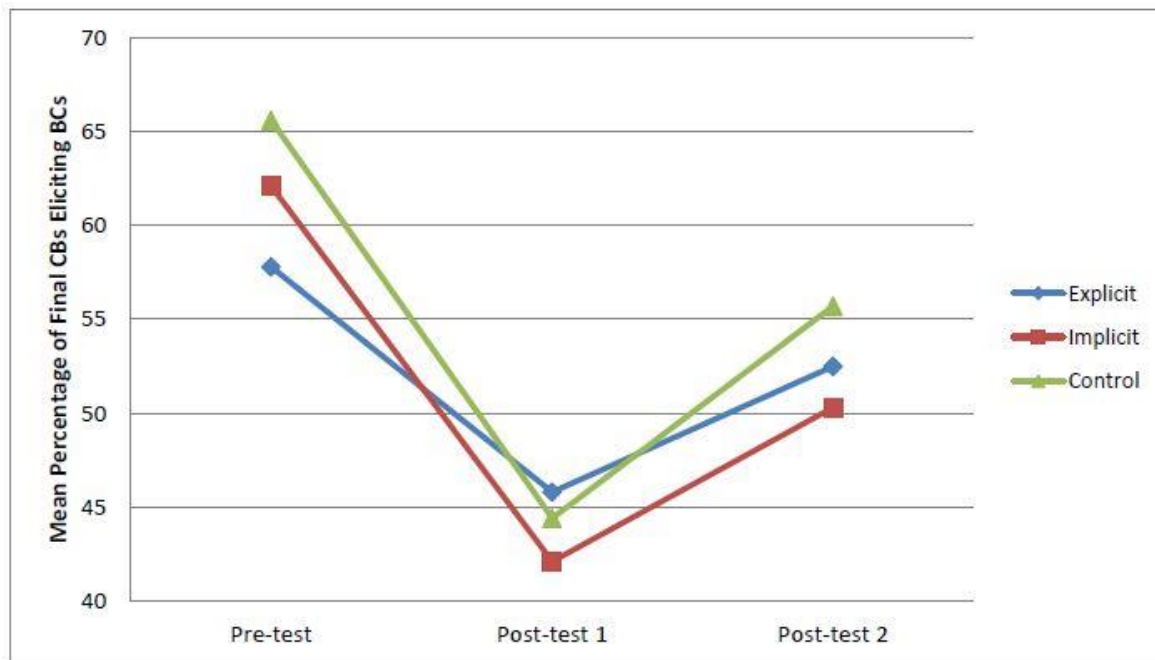


Figure 7. Proportions of final clause boundaries attracting backchannels over time

Concerning the proportions of total backchannels constituted by backchannels at final clause boundaries, Figure 8 demonstrates the disparity between the Explicit group and the other two groups. The mean percentage of the Explicit Group increased by 13.59% from the Pre-test to Post-test 1 and by 6.14% from the Pre-test to Post-test 2. In contrast, the Implicit group decreased from the Pre-test to Post-test 1 by 5.22% and by 9.4% from the Pre-test to Post-test 2, and the Control group decreased by 6.1% and 7.54% respectively.

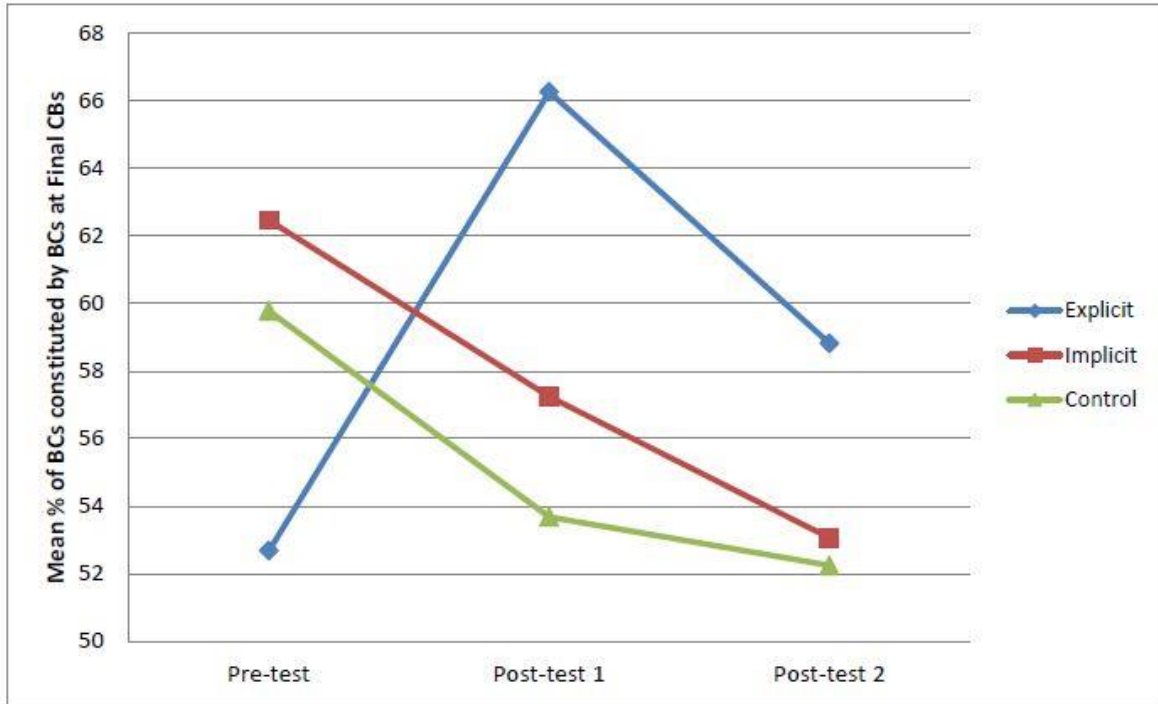


Figure 8. Proportions of BCs constituted by BCs at final clause boundaries over time

Simultaneous Speech Backchannels (SSBs)

Tables 13, 14 and 15 report on the use of Non-laughter simultaneous speech backchannels (SSBs) for Group A (Explicit), Group B (Implicit) and Group Z (Control) respectively at the three points of measurement in this study. Mean scores (with standard deviations) and the mean percentage of backchannels constituted by Non-laughter SSBs (with standard deviations) are presented.

Table 13

Group A's SSBs over time

| Types of SSBs | Pre-test | | | Post-test 1 | | | Post-test 2 | | |
|---------------|----------|----------------|-------------------------|-------------|----------------|-------------------------|-------------|----------------|-------------------------|
| | Total | \bar{x} (SD) | \bar{x} % of BCs (SD) | Total | \bar{x} (SD) | \bar{x} % of BCs (SD) | Total | \bar{x} (SD) | \bar{x} % of BCs (SD) |
| Non-laughter | 77 | 7.7 (3.83) | 17.48 (7.21) | 55 | 5.5** (3.72) | 18.69 (8.76) | 51 | 5.1** (3.14) | 14.12* (7.91) |

(\bar{x} difference of Pre-test → Post-test 1, and Pre-test → Post-test 2 significant at $p < .05$ level = *; significant at $p < .01$ level = **)

Table 14

Group B's SSBs of over time

| Types of SSBs | Pre-test | | | Post-test 1 | | | Post-test 2 | | |
|---------------|----------|----------------|-------------------------|-------------|----------------|-------------------------|-------------|----------------|-------------------------|
| | Total | \bar{x} (SD) | \bar{x} % of BCs (SD) | Total | \bar{x} (SD) | \bar{x} % of BCs (SD) | Total | \bar{x} (SD) | \bar{x} % of BCs (SD) |
| Non-laughter | 51 | 5.1 (5.45) | 13.08 (9.81) | 47 | 4.7 (3.34) | 17.39 (10.83) | 58 | 5.8 (3.01) | 15.41 (8.78) |

(\bar{x} difference of Pre-test \rightarrow Post-test 1, and Pre-test \rightarrow Post-test 2 significant at $p < .05$ level = *; significant at $p < .01$ level = **)

Table 15

Group Z's SSBs over time

| Types of SSBs | Pre-test | | | Post-test 1 | | | Post-test 2 | | |
|---------------|----------|----------------|-------------------------|-------------|----------------|-------------------------|-------------|----------------|-------------------------|
| | Total | \bar{x} (SD) | \bar{x} % of BCs (SD) | Total | \bar{x} (SD) | \bar{x} % of BCs (SD) | Total | \bar{x} (SD) | \bar{x} % of BCs (SD) |
| Non-laughter | 61 | 6.1 (3.03) | 13.66 (5.68) | 61 | 6.1 (1.45) | 19.86** (4.74) | 75 | 7.5 (2.84) | 15.9 (5.99) |

(\bar{x} difference of Pre-test \rightarrow Post-test 1, and Pre-test \rightarrow Post-test 2 significant at $p < .05$ level = *; significant at $p < .01$ level = **)

Looking at the data concerning Non-laughter SSBs collectively, Figure 9 shows that the general paths of the Implicit and Control groups mirrored one another by starting comparatively low, remaining fairly stable from the Pre-test to Post-test 1, and then sharply increasing from Post-test 1 to Post-test 2 (by 11 and 14 respectively). The Explicit group, in contrast, showed a sustained decline, of 12 from the Pre-test to Post-test 1, and then another decrease of 4 from Post-test 1 to Post-test 2. These decreases were found to be strongly significant ($p < .001$ and $p < .001$ respectively).

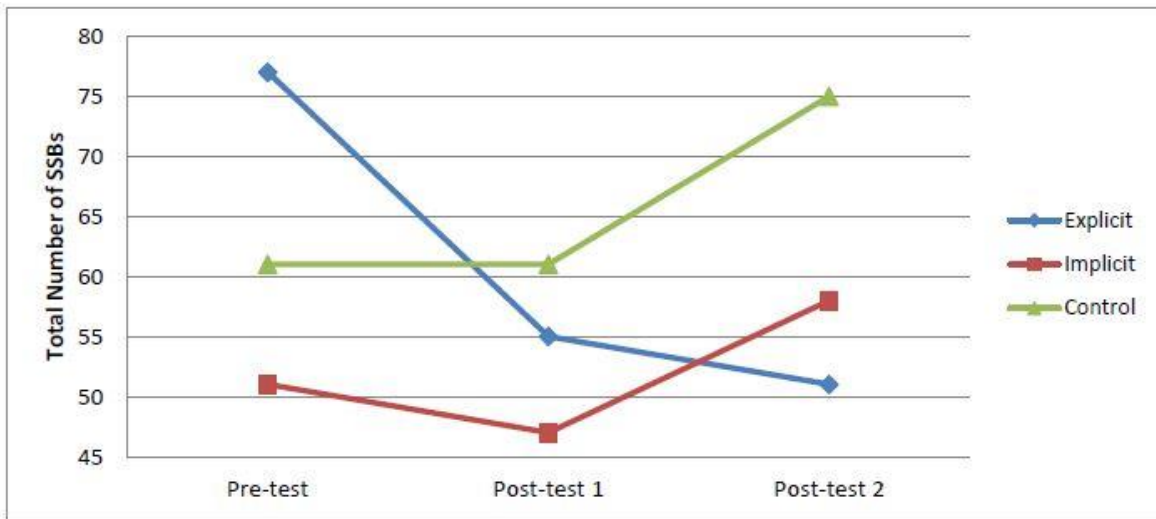


Figure 9. Non-laughter SSBs of the three groups over time

NESs Observers' Perceptions

Tables 16, 17 and 18 report on the NES observers' perceptions (according to their ratings on the modified version of Hecht's conversational satisfaction questionnaire described above) of the conversational performances for Group A (Explicit), Group B (Implicit) and Group Z (Control) respectively at the three points of measurement in this study. To provide a general idea as to how the overall ratings compared across participant groups at the three tests, the researcher has divided the 17 items in the questionnaire into two groups distinguished by the positive and negative connotations associated with each rating (see Hecht, 1978). For instance, in the items in group one (1, 2, 3, 4, 5, 6, 7, 10, 11, 13 and 16), a low score would indicate a desirable effect, whereas for the items in group two (8, 9, 12, 14, 15 and 17), a high score would convey a desirable effect. The random presentation of the video recorded conversations ensured that the NES observers did not have any knowledge as to the time-line of the conversations or which group each JEFL belonged to.

Table 16

NES observers' perceptions of Group A over time

| Items on the Questionnaire N = 10 Rating Scale: 1 (strongly agree) to 7 (strongly disagree) | Pre | | Post 1 | | Post 2 | |
|--|-----------|------|-----------|-----|-----------|-----|
| | \bar{x} | SD | \bar{x} | SD | \bar{x} | SD |
| Group 1 items: | | | | | | |
| 1. The JEFL let his/her partner know that the partner was communicating effectively. | 2.81 | .80 | 2.23** | .66 | 2.24* | .47 |
| 2. The JEFL showed his/her partner that they understood what their partner said. | 2.81 | .84 | 2.2** | .69 | 2.4 | .62 |
| 3. The JEFL showed that they were listening attentively to what their partner said. | 2.21 | .6 | 1.83 | .37 | 1.93* | .38 |
| 4. The JEFL expressed a lot of interest in what their partner had to say. | 2.64 | .71 | 2.02* | .38 | 2.15* | .53 |
| 5. The conversation went smoothly. | 3.28 | 1.13 | 2.2** | .57 | 2.52** | .56 |
| 6. The JEFL encouraged partner to continue talking. | 2.82 | .95 | 2.55 | .66 | 2.5 | .61 |
| 7. The feelings that the JEFL expressed by means of listening feedback during the conversation seemed <i>authentic</i> (...). | 3.22 | .84 | 2.32** | .75 | 2.46** | .48 |
| 10. The JEFL was polite. | 2.06 | .47 | 1.5* | .22 | 1.44** | .26 |
| 11. The JEFL appeared warm and friendly. | 2.34 | .66 | 1.83* | .32 | 1.81* | .42 |
| 13. The JEFL appeared interested and concerned. | 2.73 | .72 | 2.18* | .39 | 2.13* | .34 |
| 16. When the JEFL did not understand, they were able to clearly convey this to their conversational partner with their listening feedback. | 3.19 | .73 | 2.39* | .61 | 2.36* | .39 |
| Group 2 items: | | | | | | |
| 8. The JEFL seemed impatient. | 6.1 | .38 | 6.36* | .2 | 6.34 | .28 |
| 9. The JEFL seemed cold and unfriendly. | 6.41 | .34 | 6.46 | .22 | 6.5 | .27 |
| 12. The JEFL was impolite. | 6.52 | .29 | 6.74* | .17 | 6.75* | .2 |
| 14. The JEFL interrupted their partner at times. | 6.18 | .29 | 6.59* | .13 | 6.59** | .31 |
| 15. The JEFL seemed to want to avoid speaking. | 5.14 | .86 | 5.99** | .55 | 5.76* | .27 |
| 17. The JEFL's listening behavior seemed inadequate in some ways. | 5.1 | .72 | 5.89** | .54 | 5.7* | .69 |

(\bar{x} difference of Pre-test \rightarrow Post-test 1, and Pre-test \rightarrow Post-test 2 significant at $p < .05$ level = *; significant at $p < .01$ level = **)

Table 17

NES observers' perceptions of Group B over time

| Items on the Questionnaire N = 10 Rating Scale: 1 (strongly agree) to 7 (strongly disagree) | Pre | | Post 1 | | Post 2 | |
|--|-----------|-----|-----------|-----|-----------|-----|
| | \bar{x} | SD | \bar{x} | SD | \bar{x} | SD |
| Group 1 items: | | | | | | |
| 1. The JEFL let his/her partner know that the partner was communicating effectively. | 3.15 | .79 | 2.49* | .49 | 2.52** | .52 |
| 2. The JEFL showed his/her partner that they understood what their partner said. | 3.1 | .81 | 2.59 | .5 | 2.59* | .52 |
| 3. The JEFL showed that they were listening attentively to what their partner said. | 2.54 | .63 | 2.07 | .29 | 2.12* | .31 |
| 4. The JEFL expressed a lot of interest in what their partner had to say. | 3.06 | .7 | 2.5* | .53 | 2.52* | .61 |
| 5. The conversation went smoothly. | 3.89 | .97 | 2.98** | .72 | 2.89* | .73 |
| 6. The JEFL encouraged partner to continue talking. | 3.12 | .65 | 2.63 | .48 | 2.9 | .65 |
| 7. The feelings that the JEFL expressed by means of listening feedback during the conversation seemed <i>authentic</i> (...). | 3.5 | .84 | 2.56* | .52 | 2.71** | .6 |
| 10. The JEFL was polite. | 2.06 | .24 | 1.54** | .2 | 1.65* | .33 |
| 11. The JEFL appeared warm and friendly. | 2.4 | .58 | 1.9** | .33 | 1.97 | .45 |
| 13. The JEFL appeared interested and concerned. | 2.94 | .68 | 2.41 | .46 | 2.47 | .46 |
| 16. When the JEFL did not understand, they were able to clearly convey this to their conversational partner with their listening feedback. | 3.84 | .82 | 2.69** | .65 | 2.88** | .58 |
| Group 2 items: | | | | | | |
| 8. The JEFL seemed impatient. | 5.97 | .54 | 6.31 | .39 | 6.29 | .34 |
| 9. The JEFL seemed cold and unfriendly. | 6.18 | .33 | 6.34 | .38 | 6.45 | .2 |
| 12. The JEFL was impolite. | 6.3 | .29 | 6.63 | .31 | 6.53 | .35 |
| 14. The JEFL interrupted their partner at times. | 6.52 | .24 | 6.54 | .14 | 6.56 | .17 |
| 15. The JEFL seemed to want to avoid speaking. | 4.91 | .82 | 5.77* | .71 | 5.66 | .57 |
| 17. The JEFL's listening behavior seemed inadequate in some ways. | 4.61 | .9 | 5.76* | .48 | 5.59** | .51 |

(\bar{x} difference of Pre-test → Post-test 1, and Pre-test → Post-test 2 significant at p<.05 level = *; significant at p<.01 level = **)

Table 18

NES observers' perceptions of Group Z over time

| Items on the Questionnaire N = 10 Rating Scale: 1 (strongly agree) to 7 (strongly disagree) | Pre | | Post 1 | | Post 2 | |
|--|-----------|------|-----------|-----|-----------|-----|
| | \bar{x} | SD | \bar{x} | SD | \bar{x} | SD |
| Group 1 items: | | | | | | |
| 1. The JEFL let his/her partner know that the partner was communicating effectively. | 2.6 | .8 | 2.19 | .57 | 2.19 | .44 |
| 2. The JEFL showed his/her partner that they understood what their partner said. | 2.71 | .79 | 2.24* | .68 | 2.34* | .46 |
| 3. The JEFL showed that they were listening attentively to what their partner said. | 2.55 | 1.17 | 2.04 | .44 | 2.06 | .41 |
| 4. The JEFL expressed a lot of interest in what their partner had to say. | 2.6 | .83 | 2.34 | .65 | 2.47 | .57 |
| 5. The conversation went smoothly. | 3.2 | 1.01 | 2.57* | .71 | 2.65* | .46 |
| 6. The JEFL encouraged partner to continue talking. | 2.85 | .97 | 2.45 | .73 | 2.75 | .5 |
| 7. The feelings that the JEFL expressed by means of listening feedback during the conversation seemed <i>authentic</i> (...). | 3.24 | .8 | 2.55** | .67 | 2.56* | .52 |
| 10. The JEFL was polite. | 2.03 | .4 | 1.59* | .24 | 1.55* | .23 |
| 11. The JEFL appeared warm and friendly. | 2.3 | .47 | 1.98* | .33 | 2.06* | .43 |
| 13. The JEFL appeared interested and concerned. | 2.56 | .54 | 2.34 | .51 | 2.41 | .48 |
| 16. When the JEFL did not understand, they were able to clearly convey this to their conversational partner with their listening feedback. | 3.15 | .79 | 2.66* | .82 | 2.74* | .54 |
| Group 2 items: | | | | | | |
| 8. The JEFL seemed impatient. | 6.06 | .5 | 6.36* | .4 | 6.29 | .54 |
| 9. The JEFL seemed cold and unfriendly. | 6.22 | .64 | 6.47 | .26 | 6.48 | .3 |
| 12. The JEFL was impolite. | 6.44 | .25 | 6.58 | .19 | 6.63* | .13 |
| 14. The JEFL interrupted their partner at times. | 6.15 | .42 | 6.25 | .53 | 6.31 | .24 |
| 15. The JEFL seemed to want to avoid speaking. | 5.33 | .67 | 5.46 | .61 | 5.7 | .48 |
| 17. The JEFL's listening behavior seemed inadequate in some ways. | 5.35 | .74 | 5.7 | .62 | 5.64 | .46 |

(\bar{x} difference of Pre-test → Post-test 1, and Pre-test → Post-test 2 significant at p<.05 level = *; significant at p<.01 level = **)

As the items in the questionnaire have been separated into the two groups described above, it is possible to compare the sum totals of average responses to items in each group over time and between participant groups in this study. Figure 10 provides an illustration comparing the items in group one between participant groups over time. In this line-graph, the figures on the y-axis represent the sum of the average ratings in response to the group of items, while the x-axis again shows the differences in performance over time. The line-graph demonstrates that the general paths of the three participant groups were quite similar in that the NES observers' perceptions greatly improved from the Pre-test to Post-test 1, and then experienced a slight decrease from Post-test 1 to Post-test 2. Whereas the Control group (Group Z) followed a similar path generally, the experimental groups experienced considerably greater improvements in ratings from the Pre-test to Post-tests 1 and 2. Specifically, the average NES observers' ratings improved by 7.24 from the Pre-test to Post-test 1 and by 6.39 from the Pre-test to Post-test 2 for the Implicit group, by 6.82 from the Pre-test to Post-test 1 and by 6.02 from the Pre-test to Post-test 2 for the Explicit group, and by 4.84 from the Pre-test to Post-test 1 and by 4.01 from the Pre-test to Post-test 2 for the Control group.

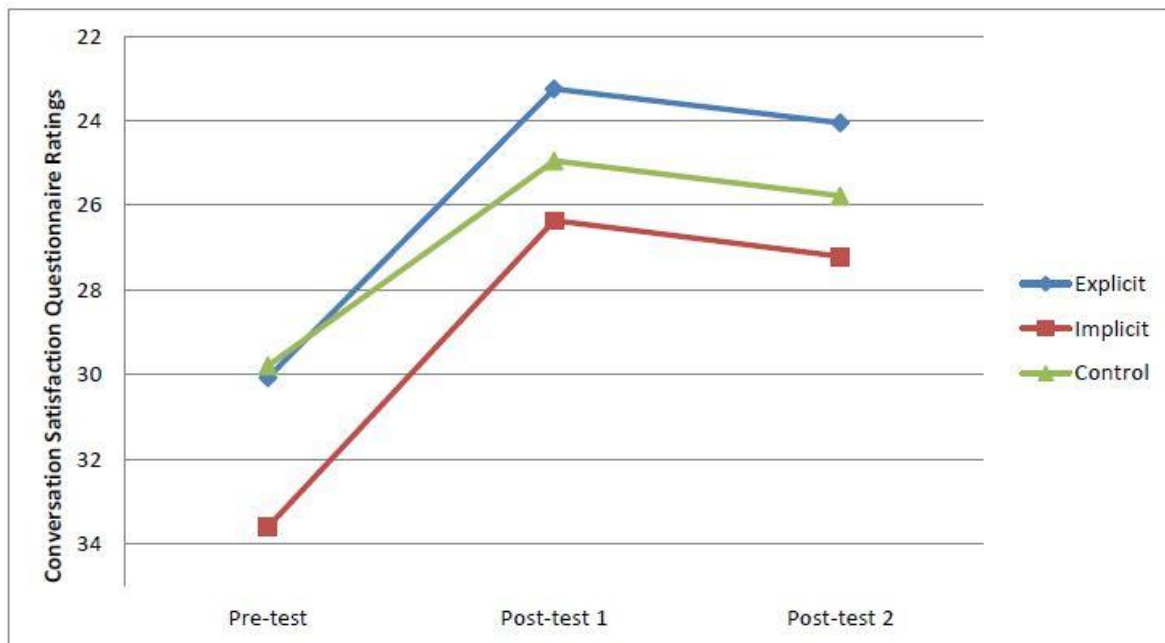


Figure 10. NES Observers' perceptions of three groups: Group 1 items

Figure 11 demonstrates the group differences pertaining to the items in group two over time.

Similar to the findings in group one above, the line-graph below also demonstrates that the NES observers' perceptions improved for all three groups from the Pre-test to Post-test 1. The experimental groups also then showed a slight decrease in positive perceptions from Post-test 1 to Post-test 2, while the Control group showed a slight increase. Overall, however, the experimental groups experienced a substantially greater improvement in ratings from the Pre-test to the Post-tests 1 and 2. That is, the average NES observers' ratings improved by 2.86 from the Pre-test to Post-test 1 and by 2.59 from the Pre-test to Post-test 2 for the Explicit group, by 2.54 from the Pre-test to Post-test 1 and by 2.19 from the Pre-test to Post-test 2 for the Implicit group, and by 1.27 from the Pre-test to Post-test 1 and by 1.5 from the Pre-test to Post-test 2 for the Control group.

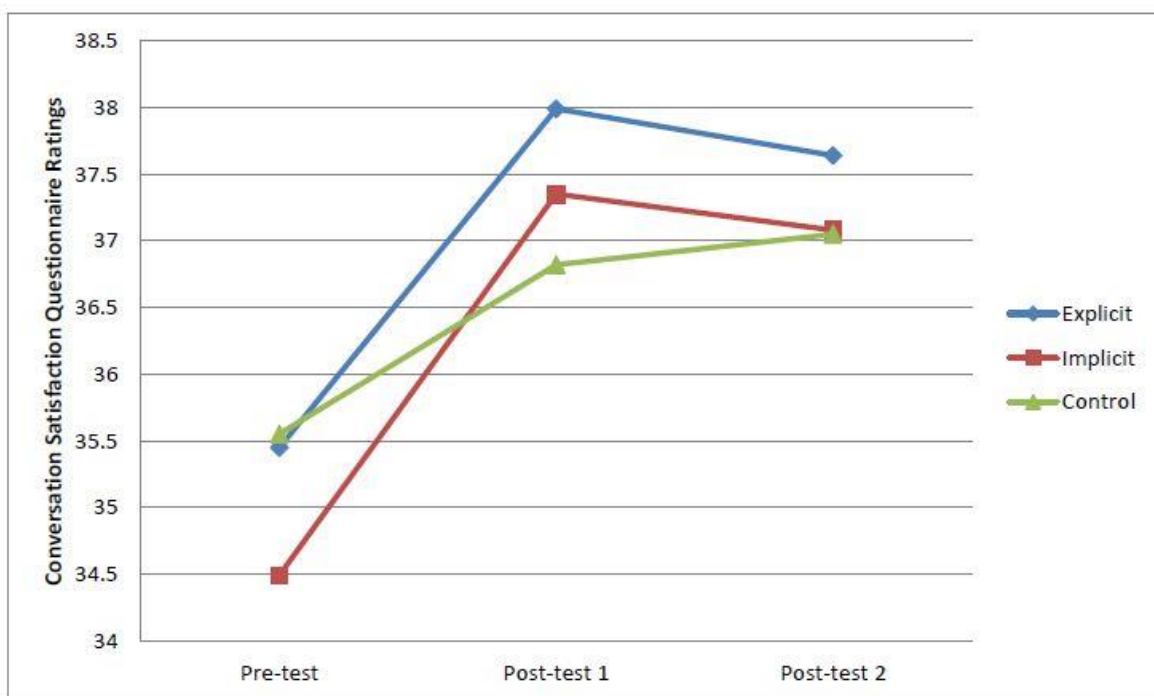


Figure 11. NESs Observers' perceptions of three groups: Group 2 items

Interviews

This section presents the quantitative and qualitative data elicited by the interviews of the participant groups at three points of measurement in this study (i.e., the Pre-test, Post-test 1 and Post-test 2). These findings are divided into three subsections. The first subsection reports on the JEFs' development in dealing with situations of non-understanding. It should be noted that

disagreement situations were excluded from this analysis because they occurred only four times in the 90 conversations. Lastly, the final two subsections present the JEFs' opinions regarding their rate of progress at the time of Post-test 1 and 2 respectively.

Dealing with Situations of Non-understanding

Tables 19, 20 and 21 report on the reactions at points of non-understanding for Group A (Explicit), Group B (Implicit) and Group Z (Control) respectively at the three points of measurement in this study. Data are presented in two areas: the use of unconventional backchannels and the use of conversational repair strategies (which are further broken down into minimal versus lengthier conversational repair strategies).

Table 19

Group A's reactions at points of non-understanding over time

| N = 10 | NONUs | Unconventional BCs | | Conversational Repair Strategies | | | |
|--------|-------|--------------------|---------------------------|----------------------------------|---------------------------|-----------------------|---------------------------|
| | | | | Minimal BCs | | Lengthier expressions | |
| | | Total | \bar{x} % of NONUs (SD) | Total | \bar{x} % of NONUs (SD) | Total | \bar{x} % of NONUs (SD) |
| Pre | 49 | 36 | 69.41 (11.66) | 13 | 30.6 (11.66) | 0 | 0 (0) |
| Post 1 | 32 | 10 | 25.17** (28.55) | 13 | 44** (11.2) | 9 | 30.83** (25.47) |
| Post 2 | 30 | 12 | 31.38** (22.95) | 11 | 50.93* (30.46) | 7 | 17.69* (19.79) |

(\bar{x} difference of Pre-test → Post-test 1, and Pre-test → Post-test 2 significant at p<.05 level = *; significant at p<.01 level = **)

Table 20

Group B's reactions at points of non-understanding over time

| N = 10 | NONUs | Unconventional BCs | | Conversational Repair Strategies | | | |
|--------|-------|--------------------|---------------------------|----------------------------------|---------------------------|-----------------------|---------------------------|
| | | | | Minimal BCs | | Lengthier expressions | |
| | | Total | \bar{x} % of NONUs (SD) | Total | \bar{x} % of NONUs (SD) | Total | \bar{x} % of NONUs (SD) |
| Pre | 46 | 34 | 73.17 (11.82) | 10 | 23.17 (10.84) | 2 | 3.67 (7.77) |
| Post 1 | 31 | 19 | 58.83 (31.27) | 9 | 31.17 (19.82) | 3 | 10 (21.08) |
| Post 2 | 29 | 20 | 68.33 (22.5) | 8 | 28.33 (18.92) | 2 | 3.33 (10.54) |

(\bar{x} difference of Pre-test \rightarrow Post-test 1, and Pre-test \rightarrow Post-test 2 significant at $p < .05$ level = *; significant at $p < .01$ level = **)

Table 21

Group Z's reactions at points of non-understanding over time

| N = 10 | NONUs | Unconventional BCs | | Conversational Repair Strategies | | | |
|--------|-------|--------------------|---------------------------|----------------------------------|---------------------------|-----------------------|---------------------------|
| | | | | Minimal BCs | | Lengthier expressions | |
| | | Total | \bar{x} % of NONUs (SD) | Total | \bar{x} % of NONUs (SD) | Total | \bar{x} % of NONUs (SD) |
| Pre | 44 | 28 | 62.83 (16.35) | 14 | 33 (15.15) | 2 | 4.17 (9) |
| Post 1 | 25 | 15 | 59.17 (16.87) | 9 | 37.5 (21.25) | 1 | 3.33 (10.54) |
| Post 2 | 27 | 17 | 62.5 (22.65) | 9 | 32.5 (24.99) | 1 | 5 (15.81) |

(\bar{x} difference of Pre-test \rightarrow Post-test 1, and Pre-test \rightarrow Post-test 2 significant at $p < .05$ level = *; significant at $p < .01$ level = **)

Figures 12 and 13 illustrate the data collectively. First, in describing Figure 12, the y-axis corresponds to the mean percentage of non-understanding situations (NONUs) eliciting unconventional responses (i.e., conveying understanding and agreement when, in fact, they had not understood), while the x-axis demonstrates the differences in performance longitudinally. Accordingly, it can be seen that the Explicit group displayed the greatest initial and sustained improvement in this area. The Implicit group exhibited a moderate decrease initially at Post-test 1, but this level of improvement was not maintained at Post-test 2. The performance of the Control group, in comparison, remained fairly constant throughout.

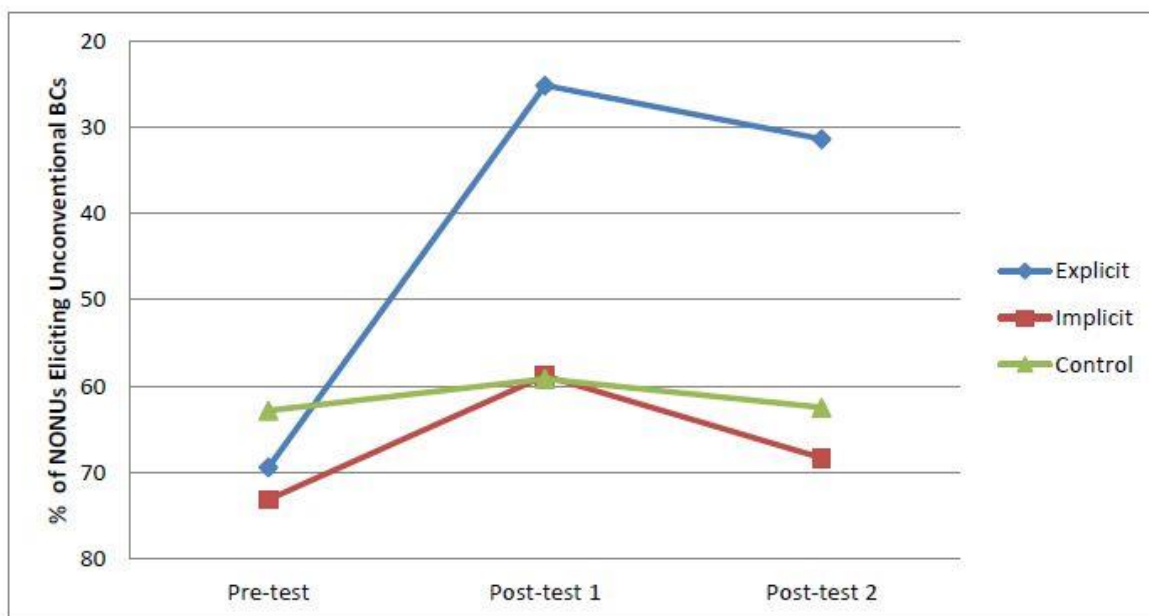


Figure 12. NONUs eliciting unconventional backchannels

Second, Figure 13 focuses on the three groups' ability to use the lengthier conversational repair strategies (CRSs). Accordingly, the figures on the y-axis represent the mean percentage of non-understanding situations eliciting lengthier conversational repair strategies for each group, while the x-axis again demonstrates the differences in performance longitudinally. Once again, the Explicit group showed the greatest and most sustained improvement, while the Implicit group showed a modest improvement initially at Post-test 1 but then a decrease at Post-test 2. The Control group again did not show any noticeable change over time.

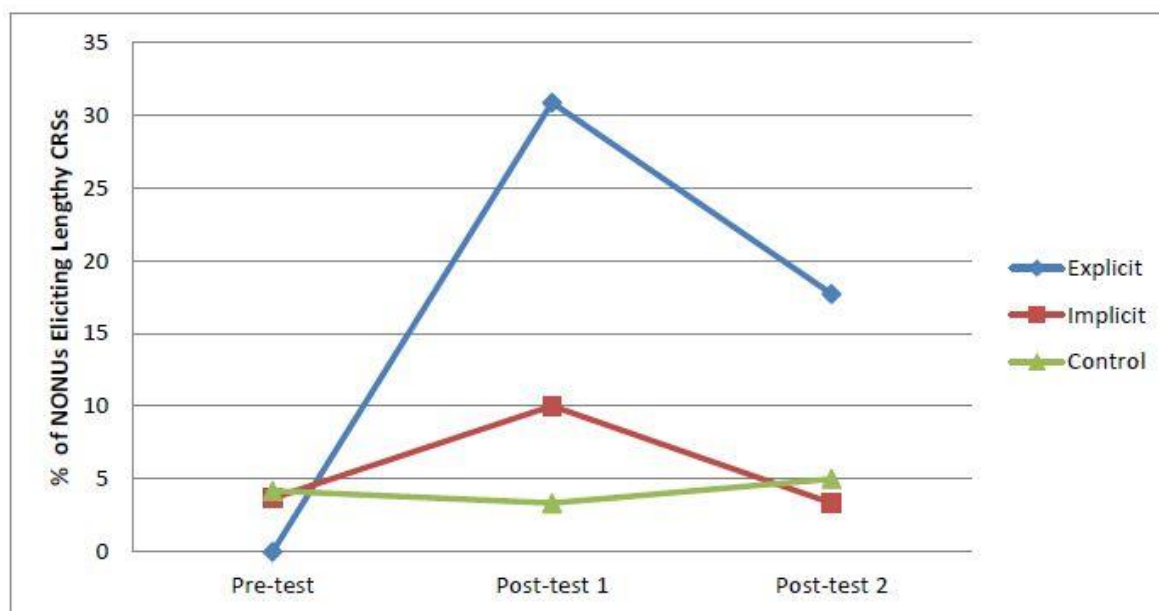


Figure 13. NONUs eliciting conversational repair strategies

Perceived Development from the Pre-test to Post-test 1

One common feeling for members of all three groups in the qualitative data of Post-test 1 was a general feeling of increased confidence due to the experience of communicating with a NES in an authentic setting. Unsurprisingly, responses from members of the Implicit group reflected that the benefits of their exposure to NESs went beyond the Pre-test and also occurred as part of the Implicit instruction they received. Concerning this treatment, the following responses show various members of this group pointing to general areas of improvement and the methods they used to achieve them:

Kouki: I tried speak more like native speaker style, and give bigger reactions.

Meo: I tried to copy natives as much as I could. From first conversation, I know I have to speak more and ask more questions, and I tried to do this.

Mayumi: The best thing that helped me was comparing what I did with what the natives did. I learned that making mistakes is not as important as keep talking.

Shio: Natives make me feel relaxed and teach me that I must speak more and try to give my opinions sometimes.

It appears, then, that several members of the Implicit group were trying to adopt the

conversational style of NESs, which they seemed to equate with speaking more.

When commenting on their Post-test 1 performances, members of the Explicit group also mentioned the instruction they received as a reason for their perceived improvements. These members, however, seemed to be able to specify in greater detail and depth the areas of their development related to what they had learned in class. The following excerpts demonstrate this group's awareness in terms of the frequency and placement by which they sent backchannels, particularly SSBs:

Aria: I tried to give aizuchi (the Japanese laymen's word for backchannel) mostly only when she finished her main idea like class taught me.

Kazuya: In first conversation, I did not even think about it, but now I was thinking how much aizuchi and when I should give.

Saya: I knew I shouldn't aizuchi so much, especially when she is speaking, so I tried to give at the end of sentences.

As demonstrated by the following responses, this group's learning awareness also extended to include other areas such as variability and WTC:

Michiko: Before like too passive, I didn't try to speak enough, and maybe I always give same boring aizuchi. Now, I'm using different aizuchi and sharing my thinking and asking questions.

Takanori: I feel better in second conversation because I can speak more and ask questions to my partner.

Miya: I used the FOQ method (Statement + Fact or Opinion or Question as a response strategy) we learned in class. It helped me keep the conversation continue.

Regarding function, several respondents commented that learning how to use conversational repair strategies facilitated communication, as follows:

Chieko: When I didn't understand before, I did not know what to do, and conversation became broken. Because of lesson, I now had a confidence what to do and what I can say in this situation.

Kazuya: I see aizuchi style is different about interrupting. Little aizuchi don't need to give so much when he speaking, but more important time like when I need help or something, I should speak this feeling.

Saya: I never changed my aizuchi at all, but in this conversation, I tried to change this and asked the person to explain more until I understand.

Aria: Big thing for my better speaking is I know better what I should say in many different situations, especially when I don't understand.

Perceived Development from Post-test 1 to Post-test 2

Analyzing the qualitative data of the interviews in Post-test 2, a pattern to emerge in the qualitative data of Post-test 2 was a general belief among the experimental groups that their communicative competence had decreased since Post-test 1. The following responses indicate that these members largely attributed their regression to not having had instruction for a great length of time (i.e., eight weeks).

Explicit group:

Hika: I pretended to understand like before. I should have asked for help, but it was too fast and I was too late.

Kazuya: Basically, I was too slow to reaction in conversation, and my aizuchi was too simple like at beginning. I needed more different kinds like my paper I have shows.

Akie: I think I probably gave too many aizuchi, maybe even a lot when we was speaking. And, I did not speak enough or ask enough questions.

Miya: I can't have good conversation like before because I did not study for it. I should look at my class papers.

Implicit group:

Meo: I did not have any practice chance for this conversation, so I forget everything from before.

Sachi: I forgot the habits of native speakers because I didn't have a chance to speak to someone.

Kouki: It was so long since I talked to native that I became very nervous again and couldn't focus enough on the conversation.

The responses of members of the Explicit group above again seemed to include greater depth and understanding in explaining where they went wrong, and many implied that the information they

needed to perform better could be found in one of the many hand-outs they had received in class. In comparison, members of the Control group did not report any major changes regarding their conversational performances from Post-test 1 to Post-test 2; however, the following replies suggested slight improvements in confidence:

Control group:

Tomomi: This was my third conversation, so I get less nervous every time.

Akanori: I could speak more because I knew what to do from other conversations.

Discussion: Improvements in L2 English Backchannel Behavior

Approximating the Listenership of NESs

While both experimental groups adopted more native-like listening behavior after treatment, the Explicit group (Group A) generally outpaced the Implicit group (Group B) in this area. The Explicit group's improvements were found to be strongly significant (i.e., at the .01 level) from the Pre-test to Post-test 1 and also through to Post-test 2 in areas of frequency, variability and SSBs. In comparison, a statistically significant improvement for the Implicit group was observed at the .05 level in the area of frequency from the Pre-test to Post-test 1; however, it was not evident at the time of Post-test 2. In the area of variability, the Implicit group's improvements were found to be strongly significant from the Pre-test to Post-test 1 (at the .01 level) and significant at the .05 level from the Pre-test to Post-test 2. Regarding the Control group (Group Z), no significant improvements were found as expected; however, in one of the subcategories of variability, the proportion of backchannels constituted by Extended responses, the Control group's performance regressed as their mean percentage decreased significantly (at the .05 level) from the Pre-test to Post-test 2.

Notably, there was one area of analysis in which a statistically significant decline was found for all three groups: backchannels provided at final clausal boundary opportunities (see Figure 7). This was unexpected because this discourse context, which has been identified in the literature as the most common discourse context in English, was explicitly taught to members of Group A. However, while all three groups sent significantly less backchannels at final clausal boundaries from the Pre-test to Post-test 1, this decline was only sustained for the Implicit group in Post-test 2. Although the across-the-board decline between the Pre-test and Post-test 1 is somewhat

difficult to explain, the fact that Post-test 2 frequencies in this area (i.e., after no subsequent treatment) returned to near original Pre-test levels for two of the three groups suggests that this discrepancy may have been an aberration. The fact that the three groups were starting at approximately the same point and the Explicit group consistently sent the most backchannels at this discourse context thereafter would seem to suggest that they achieved the highest degree of success in this area also. This is supported by the fact that the Explicit group showed marked increases in the proportions of total backchannels constituted by backchannels at final clause boundaries from the Pre-test to Post-tests 1 and 2, while the Implicit and Control groups showed a steady decline. Finally, the fact that the Implicit and Control groups' performances consistently deteriorated in this area (from the Pre-test to Post-tests 1 and 2) offers another example in which a lack of treatment may be connected with a deterioration of skills.

The overall findings, which showed that explicit treatment greatly improved JEFs' backchannel behavior, seem to go against the conventional wisdom that L2 backchannels are particularly difficult to learn in a way that is not completely implicit due to their spontaneity and automaticity in real-time conversations (Cutrone 2005; Ward et al. 2007). As most JEFs had never previously even considered their L2 backchannel behavior, it appears that raising their awareness of this all-important feature of language for the first time was enough to make a dramatic impact on their L2 conversational behavior.

There was also evidence in the JEFs' interview data to suggest strong biases towards the English used by NESs. This is consistent with the observations of Suzuki (2010), who pointed out that teacher training programs in Japan still present only standard American and/or standard British English as a single normative variety of English. Several JEFs attributed their perceived improvements in their post-treatment interviews to noticing differences between their own conversational output and NESs' output. The fact that learners consciously perceived the difference between the feature in their target language and their own can be said to be evidence of "noticing the gap" (Schmidt & Frota, 1986, p. 311). Clearly, the resolution of such an enormous issue as to which model of English should be used in Japan is beyond the focus of this current investigation; however, the writer strongly believes that this aspect of the study has touched upon an important issue that seems to be in dire need of re-examination in Japan (Jenkins, 2003; Seidlhofer, 2004; Widdowson, 1997).

Conversational Involvement

In terms of conversational involvement, both experimental groups showed benefits of instruction; however, the Explicit group (Group A) again generally outpaced the Implicit one (Group B). The Explicit group's improvements were found to be strongly significant from the Pre-test to Post-test 1 in all three subcategories of Involvement in Conversation: WTC scores, amount spoken as primary speaker, and number of follow-up questions asked. The Explicit group's initial improvements in this area varied among the sub-skills at the delayed Post-test 2. That is, from the Pre-test to Post-test 2, the improvement in mean WTC scores was again found to be significant at the .01 level, while the increase in mean number of questions was significant at the .05 level, and the difference in mean number of words was not found to be statistically significant. In comparison, the improvements in mean WTC scores for the Implicit group were found to be strongly significant from the Pre-test to Post-test 1 and from the Pre-test to Post-test 2; however, no significant improvements were evident in the other two subcategories. Finally, no significant improvements for the Control group were found in this area; however, as was the case with the Variability category, the Control group's performance significantly regressed (at the .05 level) in one of the subcategories of this skill-set: mean number of questions asked. This further suggests that a lack of treatment may eventually lead to skill loss.

Some of the results in this area were predictable. For instance, it was not surprising at all that the group that received explicit instruction was the only group to show improvements in the use of conversational management techniques. These improvements occurred not only because the JEFLs in this group were taught specific expressions to use in certain situations, but more importantly because they were explicitly made aware of the negative perceptions that their cross-cultural interlocutors would have of them if they persisted to speak English with a Japanese mind-set (i.e., negatively transfer their backchannel behavior from their L1 to L2 English). As reflected by their WTC scores and interview responses over time, members of the Explicit group (Group A) seemed to have been motivated to change their behaviour in order make a good impression across cultures. This was also demonstrated by the fact that the Explicit group was the only group to ask significantly more questions from the Pre-test to Post-test 1 and the Pre-test to Post-test 2. Still, this seems to be an area that needs to be constantly reinforced, as suggested by the decline in number of questions from Post-test 1 to Post-test 2. Another finding that was not surprising was the fact that the Implicit group's (Group B) WTC scores significantly increase

over time. The researcher believes that increased confidence (and decreased language anxiety) exhibited by the Implicit group in this study is mainly due to their extensive exposure to NESs. At the study's onset, most students admitted to being especially nervous because they did not have much, if any, experience communicating with a NES in a social context. The more experience they gained doing so, the less scared they seemed to feel about it.

Concerning how much participants in each group spoke, the results were largely inconclusive. Although members of the Explicit group produced significantly more words directly after treatment, their word output in Post-test 2 returned to near its Pre-test level. Further, the Implicit and Control groups also showed a similar increase in word output from the Pre-test to Post-test 1 and then a similar decrease from the Pre-test to Post-test 2 (see Figure 3). Thus, the increase may have had more to do with the JEFs, collectively, feeling more confident (and less anxious) in speaking in Post-test 1 because they knew what to expect from their experience in the Pre-test. Moreover, the fact that the standard deviations are quite high across groups in this category suggests considerable variability within the performances of the JEFs.

Conversational Repair Strategies

Although the success of explicit instruction on the aspects of backchannel production that are thought to be spontaneous and automatic may be somewhat surprising, it was rather predictable that the Explicit group was the only group in this study to improve in this area. In a similar study in which Takahashi (2001) investigated four input enhancement conditions for Japanese learners acquiring biclausal request forms in English, the students receiving explicit instruction were much better able to produce the set expressions (i.e., formulaic chunks of language) needed to continue the conversation and/or avoid communication breakdown. As shown in Figures 12 and 13 above, the Explicit group (Group A) clearly outpaced the other two groups (Groups B and Z) from Pre-test to Post-test 2 by demonstrating a significant increase in the amount of lengthy conversational repairs and a significant decrease in unconventional backchannels at situations of non-understanding employed. This can be explained by the fact that members of the Explicit group were taught specific linguistic expressions to use to manage a wide array of conversational situations (i.e., such as when they do not understand) and given practice opportunities and feedback to help them hone their skills. In comparison, the other groups (Groups B and Z) did not use a wide range of conversational repair strategies to manage conversational situations simply because they may have not even been exposed to such useful phrases and did not receive any

training in how and when to use them.

Intercultural Communicative Competence (ICC)

The data analysis of this category produced some clear and interesting findings (see Figures 10 and 11). First and foremost, this was another category in which the Explicit group (group A) outpaced the other two groups by a wide margin. This was demonstrated by the Explicit group's statistically significant improvement on 14 of 17 items on Hecht's (1978) modified questionnaire from the Pre-test to Post-test 1 (6 of which were significant at the .01 level), and 13 of 17 items from the Pre-test to Post-test 2 (4 of which were significant at the .01 level). It was somewhat predictable that the JEFLL group that came the closest to adopting NES norms also received the highest ratings in the NESs' conversational satisfaction questionnaires. This finding supports the hypothesis that backchannel conventions that are similar across cultures lead to positive perceptions and higher degrees of conversational satisfaction, whereas backchannel practices which differ run a much greater risk of being perceived negatively and/or resulting in miscommunication across cultures (Cutrone, 2005; White, 1989).

Interestingly, the Implicit group (Group B) and the Control group (Group Z) also registered several significant improvements in this category. Specifically, the Implicit group showed a statistically significant improvement on 9 of 17 items from both the Pre-test to Post-test 1 and the Pre-test to Post-test 2 (in both cases, 4 items were significant at the .01 level). The Control group (Group Z) showed a statistically significant improvement on 7 of 17 items from both the Pre-test to Post-test 1 and the Pre-test to Post-test 2 (only 1 item from the Pre-test to Post-test 1 was significant at the .01 level). The findings that demonstrated some level of improvement (i.e., increased ratings) for all three groups in this area were somewhat unexpected; however, these data may be linked to the NES observers (i.e., the raters) picking up on the JEFLLs' increased confidence. As mentioned above, the JEFLLs were especially nervous at the study's onset because they had no idea what to expect (and many had not even spoken socially with a NES before). However, as the study progressed, the JEFLLs reported higher degrees of confidence and less anxiety across the board because they were more familiar and comfortable with the study's process and their role in it.

Relating the Findings to SLA Theory

In relation to Schmidt's (1993) noticing hypothesis and the more general question of how input

becomes intake in the process of SLA, this is a debate that the writer does not expect to be resolved any time soon. The finding that explicit treatment (i.e., drawing attention to non-salient features of language such as listener responses) facilitates learning would seem to offer encouragement for the noticing hypothesis; however, it falls short of providing conclusive evidence supporting it. There are several issues involving the testability of the noticing hypothesis, which include an inconsistent and vague interpretation of what it means to notice in the research literature, the failure to recognize noticing as an internal process that cannot be observed directly, and the fact that noticing requires a high degree of inference from observation of behavior (Cross, 2002; Truscott, 1998). Schmidt (1990, 1993) equates noticing with attention plus awareness, which is operationalized as a cognitive process that takes place both during and directly after exposure to the input that is available for self-report. Therefore, as mentioned above, the fact that the Explicit group (Group A) received explicit input enhancement (i.e., increased opportunities to notice), subsequently demonstrated a much greater improvement than the other groups (Group B and Z) in terms of actual conversational performance (and L2 backchannel behaviour) over time, and then were able to report on specific areas of improvement by referring to what they had learned in class (i.e., what they had noticed) would seem to offer support for, at the very least, a mild version of the noticing hypothesis.

Additionally, the findings of this study would also seem to support Ellis' (2006a) belief that explicit input enhancement is particularly useful in dealing with acquisition issues related to lack of salience and L1-influenced blocking. Regarding the lack of salience issue, non-salient features of language such as backchannels seem to require intentionally focused attention to facilitate effective L2 learning. Concerning the L1-influenced blocking problem, the findings of this study suggest that consciously channelling attention to backchannel behavior may have helped change the cues that learners focused upon in their language process, which ultimately changed what their implicit learning systems were able to take in. As the Control group's lack of progress demonstrated throughout this study, it is likely that without such a change in the attentional focus of cues, learners would continue to demonstrate a great deal of L1 negative pragmatic transfer where backchannels are concerned, perhaps only gradually showing glimpses of improvement after substantial L2 experience. Thus, as MacWhinney (2001) asserts, form-focused-instruction (FFI) can be seen as a way to speed up the learning process.

As reported above, members of the Explicit group clearly showed the most improvement in

their backchannel behavior throughout the course of this study. However, it is also worthy to note that members of the Implicit group also displayed considerable improvements in several key areas of backchannel behavior. Schmidt's (1983) case study providing evidence of the implicit learning of L2 backchannel behavior in an immersion setting, and the opinions given by many of the JEFs, would seem to further support a mixed approach. Accordingly, a sensible way forward may be a combination of these methods in which implicit methods are used to supplement an explicit approach. This would seem to be consistent with Nick Ellis' (2008) explanation of how Schmidt's (1993) noticing hypothesis works:

once a stimulus representation is firmly in existence, that stimulation need never be noticed again; yet as long as it is attended for use in the processing of future input for meaning, its strength will be incremented and its associations will be tallied and implicitly catalogued. (p. 105)

In this description, Ellis also helps us understand how explicit knowledge is converted into implicit knowledge. Hence, as it relates to the findings of this study, it appears that explicit input enhancement (i.e., what draws learners attention) may be particularly helpful in facilitating the learning in the early stages of pragmatic development; however, sustained and long-term development in this area will likely depend on reinforcement opportunities and how much the learner is able to use the newly learned skill in authentic settings.

Conclusion

In summarizing the findings of this study, RQs 1 and 2 are addressed in succession below. First, concerning the effects of instruction on listening behavior, (RQ 1) the findings of this longitudinal study demonstrate that instructional treatment clearly had a positive effect on the listening behavior of both experimental groups, and (RQ 2) the group that received explicit treatment, by and large, outpaced the group that received implicit treatment. Although benefits of FFI in the teaching of L2 backchannel behavior have been shown, it is difficult to reach any firm conclusions regarding Schmidt's (1993) noticing hypothesis due to the issues mentioned above. Noticing appears to be helpful in speeding up the process of learning L2 backchannels, but it remains unclear as to whether it is actually necessary or not. While this distinction may be particularly important to theorists, it is far less relevant to classroom practitioners whose main

goal is to help students learn the target language in the most efficient way possible. Accordingly, the researcher adopts the position taken by Swan (2005) as follows:

The role of instruction in a typical language classroom is not, surely, to attempt the impossible task of replicating the conditions of natural acquisition, but to compensate for their absence. (p. 393)

In the broader context of language pedagogy, the main value of the present study would be in the general finding that pedagogical interventions, and most notably explicit input enhancement, did indeed appear to have a positive effect on L2 backchannel behavior. This not only provides support for the incorporation of backchannel behavior into the language classroom and cross-cultural communication training in the JEFLL context, but also provides ELT professionals with some pedagogical suggestions moving forward. The researcher hopes this study will serve as a platform for future investigation and diagnosis into this somewhat neglected aspect of pragmatic competence.

References

- Alun, R. E. (2008). Teaching the speaking skill to Japanese students part 1: Construct & practice. *The Journal of Kanda University of International Studies*, 20, 1-26.
- Anderson, J. R. & Lebriere, C. (1998). *The atomic components of thought*. Mahwah, NJ: Erlbaum.
- Blanche, P. (1987). The case for a pedagogy of pragmatics in foreign or second language teaching. *RELC Journal*, 18, 46-71.
- Blundell, J., Higgens, J. & Middlemiss, N. (1982). *Function in English*. Oxford: Oxford University Press.
- Boxer, D. (1993). Complaints as positive strategies: What the learner needs to know. *TESOL Quarterly*, 27, 277-299.
- Brozyna, B. (2007). The teachability of interactional discourse signals in the classroom environment. Unpublished Master's thesis, University of Reading, Reading.
- Byram, M. (1997). *Teaching and assessing intercultural communicative competence*. Clevedon: Multilingual Matters.
- Capper, S. (2000). Nonverbal communication and the second language learner: Some pedagogical considerations. *The Language Teacher*, 24(5), 19-23.
- Clancy, P. M., Thompson, S. A., Suzuki, R. & Tao, H. (1996). The conversational use of reactive tokens in English, Japanese, and Mandarin. *Journal of Pragmatics*, 26, 355-387.
- Clegg, F. (1982). *Simple Statistics: A Course Book for the Social Sciences*. Cambridge: Cambridge University Press.
- Crawford, W. (2003). Back-channel transfer in cross-cultural second language conversation: Do native Japanese transfer more than native English speakers? *Journal of Kitakyushu National College of Technology*, 36, 155-163.
- Cutrone, P. (2005). A case study examining backchannels in conversations between Japanese-British dyads. *Multilingua - Journal of Cross-Cultural and Interlanguage Communication*, 24, 237-274.

- Cutrone, P. (2010). The backchannel norms of native English speakers: A target for Japanese L2 English learners. *University of Reading Language Studies Working Papers*, 2, 28-37.
- Cutrone, P. (2011). *Towards the learning of listener responses in the Japanese EFL context*. Unpublished doctoral dissertation, University of Reading, Reading, UK.
- Cross, J. (2002). 'Noticing' in SLA: Is it a valid concept? *TESL-EJ*, 6(3). Retrieved October 25, 2010 from <http://www.cc.kyoto-su.ac.jp/information/tesl-ej23/a2.html>
- Dörnyei, Z. &Thurrell, S. (1992). *Conversation and dialogues in action*. Hemel Hempstead: Prentice Hall.
- Duncan, S. (1974). On the structure of speaker-auditor interaction during speaking turns. *Language in Society*, 2, 161-180.
- Duncan, S. & Fiske, D. (1977). *Face to Face Interaction: Research, Methods, and Theory*. Hillsdale: Erlbaum.
- Ellis, N. C. (2005). At the interface: Dynamic interactions of explicit and implicit language knowledge. *Studies in Second Language Acquisition*, 27, 305-352.
- Ellis, N. C. (2006a). Language acquisition as rational contingency learning. *Applied Linguistics*, 27, 1-24.
- Ellis, N. C. (2006b). Selective attention and transfer phenomena in L2 acquisition: Contingency, cue competition, salience, interference, overshadowing, blocking, and perceptual learning. *Applied Linguistics*, 27, 164-194.
- Ellis, N. C. (2008). Usage-based and form-focused SLA: The implicit and explicit learning of constructions. In A. Tyler, Y. Kim, &M. Takada (Eds.), *Language in the context of use: Discourse and cognitive approaches to language* (pp. 93-120). Berlin: Mouton de Gruyter.
- Ellis, R. (2008). *The study of second language acquisition (2nd edition)*. Oxford: Oxford University Press.
- Fujimoto, D. T. (2007). Listener responses in interaction: A case for abandoning the term, backchannel. *Journal of Osaka Jogakuin College*, 37, 35-54.

- Goldschneider, J. M. & DeKeyser, R. M. (2001). Explaining the “natural order of L2 morpheme acquisition” in English: A meta-analysis of multiple determinants. *Language Learning*, 51, 1-50.
- Hecht, M. (1978). The conceptualization and measurement of interpersonal communication satisfaction. *Human Communication Research*, 4, 253-264.
- Hess, L. J. & Johnston, J. R. (1988). Acquisition of back channel listener responses to adequate messages. *Discourse Processes*, 11, 319-335.
- Jefferson, G. (2002). Is “no” an acknowledgment token? Comparing American and British uses of (+)/(-) tokens. *Journal of Pragmatics*, 34, 1345-1383.
- Jenkins, J. (2003). *World Englishes*. London: Routledge.
- Krashen, S. (1982). *Principles and practice in second language acquisition*. Oxford: Pergamon.
- Lebra, T. (1976). *Japanese Patterns of Behavior*. Honolulu: The University of Hawaii Press.
- LoCastro, V. (1987). Aizuchi: A Japanese conversational routine. In L. Smith (Ed.), *Discourse across cultures: Strategies in world Englishes* (pp. 101-113). London: Prentice Hall.
- Markel, N. (1975). Coverbal behavior associated with conversation turns. In A. Kendon, R. Harris & M. R. Key (Eds.), *Organization of Behavior in Face-to-Face Interaction* (pp. 189-197). The Hague: Mouton.
- Maynard, S. K. (1986). On back-channel behavior in Japanese and English casual conversation. *Linguistics*, 24(6), 1079-1108.
- Maynard, S. K. (1987). Interactional functions of a nonverbal sign: Head movement in Japanese dyadic conversations. *Journal of Pragmatics*, 11, 589-606.
- Maynard, S. K. (1990). Conversation management in contrast: Listener responses in Japanese and American English. *Journal of Pragmatics*, 14, 397-412.
- Maynard, S. K. (1997). Analyzing interactional management in native/non-native English conversation: A case of listener response. *IRAL*, 35, 37-60.
- McCarthy, M. (2002). Good listenership made plain: British and American non-minimal response

- tokens in everyday conversation. In R. Reppen, S. M. Fitzmaurice & D. Biber (Eds.), *Using corpora to explore linguistic variation* (pp. 49-72). Philadelphia: John Benjamins.
- MacWhinney, B. (2001). The competition model: The input, the context, and the brain. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 69-90). New York: Cambridge University Press.
- McCarthy, M. (2002). Good listenership made plain: British and American non-minimal response tokens in everyday conversation. In R. Reppen, S. M. Fitzmaurice & D. Biber (Eds.), *Using corpora to explore linguistic variation* (pp. 49-72). Philadelphia: John Benjamins.
- McCarthy, M. (2003). Talking back: "Small" interactional response tokens in everyday conversation. *Research on Language and Social Interaction*, 36, 33-63.
- McCroskey, J. C. (1992). Reliability and validity of willingness to communicate scale. *Communication Quarterly*, 40, 16-25.
- Mizutani, N. (1982). The listener's responses in Japanese conversation. *Sociolinguistics Newsletter*, 13, 33-38.
- Nikolova, D. (2008). English teaching in elementary schools in Japan: A review of a current government survey. *Asian EFL Journal*, 10. Retrieved November 3, 2011 from http://www.asian-efl-journal.com/March_08_dn.php
- O'Keeffe, A. & Adolphs, S. (2008). Response tokens in British and Irish discourse: Corpus, context and variational pragmatics. In A. Barron & K. Schneider (Eds.), *Variational pragmatics* (pp. 69-98). Amsterdam: John Benjamins.
- O'Keeffe, A. Clancy, B. & Adolphs, S. (2011). *Introducing Pragmatics in Use*. Routledge: London.
- O'Keeffe, A., McCarthy, M. & Carter, R. (2007). *From Corpus to Classroom: Language Use and Language Teaching*. Cambridge: Cambridge University Press.
- Oreström, B. (1983). *Turn-Taking in English Conversation*. Lund: Lund University Press.
- Oshio, A., Abe, S. & Cutrone, P. (2012). Development, reliability, and validity of the Japanese version of ten item personality inventory (TIPI-J). *Japanese Journal of Personality*

Psychology, 21, 40-52.

- Rose, K. R. (2005). On the effects of instruction in second language pragmatics. *System*, 33, 385-399.
- Sacks, H. E., Schegloff, E. & Jefferson, G. (1974). A simple systematics in the organization of turn-taking for conversation. *Language*, 50, 696-735.
- Sardegna, V. G. & Molle, D. (2010). Videoconferencing with strangers: Teaching Japanese EFL students verbal backchannel signals and reactive expressions. *Intercultural Pragmatics*, 7(2), 279-310.
- Sato, Y. (2008). Oral communication problems and strategies of Japanese university EFL learners. Unpublished PhD Thesis, University of Reading, Reading, UK.
- Schmidt, R. (1983). Interaction, acculturation, and the acquisition of communicative competence: A case study of an adult. In N. Wolfson & E. Judd (Eds.), *Sociolinguistics and language acquisition* (pp. 137-174). Rowley: Newbury House.
- Schmidt, R. (1990). The role of consciousness in second language learning. *Applied Linguistics*, 11, 129-158.
- Schmidt, R. (1993). Consciousness, learning and interlanguage pragmatics. In G. Kasper & S. Blum-Kulka (Eds.), *Interlanguage pragmatics* (pp. 21-42). Oxford: Oxford University Press.
- Schmidt, R. & Frota, S. (1986). Developing basic conversational ability in a second language: A case study of an adult learner of Portuguese. In R. R. Day (Ed.), *Talking to learn: Conversation in second language acquisition* (pp. 237-326). Rowley, MA: Newbury House.
- Seidlhofer, B. (2004). Research perspectives on teaching English as lingua franca. *Annual Review of Applied Linguistics*, 24(2), 209-239.
- Spitzberg, B. H. (2000). A model of intercultural communication competence. In L. Samovar & R. Porter (Eds.), *Intercultural communication: A reader, 2nd edition* (pp. 7-24). Belmont, CA: Wadsworth Publishing.

- Stubbe, M. (1998). Are you listening? Cultural influences on the use of supportive verbal feedback in conversation. *Journal of Pragmatics*, 29(3), 257-289.
- Suzuki, A. (2010). Introducing diversity into ELT: Student teacher responses. *ELT Journal*, 64(3), 1-9.
- Swan, M. (2005). Legislation by hypothesis: The case of task-based instruction. *Applied Linguistics*, 26(3), 376-401.
- Takahashi, S. (2001). The role of input enhancement in developing pragmatic competence. In K. R. Rose & G. Kasper (Eds.), *Pragmatics in language teaching* (pp. 171-199). Cambridge: Cambridge University Press.
- Takimoto, M. (2009). The effects of input-based tasks on the development of learners' pragmatic proficiency. *Applied Linguistics*, 30, 1-25.
- The Japan Times (1983). Japanese mannerism is a key point in IBM case. *The Japan Times*, 23 January, p. 2.
- Thonus, T. (2007). Listener responses as a pragmatic source for learners of English. *CATESOL Journal*, 19, 132-145.
- Truscott, J. (1998). Noticing in second language: A critical review. *Second Language Research*, 14(2), 103-135.
- Ward, N., Escalante, R., Yaffa, A. & Solorio, T. (2007). Learn to show you're listening. *Computer Assisted Language Learning*. Retrieved September 8, 2010 from <http://www.cs.utep.edu/nigel/papers/call.pdf>
- Ward, N. & Tsukahara, W. (2000). Prosodic features which cue back-channel responses in English and Japanese. *Journal of Pragmatics*, 32(8), 1177-1207.
- White, S. (1989). Backchannels across cultures: A study of Americans and Japanese. *Language in Society*, 18, 59-76.
- Widdowson, H. G. (1997). EIL, ESL, EFL: Global issues and local interests. *World Englishes*, 16, 135-146.

Yngve, V. (1970). On getting a word in edgewise. *Chicago Linguistic Society*, 6, 567-578.

Appendices

Appendix A: An Example of a Transcribed Conversation Used in this Study

Transcription Conventions

- Listener responses are shown in italics below the primary speaker's talk at the point they occurred in the talk.
- To protect the identity of the participants, pseudonyms are used in the speaker labels on the left side of each transcribed line.
- To not confuse readers with the colons that are used for a different purpose described below, the speaker labels will be followed by a semi colon.
- To further preserve anonymity, pseudographs (i.e., notations in parentheses) will be used in instances where participants' private information such as name, address and/or telephone number has been uttered in the conversation.
- Numbers in parentheses indicate elapsed time in hundredths of seconds of pauses occurring in the conversations. Parentheses with a dot (.) indicate a micropause and/or hesitation under .5 seconds. Pauses are timed using transcription software in this study (Praat Version 5.0.18).
- The equal sign “=” indicates latching - i.e., no interval between the end of a prior piece of talk and the start of a next piece of talk.
- The beginnings of simultaneous speech utterances are marked by placing an opening square bracket at each of the points of overlap, and placing the overlapping talk directly beneath the talk it overlaps.
- Closing square brackets indicate the point at which two simultaneous utterances end.

Metatranscription is shown as follows:

- Empty parentheses () indicate part of the transcription that is unintelligible.
- Words between parentheses indicate the transcribers' conjecture at the words or utterances in the conversation that they are not completely certain of.
- Words between double parentheses may indicate comments and/or features of the audio materials other than actual verbalization.
- L stands for laughter.
- Other than apostrophes, which are used to show contraction between words, punctuation symbols in these transcriptions are not used as regular English punctuation markers indicating grammatical category. While other, non-regular, grammatical functions are shown by symbols such as slashes and double slashes, other punctuation symbols such as question marks and colons are used to indicate prosodic features in these transcriptions.

Nonverbal behavior is shown by the symbols indicated below.

- h stands for audible breathing. ^ stands for vertical head movement (head nod). > stands for horizontal head movement (head shake). S stands for smile. ” indicates that eyebrows are raised. G indicates body or hand gestures.
- In cases where nonverbal behavior occurs concurrently with speech, symbols are placed directly above the speech with which it co-occurs (instances where two types of nonverbal behavior occur simultaneously are shown by underlining them both). Nonverbal behavior that is continuous and occurs for a period longer than 2 seconds will be noted by signaling the beginning and the end of the behavior in parentheses where it occurs in the conversation. (N.B. The parentheses containing the

symbols below are solely used for separation purposes to make them easily identifiable in the specific examples below. Parentheses will not be used in this manner in the transcriptions as they have other specific functions, which have been outlined above.)

- A slash (/) marks the grammatical completion point of an internal clausal boundary (i.e., a clause which is continuative).
- Two slashes side by side (//) mark the grammatical completion point of a final clause boundary (i.e., a clause which is terminative). N.B. A final clause boundary is one that makes complete sense (i.e., fully meaningful) and could end the utterance there. In contrast, an internal clause is one in which the meaning is not complete, and there is a requirement for the utterance to go on in order for the meaning to be complete.
- A question mark (?) at the end of a word and/or utterance indicates a clear rising vocal pitch or intonation (i.e., one that is clearly heard, and is shown to rise by at least 600 Hz using Praat software).
- An inverted question mark (¿) at the end of a word and/or utterance indicates a clear falling pitch or intonation (i.e., one that is clearly heard and is shown to fall by at least 600 Hz using Praat software).
- A colon (:) as in the word “ye:s” indicates the stretching of the sound it follows (i.e., only marked in cases where the stretching was extended greater than .5 seconds).
- A hyphen at the end of an uncompleted word indicates the disfluency of a truncated word. For instance, if the word “word” were truncated, it may be transcribed as “wor-”.
- A part of a word and/or phrase containing CAPITAL letters indicates that it has been said with increased volume and/or more emphatically than the rest of the phrase (i.e., only marked when the highest point of the stressed part of speech is 10 decibels greater than the lowest part of the surrounding parts of speech).
- The underscore sign (_) indicates that the talk it precedes is low in volume.
- (~) indicates that the talk which follows is consistent with the person’s regular voice and tone. This symbol is used after low volume talk to indicate the point at which the volume rises back to normal. When a pause occurs after the low volume talk and the talk that follows returns to normal, this symbol is not shown.

Sample Transcription

1. Haruna; (1.89) yeah i have tried// (1.27) i had (.70) (^G)// (.) and (.) exam// (.) ^G (1.87)

[^]
uhuh

2. Haruna; maybe (1.36) (it's ok)// ^{^ ^} Lh [(^{^ ^ ^} Lh ma)]ybe (.85)

^{^ L ^ ^ ^}
[(*yeah maybe it's ok*)]

3. Andrea; [>] ((G begins)) h (urrr) what [exa]ms do you have// next week// do you have just one// (.65) ((G ends))

[()]

4. Haruna; no: (.) i have (1.67) four_i// (1.10) ^G ^{^ S}

really? (.68) eee =

5. Haruna; ^G = and two reports// (.62) (.) yeah_i (.77) (.) next (.59)

^G
next week? oh wow

6. Haruna; wednesday and (1.10) tuesday () thursday, (.)

7. Andrea; ((G begins)) ahh (isn't it)// it it's this coming week// (.) [(your)] exam week, (.99)

[^]
[(*yeah*)]

8. Andrea; ah at chodai (.) last week, = = was our exam week// =

^G
= *last week* = = (*eeeya*) (.62) *uu* =

9. Andrea; so everyone is finished// (.) so last night, we had a party// (.)

10. Andrea; like a (.) like YAAY [(^S Lh) (((G ends)) ^ ^ ^ finished)]party _ Lh (.)
 [(^{^ ^ ^} Lh) (^S finished)]

11. Andrea; [(so yeah at an izakaya) (.) so =
 [(h _ Lh)] = (wow) (.)

12. Haruna; [^G where] = did you go// (.)
 [(it was] fun) =

13. Andrea; umm (.) there's an izakaya/ i:n sumiyoshi// (.) that's called (1.38)

14. Andrea; _ ahhi forget. the name// (.71) ok (.) if there's (1.22) uu:mm (.65)
 ((G begins))

15. Andrea; if this is chitosepia/ = (.) and here's like the main road// (.) [(the)] big road, (.99)
 = yes [(^{^ ^} yeah)]

16. Andrea; and here is an the arcade// entrance to the arcade, (.) _ () (arcade) here, (.)
 ^

17. Andrea; and there's like a little (.) aah there's a bus stop// (.) and a little park// (.91)
 (yeah) =

18. Andrea; so go down this road// past chitose-, chitosepia is like here// (.) (.)
 oh ok

19. Andrea; and there's a road. here// = = go down this like the main the big chitosepia,
 = _ (mm) =

20. Andrea; i think it's (.) showamachi// (.) (.) ummm () yeah yeahyeah

yes

21. Andrea; go down that road// (.) and it's it's on the (1.07) it's on the lulu left side// (.50)

>> (G ends)

22. Andrea; i don't remember. the name// (.81) G = kind of big though, (.) (.)

"

_ (mm) = _ (oh)

23. Andrea; _ mm ~ kind of nice, = = cuz we had (1.10) sixteen people// (.)

= Lh =

24. Andrea; [so] it was very big// _ mmm (1.18) izakaya (.91)

"

[(oh)]

25. Haruna; ((G begins)) (there is a) (.69) (all) (.76) _ (kenkyusseï// no) (.) ((G ends))

26. Andrea; oh [ken]kyusseï, = = mm?mm¿mm? (.) umm maybe like (.)

[(oh)] = _ kenkyusseï =

27. Andrea; about half (.56) japanese students and half (aaa) foreign students, (.) (.)

_ (ooh)

28. Andrea; but i'm the only _ kenkyusseï// () (.56) other students,

29. Andrea; are like regular ryugakkusseï// (.65) _ i think (1.33) or just (.) japanese _ (college students)//

^^

((G begins)) S

30. Andrea; (2.19) so they all had to take the test// (.79) [() (so they were all like)]

^

[(woooLh)]

((G ends)) >>> S

31. Andrea; oooooo we're so tired// oo yay = [(we're finished)]// (.)

^ ^ ^

= ooo[oooooooooooo]

((G begins))

S ((G ends))

32. Andrea; but i said like (oh ok fine)// = = i didn't do anything// = [so]

= ok Lh =

= Lh [Lh] (.) h =

((G begins))

33. Andrea; () [(not bad)] (.) like a bonenkai, but not a bounen [(kai, like)]

[()] _ mmm

[(bounen)]

L

((G ends))

^ ^ ^

34. Andrea; a (.) (.) [(like a bon shu there you go)]// (.) so (.) (.)

[(bou)]

[()]

Lh

^ ^
] _ (h)

_ (h)

^

^

S

35. Andrea; yeah (.) _ (it was fun)// (.99) so (1.02)

^

^

^ ^

36. Andrea; did you do anything. fun// yesterday// (.54) = uu uh (1.44)

_ (uu) yesterday =

37. Haruna; yesterday (2.40) i (.) i did a part-time job// (.)

38. Andrea; oh really what's your (erbaito)// (.)

^

39. Haruna; yeah (.52) uuu (.) = (.) it's

= where (.62) _ ()

Appendix B: Willingness to Communicate Scale

DIRECTIONS: Below are twenty situations in which a person might choose to communicate or not to communicate in English. Presume that the person in each situation does not speak Japanese but can speak English. Also, presume you have *completely free choice*. Indicate the percentage of times you would choose to *communicate* in each type of situation. Indicate in the space at the left what percent of the time you would choose to communicate.

0 = never, 100 = always

- _____ 1. *Talk with a service station attendant.
 _____ 2. *Talk with a physician.
 _____ 3. Present a talk to a group of strangers.
 _____ 4. Talk with an acquaintance while standing in line.
 _____ 5. *Talk with a salesperson in a store.
 _____ 6. Talk in a large meeting of friends.
 _____ 7. *Talk with a police officer.
 _____ 8. Talk in a small group of strangers.
 _____ 9. Talk with a friend while standing in line.
 _____ 10. *Talk with a waiter/waitress in a restaurant.
 _____ 11. Talk in a large meeting of acquaintances.
 _____ 12. Talk with a stranger while standing in line.
 _____ 13. *Talk with a secretary.
 _____ 14. Present a talk to a group of friends.
 _____ 15. Talk in a small group of acquaintances.
 _____ 16. *Talk with a garbage collector.
 _____ 17. Talk in a large meeting of strangers.
 _____ 18. *Talk with a spouse (or girl/boyfriend).
 _____ 19. Talk in a small group of friends.
 _____ 20. Present a talk to a group of acquaintances.

N.B. JEFs were provided with Japanese explanations. Further, the asterisk (*) marking the filler items above, as well as the scoring table below, were not included on the questionnaires the JEFs completed.

SCORING: The WTC permits computation of one total score and seven sub-scores. The sub-scores relate to willingness to communicate in each of four common communication contexts and with three types of audiences. To compute your scores, merely add your scores for each item and divide by the number indicated below.

| <i>Subscore Desired</i> | <i>Scoring Formula</i> |
|-----------------------------|---|
| Group discussion | Add scores for items 8, 15, and 19; then divide by 3. |
| Meetings | Add scores for items 6, 11, and 17; then divide by 3. |
| Interpersonal conversations | Add scores for items 4, 9, and 12; then divide by 3. |
| Public speaking | Add scores for items 3, 14, and 20; then divide by 3. |
| Stranger | Add scores for items 3, 8, 12, and 17; then divide by 4. |
| Acquaintance | Add scores for items 4, 11, 15, and 20; then divide by 4. |
| Friend | Add scores for items 6, 9, 14, and 19; then divide by 4. |

To compute the total WTC scores, add the sub-scores for stranger, acquaintance, and friend. Then divide by 3.

Appendix C: Ten Item Personality Inventory

Name (名前) : _____ Date (記入日) : _____

Following the scale below, please write a number next to each statement below to indicate the degree to which you agree or disagree with that statement. (下の枠内の1から7までのスケールに従って、1から12までの問いに対して、最も自分に当てはまる度数の数字を、問いの数字横の空欄にそれぞれ入れてください。)

| | | | | | | |
|--------------------------------|---------------------------------------|--------------------------------|---|----------------------------|--------------------------------|----------------------------|
| Disagree Strongly (全く違うと思う) | Disagree moderately (あまりそうだとは思わない) | Disagree a little (少し違うと思う) | Neither agree nor disagree (どちらでもない) | Agree a little (少しそう思う) | Agree moderately (まあまあそう思う) | Agree strongly (強くそう思う) |
| 1 | 2 | 3 | 4 | 5 | 6 | 7 |

I see myself as (私は自分のことを・・・) :

1. _____ Extraverted, enthusiastic. (外向的、社交的、熱心だと思う)
2. _____ Critical, quarrelsome. (批判的、口やかましいと思う)
3. _____ Dependable, self-disciplined. (頼りがいがある、自立していると思う)
4. _____ Anxious, easily upset. (感情が変化しやすい、すぐいらいらすと思う)
5. _____ Open to new experiences, complex. (新しい経験や物事に挑戦する事が好き)
6. _____ Reserved, quiet. (遠慮がち、おとなしいと思う)
7. _____ Sympathetic, warm. (思いやりがある、あたたかみがあると思う)
8. _____ Disorganized, careless. (注意ミスが多い、忘れ物が多いと思う)
9. _____ Calm, emotionally stable. (おだやか、感情が安定していると思う)
10. _____ Conventional, uncreative. (新しい物事に保守的、独創性がないと思う)

(N.B. The Japanese translation of the TIPI has since been modified and renamed the TIPI-J as presented in the recent works of Oshio, Abe& Cutrone , 2012.)

Appendix D: Inventory of Conversational Satisfaction

For NES Interlocutors

Your Name:

Partner's name:

Key: 1 = Yes

7 = No

Date:

Please score the sentences below based on how often you thought they generally occurred in the conversation. Based on the key shown above, circle the number that best corresponds to your opinion.

- 1. S/he let me know that I was communicating effectively.....1 2 3 4 5 6 7
- 2. The feelings that my partner expressed by means of listening feedback during the conversation seemed authentic (i.e., they conveyed what they were truly feeling and were not just agreeing and/or pretending to understand to keep the conversation going smoothly).....1 2 3 4 5 6 7
- 3. S/he showed me that s/he understood what I said.....1 2 3 4 5 6 7
- 4. S/he showed me that s/he listened attentively to what I said.....1 2 3 4 5 6 7
- 5. S/he expressed a lot of interest in what I had to say.....1 2 3 4 5 6 7
- 6. The conversation went smoothly.....1 2 3 4 5 6 7
- 7. S/he encouraged me to continue talking.....1 2 3 4 5 6 7
- 8. S/he seemed impatient.....1 2 3 4 5 6 7
- 9. S/he seemed cold and unfriendly.....1 2 3 4 5 6 7
- 10. S/he was polite.....1 2 3 4 5 6 7
- 11. S/he appeared warm and friendly.....1 2 3 4 5 6 7
- 12. S/he was impolite.....1 2 3 4 5 6 7
- 13. S/he appeared interested and concerned.....1 2 3 4 5 6 7
- 14. S/he interrupted me.....1 2 3 4 5 6 7
- 15. My conversation partner seemed to want to avoid speaking.....1 2 3 4 5 6 7

16. Please include any other comments and/or observations regarding the Japanese participant's behavior in the conversations. (Feel free to add any comments you have regarding Japanese people's listening behavior in general.)

For JEFLs

Inventory of conversational satisfaction (対話による満足度調査アンケート)

Key: 1 = Yes (最もそう思う) 7 = No (全くそう思わない) Name(名前):

Date (記入日):

1. S/he let me know that I was communicating effectively.....1 2 3 4 5 6 7
(私との会話を全体的に理解してくれているようだった)
2. I felt I was able to present myself fairly during the conversation.....1 2 3 4 5 6 7
(私は、自分の気持ちを会話の中できちんと言えたと思う)
3. S/he showed me that s/he understood what I said.....1 2 3 4 5 6 7
(相手は、私の話した内容を理解してくれていた)
4. S/he showed me that s/he listened attentively to what I said.....1 2 3 4 5 6 7
(相手は、私の話しに注意深く耳を傾けてくれていた)
5. S/he expressed a lot of interest in what I had to say.....1 2 3 4 5 6 7
(相手は、私の話す内容にとっても興味がある様子だった)
6. The conversation went smoothly.....1 2 3 4 5 6 7
(この対話はスムーズに進んだ)
7. S/he encouraged me to continue talking.....1 2 3 4 5 6 7
(相手は、私が話を続けやすいようにサポートしてくれた)
8. S/he seemed impatient.....1 2 3 4 5 6 7
(相手は、いらいらしている様子だった)
9. S/he seemed cold and unfriendly.....1 2 3 4 5 6 7
(相手は、冷たい反応で、不親切だった)
10. S/he was polite.....1 2 3 4 5 6 7
(相手は、丁寧な対応だった)
11. S/he appeared warm and friendly.....1 2 3 4 5 6 7
(相手は、あたたかく、親切な反応だった)
12. S/he was impolite.....1 2 3 4 5 6 7
(相手は、失礼な態度だった)
13. S/he appeared interested and concerned.....1 2 3 4 5 6 7
(相手は、私の話に興味と関心を示してくれた)
14. S/he interrupted me.....1 2 3 4 5 6 7
(相手は、私の話の邪魔や妨害、さえぎりをした)
15. My conversation partner seemed to want to avoid speaking.....1 2 3 4 5 6 7
(彼ら自身が絶対話さなくていいように、他の人に話し続けてほしいと思っているという印象を与えている)

16. Please include any other comments and/or observations regarding the participant's behavior in the conversations. (上記に無いコメントや意見、感想などがあれば書いてください)

For NES Observers

Date: _____ Name: _____ **Key:** 1 = Yes 7 = No

Please score the sentences below based on how often you thought they generally occurred in the conversation. Based on the key shown above, circle the number that best corresponds to your opinion.

- 1. The Japanese person let his/her partner know that the partner was communicating effectively.
.....1 2 3 4 5 6 7
- 2. The Japanese person showed his/her partner that they understood what their partner said.
.....1 2 3 4 5 6 7
- 3. The Japanese person showed that they were listening attentively to what their partner said.
.....1 2 3 4 5 6 7
- 4. The Japanese participant expressed a lot of interest in what their partner had to say.
.....1 2 3 4 5 6 7
- 5. The conversation went smoothly.....1 2 3 4 5 6 7
- 6. The Japanese encouraged his/her partner to continue talking.....1 2 3 4 5 6 7
- 7. The feelings that the Japanese person expressed by means of listening feedback during the conversation seemed *authentic* (i.e., they conveyed what they were truly feeling and not just agreeing and/or pretending to understand for the sake of harmony and/or to keep the conversation going smoothly).....1 2 3 4 5 6 7
- 8. The Japanese person seemed impatient.....1 2 3 4 5 6 7
- 9. The Japanese person seemed cold and unfriendly.....1 2 3 4 5 6 7
- 10. The Japanese person was polite.....1 2 3 4 5 6 7
- 11. The Japanese person appeared warm and friendly.....1 2 3 4 5 6 7
- 12. The Japanese person was impolite.....1 2 3 4 5 6 7
- 13. The Japanese person appeared interested and concerned.....1 2 3 4 5 6 7
- 14. The Japanese person interrupted their partner at times.....1 2 3 4 5 6 7
- 15. The Japanese person seemed to want to avoid speaking.....1 2 3 4 5 6 7
- 16. When the Japanese person did not understand something, they were able to clearly convey this to their conversational partner with their listening feedback.....1 2 3 4 5 6 7
- 17. The Japanese person's listening behavior seemed inadequate in some ways.....1 2 3 4 5 6 7
If you answered "yes" (i.e., 1, 2 or 3) to question 17, please explain how and/or why you think their listening behavior seemed inadequate.

18. Any other comments and/or observations regarding the Japanese participant's behavior in the conversation.

Appendix E: Data Record Sheet to Record Instances of Miscommunication

| Data Record Sheet | | | | | | | |
|--|--|--|--|--|--|--|--|
| 1. Wanting the speaker to continue = CONT 2. Indicating understanding of content = UND 3. Indicating agreement = AGR 4. Showing empathy and support to the speaker's evaluative statement = EAS 5. Showing a strong emotional response = SER 6. As a minor addition = MA 7. Asking for clarification = CLAR 8. Indicating non-understanding = NONU 9. Indicating disagreement or dissatisfaction = DOD 10. Other = OTH 11. Indicating boredom and/or disinterest in the conversation = BODI 12. Indicating impatience and a desire for the speaker to finish quickly = IMP 13. Giving the impression that their response was insincere = INSI 14. Giving the impression of not understanding but pretending to = PRET 15. Giving the impression that they want to keep the other person speaking strictly to avoid speaking themselves = AVSP (N.B. More than one function is possible. It also possible for the interviewee to be unsure, which will be demonstrated with a question mark.) | | | | | | | |
| Time | | | | | | | |
| JEFL BC Intentions | | | | | | | |
| NES BC Perceptions | | | | | | | |
| Analysis (notes) | | | | | | | |
| Time | | | | | | | |
| JEFL BI | | | | | | | |
| NES Perceptions | | | | | | | |
| Analysis (notes) | | | | | | | |
| Time | | | | | | | |
| JEFL BI | | | | | | | |
| NES Perceptions | | | | | | | |
| Analysis (notes) | | | | | | | |

Appendix F: Tables Demonstrating Characteristics of Three Groups

Characteristics Pertaining to Members of Group A

| Pseudonym | Gender | TOEIC Scores | L2 WTC (dyadic) | Personality (Extraversion) | Age | Year of Study |
|-----------|--------|--------------|-----------------|----------------------------|------|-----------------|
| Michiko | F | 700 | 65 | 6.5 | 20 | 3 rd |
| Hika | F | 650 | 60 | 6 | 20 | 3 rd |
| Aria | F | 625 | 44 | 2 | 19 | 2 nd |
| Haruna | F | 555 | 61 | 6 | 19 | 2 nd |
| Miya | F | 500 | 54 | 4 | 19 | 2 nd |
| Saya | F | 470 | 45 | 4.5 | 19 | 3 rd |
| Chieko | F | 445 | 34 | 3 | 20 | 2 nd |
| Ayuka | F | 360 | 42 | 5 | 19 | 2 nd |
| Takanori | M | 504 | 40 | 4 | 20 | 3 rd |
| Kazuya | M | 480 | 39 | 2.5 | 19 | 2 nd |
| AVG. | | 528.9 | 48.4 | 4.35 | 19.4 | |

Characteristics Pertaining to Members of Group B

| Pseudonym | Gender | TOEIC Scores | L2 WTC (dyadic) | Personality (Extraversion) | Age | Year of Study |
|-----------|--------|--------------|-----------------|----------------------------|------|-----------------|
| Rika | F | 685 | 50 | 5 | 20 | 2 nd |
| Shio | F | 630 | 49 | 3 | 20 | 3 rd |
| Mayumi | F | 585 | 35 | 2 | 20 | 3 rd |
| Meo | F | 550 | 52 | 2.5 | 20 | 3 rd |
| Madora | F | 500 | 60 | 4.5 | 19 | 2 nd |
| Sachi | F | 475 | 59 | 5 | 19 | 2 nd |
| Keiko | F | 460 | 39 | 4.5 | 19 | 2 nd |
| Mika | F | 390 | 51 | 3.5 | 19 | 2 nd |
| Taro | M | 525 | 58 | 5 | 19 | 2 nd |
| Kouki | M | 450 | 51 | 4 | 20 | 3 rd |
| AVG. | | 525 | 50.4 | 3.9 | 19.4 | |

Characteristics Pertaining to Members of Group Z

| Pseudonym | Gender | TOEIC Scores | L2 WTC (dyadic) | Personality (Extraversion) | Age | Year of Study |
|-----------|--------|--------------|-----------------|----------------------------|-----|-----------------|
| Yukari | F | 695 | 64 | 6 | 20 | 3 rd |
| Mikki | F | 645 | 63 | 4 | 20 | 3 rd |
| Akie | F | 595 | 60 | 5 | 20 | 2 nd |
| Yuki | F | 565 | 53 | 4.5 | 19 | 2 nd |
| Tomomi | F | 540 | 61 | 5.5 | 20 | 3 rd |
| Yoko | F | 516 | 48 | 5 | 19 | 2 nd |
| Misako | F | 485 | 54 | 5 | 19 | 2 nd |
| Yukiko | F | 445 | 41 | 6 | 19 | 2 nd |
| Akanori | M | 610 | 58 | 1.5 | 19 | 2 nd |

| | | | | | | |
|------|---|-------|------|------|------|-----------------|
| Hiro | M | 560 | 60 | 5 | 20 | 3 rd |
| AVG. | | 565.6 | 56.2 | 4.75 | 19.6 | |

English vocabulary instruction through storybook reading for Chinese EFL kindergarteners: Comparing rich, embedded, and incidental approaches

Susanna S. Yeung

Department of Psychological Studies, The Hong Kong Institution of Education, Hong Kong

Mei-lee Ng

Department of Early Childhood Education, The Hong Kong Institute of Education, Hong Kong

Ronnel B. King

Department of Curriculum and Instruction, The Hong Kong Institute of Education, Hong Kong

Bio data

Dr. Susanna S. YEUNG is currently Assistant Professor of Department of Psychological Studies at The Hong Kong Institute of Education (HKIEd). Before joining HKIEd, she has taught at The University of Hong Kong, The Hong Kong Polytechnic University and The Open University of Hong Kong. She received her B.SocSci, MPhil, and EdD degrees from the University of Hong Kong. Her research interests revolve around early English language and reading development and classroom intervention and learning outcomes. siusze@ied.edu.hk

Dr. Mei Lee NG is an Associate Head and Assistant Professor in the Early Childhood Education Department of the Hong Kong Institute of Education. She holds two MAs(in TESOL & in Child development with ECE) from the Institute of Education of the London University. Before she shifted her focus of interest in young children, she has been an English language teacher in secondary school for 12 years. She completed her PhD in the University of Hong Kong and her main research interests are in early second language learning and parental involvement in children's learning and development.

Dr. Ronnel B. KING received his PhD (major in educational psychology) at The University of Hong Kong in 2012, and was conferred the "Highly Commended PhD Award" by the Global

Address correspondence: 10 Lo Ping Road, Tai Po, New Territories, Hong Kong

SELF Research Network for his doctoral dissertation. He is currently working as an assistant professor in the Department of Curriculum and Instruction, The Hong Kong Institute of

Education (HKIED). Prior to joining HKIED, he was a research scientist at the National Institute of Education, Singapore. His research interests are on student motivation, well-being, and learning.

Abstract

Vocabulary is important for both reading and achievement outcomes. To date, there has been a lack of research examining the effects of vocabulary instruction within a storybook reading context for young Chinese children who learn English-as-a-foreign-language (EFL). The present study aims to evaluate the effects of rich and embedded approaches as compared with an incidental approach in enhancing knowledge of word meanings for Chinese EFL kindergarteners. The vocabulary instructions were done within a storybook reading context. Within-subject design was adopted in which participants underwent vocabulary instruction using rich, embedded (also known as extended), and incidental approaches. Instructional effects were measured by the increase in knowledge of word meaning of target words taught by each of the instructional approaches. Forty-three Hong Kong Chinese EFL children from three classrooms in a kindergarten participated in the study. Findings showed that rich instruction resulted in significant increases in receptive and expressive knowledge of word meanings. Both extended instruction and incidental instruction were not able to significantly enhance target word knowledge. Implications for evidence-based vocabulary instruction for Chinese EFL children are discussed.

Keywords: vocabulary instruction; Chinese EFL learners; storybook reading

Introduction

Sufficient vocabulary is important for both reading outcomes as well as academic achievement (August & Shanahan, 2006). Vocabulary learning is known to be a complex and lengthy process (Beck & McKeown, 2007). The size and quality of vocabulary for children who learn two languages simultaneously may suffer due to reduced exposure to oral input as compared to their monolingual counterparts (August, Carlo, Dressler, & Snow, 2005). An increasing number of studies have examined how to accelerate vocabulary learning among children who learn English-as-a-foreign-language (EFL; Slavin & Cheung, 2003). However, majority of vocabulary training studies have focused on vocabulary development of English native speakers with limited research on the development of EFL learners, especially for those who are from Asian educational contexts. This is alarming given that a growing population of EFL learners around the world, thus the need to have adequate knowledge of how to improve their vocabulary attainment. Moreover, extant research has focused on examining the improvements on comprehension of texts but not on learning the meanings of words (Marulis & Neuman, 2010). For young children, learning word meanings is important for text comprehension (Babayigit, 2015, Jeon & Yamashita, 2014). Currently, few studies have examined the effectiveness of vocabulary instruction among EFL children (August & Shanahan, 2006). The present research extends previous literature by increasing our understanding of the efficacy of theory-driven vocabulary instruction for EFL kindergarten children from Hong Kong, with the learning of word meanings as outcomes.

In Hong Kong, most children start to learn two languages (Chinese and English) in both oral and written forms around the age of 3. Ng & Rao (2013) reported the results of a recent survey showing that 100% of the 256 sampled kindergartens offered English teaching at the K2 and K3 levels (ages 4 and 5). English was taught as a separate subject in specific time slots during the week. The teaching of English relied much on the use of textbooks which generally focused more on print learning using whole word method (e.g., copying target words) than oral language skills. Formal exercises, such as copying words and sentences, were found to be more dominant than oral activities such as singing, reading stories or playing language games. Despite an early start, children are usually exposed to limited English oral language outside English lessons, and English is rarely used for daily communication. Consequently, Hong Kong children's English oral language proficiency is generally low. Given this unique educational context, the aim of the present research is to examine the effects of implicit and explicit vocabulary instructions on the

learning of word meaning among Chinese EFL kindergarteners. In this study, we examined to what extent different instructional approaches are able to enhance vocabulary knowledge of target words. Hence, the instructional approach is said to be effective if it enhances the knowledge of word meaning for the target words. To our knowledge, there has been no research on the effects of explicit/implicit instructional approaches on second language vocabulary learning in the Asian context.

Literature Review

Vocabulary Learning

Storybook reading is traditionally a tool for enhancing children's language skills. Storybook reading activities provide a meaningful context for young children to understand the meanings of vocabulary, actively engage in the learning process, and facilitate rich dialogues (Dickinson, Griffith, Golinkoff, & Hirsh-Pasek, 2012). A number of vocabulary teaching approaches within a storybook reading context have been shown to be effective, namely incidental instruction, embedded instruction, and rich instruction (also known as extended instruction; e.g, Elley, 1989; Beck & McKeown, 2007).

The major distinction among these instructional approaches is whether new words are explicitly taught or not. Embedded instruction and rich instruction adopt explicit teaching of vocabulary in which the meanings and definitions of new words are given before, during or after storybook reading. Follow-up activities are usually conducted to review these words. The two approaches vary in terms of the depth and breadth, with embedded instruction having greater breadth and rich instruction having more in-depth teaching of words. It means that within a certain amount of instructional time, embedded instruction is able to enhance the knowledge of larger number of words but students would have limited knowledge of each target word. In contrast, rich instruction is able to have a more complete knowledge of words (multiple meanings and word usage in different contexts; Coyne, McCoach, Loftus, Zipoli, & Kapp, 2009; Zipoli, Coyne, & McCoach, 2011).

On the other hand, the incidental approach, which is also known as the implicit or indirect method, does not involve deliberate teaching of the meanings of target words (Elley, 1989). Children acquire the word meanings through inferences from the story context. This distinction

(implicit vs. explicit or direct vs. indirect) is a key component identified in the field of vocabulary instruction (National Reading Panel, 2000).

There has been mixed findings regarding the effects of three types of instructions in word learning. The incidental approach is argued to be effective because it enhances the association of a word with its meaning by providing repeated encounters (Biemiller & Boote, 2006). The embedded and rich instructions follow a cognitive processing framework to engage children in active mental manipulation of new meanings. Thus, the active interactions with the new words allow the learner to have the ability to use the word in other contexts and form generalizations (Miller, 2003; Nagy & Scott, 2000). Recently, related research has involved younger students (e.g., McKeown & Beck, 2014). However, there is a lack of such investigations which directly compared instructions for learning vocabulary in a second language varied along the implicit/explicit dimension in the Asian educational settings. The following sections will present a more detailed review of the three instructional approaches and related empirical evidence.

Incidental instruction

Early research work in vocabulary teaching has shown that children can learn vocabulary through storybook reading (teacher read aloud) and discussing stories (e.g., Bus, van Ijzendoorn, & Pellegrini, 1995; Elley, 1989; Robbins & Ehri, 1994; Senechal & Cornell, 1993). There is also evidence that children are able to learn word meanings from viewing video narratives (Oetting, Rice, & Swank, 1995). These findings support that incidental exposure without intentional teaching of target words is beneficial to vocabulary development. The exposure to target words through repeated readings help students associate target words with its meaning (Biemiller & Boote, 2006).

However, a recent meta-analysis study on vocabulary intervention reported that incidental instruction was not as effective as the explicit approach in accelerating the learning of word meaning (Marulis & Neuman, 2010). Moreover, students with low initial vocabulary are shown to have difficulties in learning words through incidental exposure with storybook telling alone without extra supports such as pronouncing, defining and discussing target words after the storytelling (e.g., Beck & McKeown, 2007; Senechal, Thomas, & Monker, 1995). They might have difficulties inferring word meanings from stories because of their inadequate oral vocabulary and content knowledge (Stahl, 1991). EFL children generally have lower L2

vocabulary (Oller & Eilers, 2002). Therefore, in the current investigation, the incidental approach was regarded as a control condition, in comparison with explicit methods which have been shown to be more effective among English speaking populations.

Embedded instruction

Vocabulary learning can be facilitated by direct and explicit instruction of new words encountered in storybook reading (Beck, McKeown, & Kucan 2002). The embedded instructional approach provides children with simple explanations of target words or presents simple synonyms during storybook reading. Follow-up activities are designed to review the target words. Examples of activities include role playing or acting out the meanings, pointing to the picture in the book, or other activities that allow students to pronounce and remember the word meanings.

Some studies have reported evidence that providing target word explanations in the context of storybook reading is more effective compared to when no explicit explanations are provided (incidental exposure) among English native speaking young children (e.g., Penno, Wilkinson, & Moore., 2002; Justice, Meier, & Walpole, 2005). For instance, Biemiller and Boote (2006) demonstrated that story book reading followed by simple meaning explanations was effective in teaching word meanings for junior primary grade students and the gains were maintained over a period of 4 weeks after posttest.

However, some researchers have argued that embedded instruction may not be sufficient for students whose initial vocabulary is low or those who are at risk for disability (Robbins & Ehri, 1994). A more intensive teaching of target words is needed for students to accelerate their vocabulary development (Gersten, 1998; Gersen & Baker, 2000). The present study of vocabulary instruction with young EFL learners in Hong Kong compares the effects of instruction with varying intensity in enhancing knowledge of word meaning. An alternative approach to embedded instruction is rich or extended instruction, which exposes children to more intensive and extended teaching of word meanings (Beck et al., 2002).

Rich instruction

Rich instruction within the context of storybook reading proposed by Beck, Perfetti, and McKeown (1982) involves teaching vocabulary with elaborate explanations, using the target words in different contexts and playing meaningful word games to allow students to apply the

target words. This approach is characterized by providing both definitional and contextual definitions of the new words, giving frequent and wide-ranging opportunities for children to use the target words in different contexts and thus encouraging deep processing of meanings of the target words. In particular, a key distinction of this approach compared to embedded instruction within the storybook reading context is that the application of target words to contexts beyond the associated story is demonstrated among young first language learners in rich instruction but not in embedded instruction. However, a direct comparison between the two approaches (rich and embedded instruction) has not been conducted for EFL learners.

Beck and colleagues (Beck & McKeown, 2007) investigated the effects of rich instruction as compared to incidental approach among kindergarten and first grade students. They found that rich instruction had better effects on vocabulary acquisition than incidental learning among typical L1 readers. This method is also demonstrated as more effective in enhancing vocabulary knowledge than the incidental approach among L1 students who are at-risk for reading disability (Maynard, Pullen, & Coyne, 2010).

A major difference between the embedded and rich instruction is that the former emphasizes breadth and efficiency and the latter focuses more on depth and effectiveness. Given the fact that both rich and embedded instruction belong to the explicit spectrum, some recent studies have directly compared the effectiveness of these two methods which vary mainly in the extent to which new words are being reviewed and applied in other contexts. They found that rich instruction is more effective than embedded instruction in word learning and students' understandings of word meanings were maintained six weeks after the instruction (Coyne, McCoach, & Kapp, 2007). However, in another study, Coyne and colleagues (Coyne et al., 2009) reported that rich instruction was effective in teaching full and refined meaning of words while embedded instruction was more effective in enhancing partial knowledge of target words among kindergarten students as compared to the incidental approach. Specifically, rich instruction was more effective in enhancing performance in vocabulary measures requiring the production of definitions whereas embedded instruction resulted in higher number of word meanings acquired (partial knowledge). They argued that embedded instruction is time efficient while rich instruction results in more refined knowledge of target vocabulary.

Effective methods in vocabulary instruction for EFL children

Studies reviewed above involved native English speaking children. In the EFL learning context, studies have generally shown that vocabulary instructions that work for native English speaking children are also effective for EFL children (August & Shanahan, 2006; Silverman, 2007). Certain strategies associated with embedded and rich instruction such as illustrating and acting out were found to be especially important for EFL children because of their lower vocabulary knowledge (Silverman, 2007).

However, many of these studies involved children educated in English-speaking environments with different home languages. The educational context in Hong Kong is different in terms of language exposure and the use of L2 in daily communication. It is conceivable that Hong Kong EFL children may need extra support in vocabulary acquisition. To our knowledge, there has been no empirical study comparing the direct instructional approach to the incidental approach among Asian EFL children who are not immersed in an English speaking environment.

The Present Study

The purpose of the present study was to compare the effects of three pedagogical methods (incidental, embedded, and rich) on learning the meanings of new words among Chinese EFL kindergarten children. Using a within-subject design, we compared the improvements on word knowledge after undergoing rich, embedded, and incidental instructions. Given that the targets words taught were unfamiliar to the participants, we assumed that the gains in knowledge of word meaning were a result of the instruction provided to the participants. Specifically, our research question is: Are rich and embedded instructions more effective in enhancing word knowledge among Chinese EFL children in kindergarten classrooms than incidental instruction? We predicted that the increase in word knowledge resulting from rich and embedded instruction would be higher than the increase in word knowledge resulting from incidental instruction when initial knowledge of word meaning is controlled.

Method

Participants

Children participants

Convenience sampling was used. Forty-three children were recruited from three K3 classes (final year of preschool education) in a Hong Kong kindergarten. Parental consent was sought for each child. The participating school was a laboratory kindergarten affiliated to a teacher education institution. The school adopted a child-centered exploratory curriculum which designated children's mother tongue as the main medium of expression. Hence, Cantonese was used as the main medium of teaching and learning. English learning was provided as an additional activity. English lessons were conducted three times a week, with each session lasting for 30 minutes per session. Lessons were conducted by the home room teacher who was a Cantonese speaker. No textbooks were used nor any formal and systematic phonic teaching provided.

All participating children except three in this study spoke Cantonese at home. The reported home language used was Cantonese except for the three children who had non-Cantonese speaking parents. These three children were excluded from the final analysis because the current study focused on the Chinese EFL population. Another three children were absent for the pretest. The final number of children who completed the pretest was 37. During instruction implementation, 4 children were absent for more than one session for the week. Post-test and delayed post test scores of those children were excluded from final analysis. Another 3 children were absent for the delayed posttest. Therefore, the final sample comprised of 30 children, with 8 children from Class A, 12 from Class B and 10 from Class C. There were no meaningful differences in terms of the word knowledge of target words and general receptive vocabulary for students from Class A, B, and C during the pre-test.

There were 12 boys and 18 girls in the final sample with a mean age of 5.10 years (range = 4.67-5.42 years; $SD = 0.20$). All of them were reported by teachers as having no sign of developmental disabilities.

The project teacher

An experienced kindergarten teacher who possessed a Master's degree in Early Childhood Education and had 7 years of teaching experiences in preschool was trained by the second author who developed the lesson plans. The project teacher delivered the three instructional conditions

to all the three classes. Prior to the intervention, two training meetings were held with the project teacher who was guided through the lesson plans developed by the second author. As she was the one who did the teaching all throughout, we believe that there was a fairly high fidelity of implementation and the standardization of administration across interventions. Focus of each instructional condition were explained and examples given.

Research Design

The study adopted a within–subject quasi-experimental intervention design. Children from three classes (Class A - morning class, Class B- afternoon class and Class C - whole day class) were exposed to all the three intervention conditions conducted by the project teacher: rich instruction (R), embedded instruction (E) and incidental instruction (I) over a period of three weeks. Each participant served as his/her own control. Three different storybooks which were suitable for the English proficiency level of the participants, one for each week, were used for the three-week intervention. Each reading session lasted for 30 minutes. A word list of four target words was compiled for each story and the words were exposed to the three groups of children in different intervention conditions in a random order. The instructional structure and design were identical across the three conditions. Each session lasted approximately 30 minutes, with 15 minutes dedicated to reading the storybook and introducing target words and 15 minutes dedicated to postreading and vocabulary activities (see Table 1). All instructions were implemented by a part-time English teacher who was provided with training in implementing the instructions by the researchers.

The scoring system of the present study was based on the instructional approach. The word knowledge scores were calculated for each participant based on the instructional approach for which the word was taught. The order of instruction is presented in Table 1.

Table 1

Instructional approaches for the three classrooms

| Weeks | Class A | Class B | Class C |
|--------|------------------------|------------------------|----------------------|
| Week 1 | Rich instruction | Embedded instruction | Incidental |
| Week 2 | Embedded instruction | Incidental instruction | Rich instruction |
| Week 3 | Incidental instruction | Rich instruction | Embedded instruction |

The Three Instructional Conditions

The intervention was developed around three storybooks on the theme of “animals” as this theme was very popular among kindergarten children. The three stories chosen were: *Mr Gumpy’s Outing* by John Burningham (1970), *I Know an Old Lady who Swallowed a Fly* by Stephen Gulbis (2001) and *On the First Day of Grade School* by Emily Brenner (2004).

Four target words were selected from each story, adding to a total of 12 words. Wordlist A consisted of four target words chosen from *Mr Gumpy’s Outing*. They were *squabble*, *hope*, *tease* and *flap*. Wordlist B contained words chosen from *I Know an Old Lady who Swallowed a Fly*. These were *lady*, *throat*, *swallow* and *tickle*. Wordlist C was comprised of words taken from the story *On the First Day of Grade School*. They were *python*, *squeeze*, *tiptoe* and *snoring*.

Our criterion for selection was to find words that were unfamiliar to children but whose meanings could be easily understood. The fact that students had no prior knowledge of any of the target words made it easier for us to test the strength of the intervention accurately. Teachers were consulted about the choice of words as they had first-hand knowledge on what was familiar to children and what was not. The twelve target words included two nouns (*lady*, *python*) and ten verbs (*squabble*, *hop*, *tease*, *flap*, *swallow*, *tickle*, *tiptoe*, *squeeze*, *snoring*).

Each of the three wordlists was developed into three versions of the instructional conditions, two intervention conditions (i.e., words taught with embedded instruction or rich instruction) and a comparison condition (i.e., words receiving only incidental exposure). We developed detailed and scripted lesson plans for all instructional conditions. Based on the principles for each instructional approach, learning activities were designed for the chosen target words. The learning activities were age-appropriate and fun in nature. Each version included four words that were taught using rich instruction, four words taught using embedded instruction, and four words

receiving incidental exposure.

Rich (extended) instruction

Before the storybook reading, target words were introduced to children with the help of picture cards as magic words to be learnt. While reading, the project teacher paused at the target words and gave the definition in a child-friendly language, helped children to pronounce the words, and discussed the meaning of the words in-depth by giving additional examples of the words' usage beyond the story context. For example, in explaining the word 'squeeze', the following examples were used. "the python squeezes something by curling up its body round the thing", "we squeeze the bottle of ketchup by closing our fist". After reading, extended activities that went beyond the story context such as games, role play or singing were used to offer opportunities for children to apply the target words in other contexts. For example, to consolidate children's understanding of the target word "tiptoe," a game was designed to let children experience the meaning of the word in a context other than the story. Children were instructed to tiptoe around a sleeping pig (role-played by the teacher). Children were taught to chant an invented rhyme "snoring pig, snoring pig, you can't hear me, tiptoe one, tiptoe two..." When the count reached to 10, the teacher who pretended to be the snoring pig woke up and tried to catch the child. The child caught will take on the role of the snoring pig. The game was repeated until all the children were caught.

Embedded instruction

The project teacher followed similar procedures as that in the rich instruction by introducing the target words at the beginning, providing definitions to the target words with the help of the picture cards as the words appeared in the story and conducting the extended activities such as role playing and language games to assist children in their understanding of the target words. However, embedded instruction was distinct from rich instruction in its depth in the explanation of the target words. While reading, simpler and briefer explanations were given for the target word (e.g. "a python is a very big snake"). The extended activities did not go beyond the story contexts and children were not encouraged to apply the words in other contexts. An example of the extended activities in the embedded exposure condition was that the teacher re-told the story with children pretending to be the animals by wearing different animal headbands.

Incidental instruction

Words in the incidental exposure condition appeared in the story but were not highlighted and taught explicitly. Children heard these words in the context of the story once per storybook reading. The project teacher did not directly discuss these words at any time.

Fidelity of Implementation

Lesson observations were conducted once per week for each class by the researchers. Altogether nine lessons were observed, one lesson for each classroom every week. We developed a simple checklist for fidelity check. There were 5 items on the checklist and the observer responded either Yes or No to all the items: (1) did the teacher deliver each instructional element; (2) did the teacher model procedures appropriately; (3) did the teacher maximize opportunities to respond; (4) did the teacher provide error correction and (5) did the teacher read storybooks with enthusiasm? Over 95% of “Yes” was recorded and thus indicated a substantial level of implementation fidelity.

Measures

Participating children were assessed before and after the intervention. An eight-week follow-up posttest was conducted. All measures were administered individually with instructions in Cantonese, which is the participants’ native language, by trained experimenters (training provided by the first author on conducting children assessments). The English items were orally presented in English. Three researcher-developed measures were designed based on the wordlists to assess the effects of the intervention on children’s vocabulary learning. The Peabody Picture Vocabulary Test (PPVT), a standardized test of receptive vocabulary, was used to assess the overall receptive vocabulary knowledge before and after the intervention was implemented.

Expressive measure of story word definitions (expressive definitions)

The expressive definitions measure assessed students' knowledge of the target word definitions. The experimenter pronounced the 12 target words one by one and children were asked to give the definition for each target word. For example, for the target word *squabble*, children were asked, "What does the word squabble mean?" Responses were recorded verbatim. If the child failed to give any response after five seconds, the experimenter then asked a follow-up question, "Tell me

anything else you know about the word squabble." Responses to both prompts were scored together. Two points were given for a complete response (e.g., "squabble means quarreling"), one point for a partial or related response (e.g., "people squabble when they are angry"), and zero point for an unrelated response or no response. Given the low oral proficiency of the current sample, Cantonese responses were scored based on the corresponding meanings given to a target word.

Receptive measure of target word definitions (receptive definitions)

The receptive definition measure assessed receptive knowledge of the target words. Children were read questions that required a yes or no answer. Each target word was represented by two questions. One of them corresponded to the correct definition and one corresponded to an incorrect definition (Beck & McKeown, 2007). For example, for the target word squabble, the questions were "Is squabble quarreling with each other loudly?" and "Does squabble mean sad?". Children received one point for each correct answer and zero point for each incorrect answer. The questions for each of the target words were distributed across the measure. As mentioned, due to low English proficiency of the participants, the questions were translated into Cantonese and each question was orally presented to participants. A total score was calculated for target words introduced in different instructional conditions (i.e., rich, extended, incidental instruction). Maximum score for each target word is two points. The maximum score a child could receive for each condition was eight (i.e., two points possible for each of four words).

Receptive measure of target words in context

The context knowledge measure assessed receptive knowledge of the target words provided in neutral contexts. This test was designed to assess high levels of target word knowledge by requiring children to make finer discriminations about word meanings. Children were asked questions that required a yes or no answer (Beck & McKeown, 2007). Each target word was represented by two questions. For example, for *squabble*, the two questions were, "If someone makes you angry, would you squabble?" and "If your teacher praises you, would you squabble?" Children received one point for each correct answer and zero points for each incorrect answer. Each question was presented orally in both English and Cantonese. Responses in Cantonese were coded based on the corresponding meaning in English. Each child received a separate total score

for target words introduced within each instructional condition (i.e., rich, extended, incidental instruction). The maximum score for each instructional condition was eight (i.e., two points possible for each of four words).

Overall English receptive vocabulary

The Peabody Picture Vocabulary Test (PPVT-III) Form IIIA (Dunn & Dunn, 1997) was used to measure English receptive vocabulary. The experimenter orally presented a series of 36 words from the two to six year-old subset. Students were asked to point to one of four pictures to identify each vocabulary word. One point was given for every correctly identified word; the maximum possible score was 36. The stopping rule was set at 8 consecutive wrong responses. The internal consistency reliability was acceptable ($\alpha = 0.80$).

Procedures

Data were collected one week before the intervention, after the intervention and eight weeks after the intervention. The PPVT was only administered during the pretest and all other measures were administered across all the three time points. The assessment was finished within 20 minutes in the observation room of the kindergarten for each child. Scoring was performed by the first author who was blind to the instructional conditions of the target words.

Results

Class Comparisons before Instruction

We administered one measure tapping general vocabulary and three measures tapping the knowledge of target word meanings. The mean scores of these measures are presented in Table 2. Generally speaking, the initial knowledge of the target words was low. There were high standard deviations for the expressive definition measure. It is possibly due to high individual differences in word knowledge in each class. Results from one-way Analysis of Variance (ANOVA) indicated that the three classes were not significantly different in PPVT scores ($p = 0.95$), indicating that their general vocabulary level was similar. No significant differences were found on the measures of expressive definition ($p = 0.76$), receptive definition ($p = 0.91$) and receptive knowledge in context ($p = 0.57$). The mean score of the expressive definition measure was 2.97, indicating that participating children could only provide meaning for fewer than 2 target words

(full score = 24) on average. The mean scores of the two receptive measures for all participants were only slightly higher than chance level (receptive definition, $M = 14.40$ out of 24; receptive knowledge in context, $M = 14.73$ out of 24). Scores on target word meanings at pretest indicated that participating children had a very low level of knowledge of target words before the implementation of the intervention.

Table 2

Means and Standard Deviations of PPVT and Target Word Meaning Measures at Pretest by Class

| | Range | Class A | | Class B | | Class C | |
|--------------------------------|-------|----------|-----------|----------|-----------|----------|-----------|
| | | <i>M</i> | <i>SD</i> | <i>M</i> | <i>SD</i> | <i>M</i> | <i>SD</i> |
| PPVT | 16-34 | 24.43 | 3.78 | 24.15 | 5.64 | 23.60 | 6.06 |
| Expressive definitions | 0-15 | 2.00 | 4.04 | 3.23 | 3.22 | 3.30 | 4.60 |
| Receptive definition | 9-22 | 14.00 | 2.45 | 14.60 | 3.81 | 14.75 | 2.84 |
| Receptive knowledge in context | 11-22 | 14.00 | 2.58 | 14.61 | 3.07 | 15.4 | 2.22 |

Comparison of Instructional Conditions on Knowledge of Target Word Meaning

Descriptive statistics of knowledge of target words at posttest and delayed posttest by instructional conditions are presented in Table 3. A series of repeated measure ANOVAs were conducted to examine the differences in knowledge of word meaning in the three conditions and whether the differences were maintained over time. The two within-subject variables were instructional conditions (incidental vs. embedded vs. rich) and time (posttest vs. delayed posttest). Cohen’s d effect sizes are summarized in Table 4.

Table 3

Means and Standard Deviations of PPVT and Target Word Meaning Measures at Posttest and Delayed Posttest by Instructional Conditions

| | Range | Incidental | | Embedded | | Rich | |
|-----------------------------------|-------|------------|-----------|----------|-----------|----------|-----------|
| | | <i>M</i> | <i>SD</i> | <i>M</i> | <i>SD</i> | <i>M</i> | <i>SD</i> |
| Expressive definitions | | | | | | | |
| Posttest | | 2.10 | 0.38 | 2.40 | 0.44 | 3.13 | 0.43 |
| Delayed Posttest | | 2.23 | 0.40 | 2.33 | 0.42 | 3.57 | 0.45 |
| Receptive definition | | | | | | | |
| Posttest | | 4.93 | 0.26 | 5.35 | 0.24 | 5.83 | 0.23 |
| Delayed Posttest | | 5.31 | 0.22 | 5.28 | 0.25 | 6.00 | 0.24 |
| Receptive knowledge in context | | | | | | | |
| Posttest | | 5.23 | 0.20 | 5.47 | 0.23 | 5.57 | 0.22 |
| Delayed Posttest | | 5.60 | 0.21 | 6.07 | 0.22 | 5.77 | 0.24 |

Table 4

Cohen's *d* Effects Sizes of Instructional Conditions at Posttest and Delayed Posttest

| | Incidental vs. Embedded | Embedded vs. Rich | Rich vs. Incidental |
|------------------------|----------------------------|----------------------|------------------------|
| Expressive definitions | | | |
| Posttest | 0.73 | 1.68 | 2.54* |
| Delayed Posttest | 0.24 | 2.85* | 3.15** |
| Receptive definition | | | |
| Posttest | 1.68 | 2.04 | 3.67** |
| Delayed Posttest | 0.13 | 2.94* | 3.00** |

* $p < 0.05$; ** $p < 0.01$

For the expressive definition measure, the main effect of instructional condition was significant (Wilk's $\Lambda = 0.78$, $F = 4.00$, $p = 0.03$) whereas the main effect of time was insignificant (Wilk's $\Lambda = 0.97$, $F = 0.93$, $p = 0.34$). The interaction effect was also not significant (Wilk's $\Lambda = 0.93$, $F = 1.10$, $p = 0.07$). The significant main effect of instructional condition

indicated that there were significant differences on expressive definition of words for students who were taught using different instructional conditions. Follow-up paired-sample t-tests indicated that, at posttest, expressive definition of word was significantly higher for students in the rich instruction condition compared to the incidental instruction ($t = 2.33, p = 0.002$). At delayed posttest, there were significant differences on expressive definition of words between students who used the embedded method and the rich method ($t = 2.16, p = 0.04$) and on words taught by the rich method and the incidental method ($t = 2.86, p = 0.008$). The non-significant main effect of time indicated that the effects of instruction were maintained over time.

For the repeated measure ANOVA on receptive definition, the main effect of instructional condition was significant (Wilk's Lamda = 0.63, $F = 7.93, p = 0.03$) whereas the main effect of time was not significant (Wilk's Lamda = 0.95, $F = 1.42, p = 0.24$). The interaction effect was also not significant (Wilk's Lamda = 0.95, $F = 0.65, p = 0.53$). Follow up paired sample t-test was conducted to examine mean differences between instructional conditions. Again, the only significant mean difference was found between receptive definition of words for students who were taught using incidental instruction and those using the rich instruction ($t = 2.73, p = 0.01$) at posttest. At delayed posttest, there were significant differences between words taught by embedded and rich instruction ($t = 2.58, p = 0.02$) and words taught by incidental and rich instruction ($t = 3.84, p = 0.001$).

Results from the repeated measures ANOVA indicated that there was no significant effect of instructional condition and time on the receptive knowledge of words in context.

Discussion

The purpose of the present study was to evaluate the effectiveness of embedded and rich instruction as compared to incidental instruction for the teaching of word meaning to Chinese EFL kindergarten students in an authentic classroom setting. There are three major findings: (1) as expected, rich instruction was more effective than incidental instruction in teaching word meanings; (2) unexpectedly, embedded instruction was not more effective than the incidental instruction; and (3) the effect on word learning from rich instruction was maintained over an eight-week period. Taken together, our results suggest that rich instruction could be more effective than both the embedded and incidental approach in teaching vocabulary to young Hong Kong Chinese EFL children.

We have demonstrated that rich instruction resulted in more word knowledge gain than embedded and incidental instruction on two of our measures: receptive definition and expressive definition. Large effect sizes were observed at posttest and delayed posttest, favoring rich instruction when compared with the other two instructional methods. Our findings are in line with past studies showing that EFL and bilingual children need explicit instruction in vocabulary learning (August & Shanahan, 2006; Collins, 2005). For beginning English readers, effective vocabulary instruction with storybook reading should be direct and explicit, provide children with explanations, engage them in deep processing of words, and relate the target words to their background knowledge (Senechal, 1997). We have extended the existing literature by showing that direct and explicit vocabulary instruction is effective in enhancing knowledge of word meaning for Hong Kong Chinese EFL children who tend to have low English proficiency level.

Nevertheless, regarding how explicit or how much support is needed, our findings suggest that Chinese EFL children, particularly those with low proficiency, need to be provided with detailed explanations of target words and be given support in applying the target words in other settings, which are the crucial elements of rich instruction (Beck & McKeown, 2007, Beck et al., 2002; Silverman, 2007). Embedded instruction, on the other hand, provides brief explanations of meanings and learning activities of word exploration within the story context. Our findings suggest that embedded instruction may lack the intensity of vocabulary teaching needed and therefore is not that effective in increasing knowledge of word meaning for Hong Kong Chinese EFL children who are just beginning to learn English and who have limited exposure to oral English in daily life.

It is also important to note that although there were significant gains in knowledge of expressive and receptive definition of target words, the posttest and delayed posttest scores of these two measures were substantially below the full score (three and six points out of eight for expressive and receptive definition respectively). This means that children were not able to identify or produce word meanings for some words and the rich instruction may only be effective in teaching some of the target words. It is also possible that the instructional time was not sufficient for children to learn all the target words. Our study taught four target words per week which was more than other similar studies on L1 children with comparable instructional time (e.g., Cena et al., 2013; Coyne et al., 2007). Another major difference between the current study and existing training studies is that our study was implemented in a whole class setting with 10-

12 children whereas previous studies were usually conducted in small groups of 4-6 children (e.g., Coyne et al., 2009). More research is needed to document characteristics of effective vocabulary instruction for EFL learners (Marulis & Neuman, 2010).

A surprising finding relates to the non-significant change in receptive definition in context at posttest and delayed posttest. This measure was more difficult than the receptive definition task and required a deeper understanding of words applied in novel contexts. The findings suggest that all three instructional approaches were only able to enhance partial word knowledge and were not helpful to students in developing more complete word knowledge. It might be particularly difficult for Chinese EFL children to apply the newly learned words to novel contexts and more instructional effort is needed to help these children to develop complete knowledge of word meanings.

It is encouraging that students' enhanced knowledge in word meaning can be maintained over time. Our results indicated that the loss of knowledge between posttest and delayed posttest was not statistically significant. Examination of mean scores across conditions showed that, surprisingly, there were slight increases in knowledge of word meanings. This is inconsistent with past research demonstrating a deterioration of word knowledge over time after intervention, particularly for expressive measures (e.g., Coyne et al., 2009). Class teachers from the participating kindergarten reported that they did some forms of review with children. Children may also encounter the target words incidentally in or out of the classroom because the target words in the present study were simple English words. These may account for the slight increase at delayed posttest in word knowledge. It should be noted that although children were provided with some forms of review, the review was similar across conditions because the participating kindergarten had the same curriculum and learning activities across different classrooms.

The findings of the present study have strong implications for the teaching of vocabulary to young EFL learners who are not educated in an English environment. We have presented evidence that EFL children are able to learn word meanings in a short period of time with appropriate and carefully designed instruction within a storytelling context. Our study was conducted in a real classroom setting and the intervention was delivered by an early childhood teacher who was provided with training and support. Our results demonstrated that Hong Kong Chinese EFL children may need to be taught explicitly and be engaged in deep processing of words. Storytelling provides an interesting context for young EFL children to engage in various

kinds of learning activities. Our findings show that embedded instruction is not as effective instruction as rich instruction in enhancing knowledge of word meaning in Hong Kong Chinese EFL children who need extra support in word learning because of their limited exposure time to English and weak initial vocabulary. Given the high emphasis on learning of print in Hong Kong (Ng & Rao, 2013) and other Asian educational contexts, our study calls for a reconsideration of English learning curriculum in these places and demonstrates the need to incorporate rich word learning activities for young EFL children.

Limitations

There were a number of limitations in this study. First, the sample size of the study was small and there was considerable attrition. In addition, we only involved children from one kindergarten. It is important to replicate current findings with a larger sample drawing from different schools and varying social backgrounds.

Second, given the short duration of the present instruction, we have not included standardized measure as outcomes and only used proximal measures to assess the instructional effects. The effects of vocabulary instruction on standardized measures are reported to be significantly lower than author-created measures (Marulis & Neuman, 2010). Author-created measures are regarded as more sensitive and are construed as more proximal indicators of instructional effects (National Reading Panel, 2000). Future research could investigate the effects of long-term instruction and its impact on standardized tests and more distal measures.

Conclusion

To conclude, the current research presents encouraging findings that a rich instructional method within the context of storybook reading may be a promising approach for Hong Kong Chinese EFL kindergarteners to learn English vocabulary. In our study, both embedded and incidental instructional methods were not as effective in enhancing vocabulary knowledge. With carefully-designed instruction, our results suggest that young EFL children are able to learn word meanings with multiple story readings and follow-up learning activities that promote deep processing of words in a short duration.

References

- August, D., Carlo, M., Dressler, C., & Snow, C. (2005). The Critical Role of Vocabulary Development for English Language Learners. *Learning Disabilities Research & Practice, 20*, 50-57.
- August, D., & Shanahan, T. (2006). Synthesis: Instruction and professional development. In D. August & T. Shanahan (Eds.), *Developing Literacy in Second-Language Learners: Report of the National Literacy Panel on Language-Minority Children and Youth*, (pp. 351-363). Mahwah, NJ: Lawrence Erlbaum Associates Publishers.
- Babayigit, S. (2015). The relations between word reading, oral language, and reading comprehension in children who speak English as a first (L1) and second language (L2): A multigroup structural analysis. *Reading and Writing, 28*, 527-544.
- Beck, I. L., & McKeown, M. G. (2007). Increasing young low-income children's oral vocabulary repertoires through rich and focused instruction. *Elementary School Journal, 107*, 251-271.
- Beck, I. L., McKeown, M. G., & Kucan, L. (2002). *Bringing words to life: Robust vocabulary instruction*. New York. Guilford.
- Beck, I. L., Perfetti, C. A., & McKeown, M. G. (1982). Effects of long-term vocabulary instruction on lexical access and reading comprehension, *Journal of Educational Psychology, 74*, 506-520.
- Biemiller, A., & Boote, C. (2006). An effective method for building meaning vocabulary in primary grades. *Journal of Educational Psychology, 98*, 44-62.
- Bus, A. G., van Ijzendoorn, M. H., & Pellegrini, A. D. (1995). Joint book reading makes for success in learning to read: A metaanalysis on intergenerational transmission of literacy. *Review of Educational Research, 65*, 1-21.
- Collins, M. (2005). ESL preschoolers' English vocabulary acquisition from storybook reading. *Reading Research Quarterly, 40*, 406-408.
- Cena, J., Baker, D. L., Kame'enui, E. J., Baker, S. K., Park, Y., & Smolkowski, K. (2013). The impact of a systematic and explicit vocabulary intervention in Spanish with Spanish-speaking English learners in first grade. *Reading and Writing, 26*, 1289-1316.
- Coyne, M. D., McCoach, D. B., & Kapp, S. (2007). Vocabulary intervention for kindergarten students: Comparing extended instruction to embedded instruction and incidental

- exposure. *Learning Disability Quarterly*, 30, 74–88.
- Coyne, M. D., McCoach, D. B., Loftus, S., Zipoli, R., Jr., & Kapp, S. (2009). Direct vocabulary instruction in kindergarten: Teaching for breadth versus depth. *Elementary School Journal*, 110, 1-18.
- Dickinson, D. K., Griffith, J. A., Golinkoff, R. M., Hirsh-Pasek, K. (2012). How reading books fosters language development around the world. *Child Development Research. Volume 2012*, Article ID 602807. doi:10.1155/2012/602807
- Dunn, L. M. & Dunn, L. M. (1997). *Peabody picture vocabulary test*. (3rd edn). Circle Pines, MN: American Guidance Service.
- Elley, W. B. (1989). Vocabulary acquisition from listening to stories, *Reading Research Quarterly*, 24, 174-187.
- Gersten, R. (1998). Recent advances in instructional research for students with learning disabilities: An overview. *Disabilities Research and Practice*, 13, 162-170.
- Gersten, R., & Baker, S. (2000). What we know about effective instructional practices for English-language learners. *Exceptional Children*, 66, 454–470.
- Jeon, E. H., & Yamashita, J. (2014). L2 reading comprehension and its correlates: A meta-analysis. *Language Learning*, 64, 160-212.
- Justice, L. M., Meier, J., & Walpole, S. (2005). Learning new words from storybooks: An efficacy study with at-risk kindergartners. *American Speech-Language-Hearing Association*, 36, 17–32.
- Marulis, L. M, & Neuman, S. B. (2010). The effects of vocabulary intervention on young children’s word learning: A meta-analysis. *Review of Educational Research*, 80, 300-335.
- Maynard, K. L., Pullen, P., & Coyne, M. D. (2010). Teaching vocabulary to first-grade students through repeated shared storybook reading: A comparison of rich and basic instruction to incidental exposure. *Literacy Research and Instruction*, 49, 209-242.
- McKeown, M. G., & Beck, I. L. (2014). Effects of vocabulary instruction on measures of language processing: Comparing two approaches. *Early Childhood Research Quarterly*, 29, 520-530.
- Miller, G. A. (2003). The cognitive revolution: A historical perspective. *Trends in Cognitive Science*, 7, 141-144.
- Nagy, W. E., & Scott, J. A. (2000). Vocabulary processes. In M. L. Kamil, P. B. Mosenthal, P. D.

- Pearson, & R. Barr (Eds.), *Handbook of reading research* (Vol. 3) (pp. 69-84). Mahwah, NJ: Erlbaum.
- National Reading Panel. (2000). *Teaching children to read*. Washington, DC: National Institute of Child Health and Development.
- Ng, M. L., & Rao, N. (2013) Teaching English in Hong Kong Kindergartens: A Survey of Practices. *International Journal of Literacies, 19*, 25-47.
- Oetting, J. B., Rice, M. L, & Swank, L. K. (1995). Quick incidental learning (QUIL) of words by school-age children with and without SLI. *Journal of Speech and Hearing Research, 38*, 434-445.
- Oller, D. K., & Eilers, R. E. (2002). *Language and literacy in bilingual children*. Tonawanda, NY: Multilingual Matters.
- Penno, J. F., Wilkinson, .A.G., & Moore, D, W. (2002). Vocabulary acquisition from teacher explanation and repeated listening to stories: Do they overcome the Matthew Effect? *Journal of Educational Psychology, 94*, 23-33.
- Robbins, C, S., & Ehri, L. C. (1994). Reading storybooks to kindergartners helps them learn new vocabulary words. *Journal of Educational Psychology, 86*, 54-64.
- Senechal, M., & Cornell, E, H. (1993). Vocabulary acquisition through shared reading experiences. *Reading Research Quarterly, 28*, 360-374,
- Senechal, M., Thomas, E., fs Monker, J. (1995). Individual differences in 4-year-old children's acquisition of vocabulary during storybook reading. *Journal of Educational Psychology, 87*, 218-229.
- Silverman, R. (2007). Vocabulary development of English-language and English-only learners in kindergarten. *Elementary School Journal, 107*, 365–384.
- Slavin, R. E., & Cheung, A. (2003). A synthesis of research on language of reading instruction for English language learners. *Review of Educational Research, 75*, 247-284.
- Stahl, S. A. (1991), Beyond the instrumentalist hypothesis: Some relationships between word meanings and comprehension. In P. J. Schwanenflugel (Ed.), *The psychology of word meanings*, (pp. 157-186). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Zipoli, R. P., Coyne, M. D., McCoach, D. B. (2011). Enhancing vocabulary intervention for kindergarten students: Strategic integration of semantically related and embedded word review. *Remedial and Special Education, 32*, 131-143.

The Effects and Implications of Implementing Oral Presentations in an Omani ICLHE Classroom

Ali-Al-Issa

Sultan Qaboos University

Bio data

Dr. Ali Al-Issa is currently an Associate Professor of TESOL at Sultan Qaboos University, Sultanate of Oman. He has presented and published widely about ELT in Oman. He received several international and local research and teaching awards and recognitions. His research interests include language education and teacher training and education.

Abstract

This study examines the role of oral presentations in an Integrating Content and Language in Higher Education (ICLHE) classroom at the College of Law, Sultan Qaboos University (SQU) Sultanate of Oman. Forty four mixed-ability English for Law students were divided into 11 groups with each group required to present a different topic pertinent to Law in English. Each of the students was then asked to complete a 36-item self report questionnaire about the experience. The analysis of the data and discussion revealed that the participants found the experience relevant for their academic and professional lives. The data analysis and discussion further showed that such experience has positive effects on English language education policy implementation in Oman.

Keywords: Oral presentations; ICLHE; CLIL; College of Law; English

Address correspondence:
P.O. Box 3058, Ruwi 112
Sultanate of Oman

Introduction

Education in general and higher education (HE) in particular have evolved and become more student-centred, collaborative, and collective. Language education is no exception. A distinct feature of this has been student oral presentations. Students in different HE disciplines around the world present topics in a target foreign language – mainly English – as a part of their course requirements through which they achieve a dual aim – demonstrate ability to use the target language competently within their subject matter and achieve mastery over that particular subject matter. This aim is particularly the case in the globalization era and age of internationalization where English has become the first and dominant international language and a means for building a global citizen. This situation is best achieved through serving multiple purposes like inter-lingual and international communication, pursuing higher education domestically and abroad, acquiring science and technology from infinite sources, finding a white-collar job locally and globally, and cultural analysis and understanding of the different independent and fused and integrated English language cultures worldwide. This is particularly evident in Third World contexts, like the Sultanate of Oman and its five neighbours in the Gulf, which have embraced English as a tool for modernization and a medium of instruction in their various HE disciplines and programmes. This is since English and HE move parallel today and strongly complement each other in a world largely dominated by business relationships and mutual economic interests and benefits (Coleman, 2006).

Review of Literature

Presenting about various legal topics in English is one important strategy and activity that provides students with a wide array of opportunities to develop their linguistic and subject matter knowledge and beyond, especially if the topics are consistent with and covered in the syllabus (Soureshjani & Ghanbari, 2011-12). Radzuan and Kaur (2011) defined oral presentations as “a planned and rehearsed talk or speech that is not committed to memory or read directly from script, given by a presenter (sometimes more than one) to an audience or two or more people” (p. 1437). The two authors described oral presentations as a “stressful communicative event” due to the anxiety they cause to the presenting students as a result of being formally assessed by an expert.

Siriphotchanakorn (2005) considered oral presentations as an activity that can be included in the communicative language teaching-oriented classroom to help develop oral proficiency,

practice speaking, practice pronunciation, foster interaction, transfer information to the audience in a professional manner, promote productive partnership in the English as a foreign language (EFL) classroom, share ideas and opinions, promote analytical skills, reduce speech anxiety, promote linguistic fluency and accuracy, promote student-centred learning, provide beneficial learning experience, enable students to use language in real life situations, build students' confidence, increase students' speaking time, and prepare students for the job market.

Miles (2009) acknowledged that “oral presentations are becoming a more important part of language teaching, especially in the university environment” (p. 103). Oral presentations, according to Morita (2000), reflected “intellectual values and academic skills” (p. 287) and are an important part of the graduates' required daily interactive activities and contribute to their successful course and program completion and “disciplinary enculturation and apprenticeship into academic discourses and cultures” (p. 280). Morita (2000) acknowledged the complexity of oral presentations as socio-cultural and language-mediated activities and processes that contribute to students' language acquisition and cognitive development and require “analytical and critical reading and thinking skills” (p. 287).

Morita (2000) highlighted the significant role of teachers in their students' language socialization process as “competent members of the social group”, who positively affect and get affected by the entire process, which has important implications for lifelong learning. Instructors, according to Morita (2000), expect their students to demonstrate “the ability to extract main points from the literature and synthesize them in a meaningful way and the ability to articulate opinions in spoken and written communication” (p. 287) to contribute to their academic community and work independently and collaboratively, despite the complexity of the latter type of work due to the limited chances outside the classroom.

Pragorbsuk and Moore (2002) employed oral presentations on an English for Academic Purposes (EAP) course to enhance their students' English language proficiency and practice speaking. They concluded that it would be beneficial for students to deliver more than three presentations during the course since the higher the frequency of conducting oral presentations, the greater chances the students have to develop their language and content knowledge and analytical skills.

Oral presentations in an integrated content and language in higher education (ICLHE) classroom like the one under investigation in this study, therefore, can have a significant role in

creating, constructing, and interpreting students' knowledge (Coyle, 2006), empowering knowledge acquisition (Gaballo, 2010), developing thinking (Nordmeyer, 2010), and using the language authentically and as a medium of communication and information (de Graaff, Koopman, Anikina, & Westhoff, 2007) in a collaborative, real-life, and meaningful academic context and settings, which can positively influence students' confidence, expectations (Wiesemes, 2005), motivation, real-life goals, and cultural and multicultural awareness (Pinkley, 2012).

Al-Issa (2014) collected the responses of 182 graduates of the College of Law about the uses of English in the Omani job market. He found that ICLHE had a strong potential to contribute to the development of the College students' academic, cognitive, cultural, linguistic, social, and psychological aspects to meet the demands of the job market and beyond, provided faculties make the right theoretical and practical language and content choices and decisions to influence productive learning and natural and innovative language production, and hence, positive policy implementation.

de Graaff et al. (2007) stated that the main goals of Content and Language Integrated Learning (CLIL) within the Dutch context, which are also largely applicable to the Omani context, are to help students develop a better command of the target language, prepare students to study courses in English abroad, and achieve internationalization. The four authors argue that fluency in a CLIL classroom can be achieved best by using authentic materials leading to tasks that engage students in interactive language use. de Graaff et al. (2007) stated that “. . . CLIL involves additional language learning objectives and specific opportunities for communication and language use” (p. 606) and that experience generating from using communicative language teaching, content-based language teaching, and task-based language teaching has led to the emergence of CLIL.

de Graaff et al. (2007) emphasized that the role of effective teachers in a CLIL classroom is represented in being motivators and providers of opportunities for exposure to language form and meaning and functional and communicative use of the target language to help facilitate their students' language acquisition. de Graaff, Koopman and Westhoff (2007) considered CLIL teachers in general responsible for carefully preparing and presenting adequate process-based tasks and activities to their students that not only successfully integrate language and content (Lorenzo, Casal, & Moore, 2009), but also strike the right balance between them (Gaballo, 2010).

de Graaff, Koopman, Anikina, and Westhoff (2007) constructed an observation tool to find evidence for teaching performance promoting language acquisition within three medium-sized Dutch CLIL secondary schools contexts. The lessons were videotaped and analyzed and the target teachers were interviewed. Sufficient evidence was collected about the observed teachers' effective performance and ability to facilitate form-focused processing than teaching grammar explicitly or explaining grammar rules. Use of authentic materials and adapted texts up to the level of the students and scaffolding on the content and language level facilitated exposure to language input at a challenging level. Teachers also facilitated meaning-focused processing through using implicit and explicit corrective feedback on incorrect meaning identification and practice the target language through relevant written and oral assignments. Furthermore, teachers gave examples, used recasts and confirmation checks, clarification requests, and gave feedback to facilitate form-focused processing. As far as output production is concerned, teachers engaged students in feasible interactive and communicative tasks like presentations, for instance. Teachers here gave students time to complete those tasks, encouraged them to interact in the target language, provided feedback on students' incorrect language, and stimulated peer feedback. Last but not least, teachers facilitated the use of compensation strategies to help their students overcome language comprehension and production problems.

Mehisto and Asser (2007) stated that one of the aims of the Estonian CLIL program, which are also applicable to the Omani context, is to help students achieve "the cognitive and social skills and habits required for success in an ever changing world" (p. 684). Amongst the best practice The Estonian Centre outlined are focusing on content, supporting language learning, building learning on students' existing knowledge, skills, experience, and attitudes, building student confidence to experiment with language, negotiating the meaning of language and content with students, creating opportunities for students to communicate, organizing learning through themes and projects, focusing on accommodating student interests, letting the students indicate the language they need, and following the national curriculum. The same center deemed strategies like active learning, peer cooperative work, students centered learning, responding to different learning styles, and teachers acting as facilitators as complementary best practices for CLIL programming (Mehisto & Asser, 2007).

Mehisto and Asser (2007) acknowledged the central role of teachers to educational programme development. The two authors asked 41 teachers to describe their application of

CLIL methodology based on a predetermined set of selected strategies. They found that students were engaged in learning due to the fact that teachers successfully delivered quality CLIL programming as represented in setting content and language goals, evaluation of students, analysis of learning process, and pair and group work. Teachers were further found demonstrating a genuine interest in their students' welfare.

Gaballo (2010) analyzed an Italian ICLHE scenario based on the use of information and communication technology (ICT) through involving students in activities, which provided them with “. . . a combination of professional expertise and linguistic competence” (p. 1). Students enrolled in this course had limited English language proficiency and their exposure to English was restricted to the classroom and the Internet, a case which is largely similar to the one under investigation in this study. Students were required to be active participants and inquirers through being exposed to complex information and demanding activities focusing on critical thinking, to help them develop as dynamic knowledge acquirers and problem solvers – two integral skills for their future jobs. Students were assessed on the basis of their “. . . understanding of the subject matter and their capacity to engage with the subject matter through analysis, evaluation and speculation” (p. 5). Students, according to the writer, developed intrinsic motivation and acquired “a third language”, which is “a specialized variant of L2 that is regulated by its own grammar and style” (p. 4).

Kasper (1997) assessed the effect of content-based instruction (CBI) – another acronym for CLIL (Banegas, 2012; Eyjólfsdóttir, 2011) – of 184 randomly selected students from an English as a second language (ESL) community college enrolled in intermediate ESL reading and writing course. The data collected suggested that CBI at the intermediate level not only enhances performance, but may also help students' subsequent performance in the college academic mainstream and increase their likelihood of earning a college degree, as it is the case with the context under investigation. Kasper (1997) argued that the type of materials CBI students read and the choices of materials that are grounded in mainstream academic disciplines have important implications for their English language development. Moreover, students were “. . . able to acquire sufficient and cumulative knowledge on the discipline(s) of interest, enabling them to handle the topic(s) on a sophisticated, university level” (p. 317). Kasper (1997) argued that students must learn to use the target language “interpretatively” and “critically” when “. . . faced with academically-oriented linguistic and rhetoric tasks” (p. 317).

Kasper (1997) also stressed the importance of ESL students using more advanced levels of language processing to help them comprehend discipline-based materials. Kasper (1997) acknowledged that as ESL students work through a discipline-based text, they become “. . . aware of how to construct meaning from information stored in memory, how to extract relevant information from the larger text context, and how to filter out redundant or irrelevant information” (p. 317-8).

Moreover, Kasper (1997) emphasized the importance of incorporating activities into the CBI ESL course and programme that promoted interactive engagement leading to facilitating information synthesis using multiple texts, “meaning construction”, and “enhanced linguistic proficiency” (p. 318).

Kasper (1997) additionally highlighted the significant choice of discipline-based texts in CBI ESL courses and programs to help motivate and encourage students to relate those texts to their own experience, knowledge, ideas, reflections, and information to help them view the information presented from different perspectives and synthesize it to make the necessary inter-textual and interdisciplinary connections. Such choices, according to Kasper (1997), also impacted students’ construction of schemata and growth of cognitive-intellectual interactive skills and abilities leading to efficient use of comprehension strategies and development of the target language proficiency and critical thinking skills, which are necessary for a successful academic experience.

In his review of literature about CBI, Stoller (2004) discussed different case studies that address different challenges CBI faculties encountered at the tertiary level. Amongst those challenges were orchestrating a careful transition from familiar to unfamiliar content to help students overcome their frustrations through minimizing the sudden jumps from easy to difficult. Stoller (2004) also discussed implementing a theme-based approach, exposing students to real data, and involving them in its collection to help them acquire a range of literacies associated with learning. Stoller (2004) additionally discussed various case studies, which integrated one subject area into language classes for an extended period of time. Teaching Law through English over two semesters, as it is the case at the College of Law at SQU, serves a good example. Stoller (2004) described this sustained content-language teaching approach as stimulating, versatile, useful for integrating language and content, explicitly teaching language and academic skills, and one which deeply engages students with content.

Thus, teachers at the tertiary level face three challenges when they switch to a content-based curriculum. Stoller (2004) included (1) the determination of course content in response to diverse student interests, (2) the selection of content resources and the designation of targeted grammar points for students with varied proficiency levels, and (3) the sequencing of structured input and output activities (p. 266).

On the other hand, Seikkula-Leino (2007) investigated how 217 pupils from Grades Five and Six in a Finnish comprehensive school, where 116 of them were enrolled in CLIL classes, had successfully learnt content in CLIL. She found that pupils taught in Finnish overachieved more strongly than their CLIL counterparts. She further found that while CLIL students were moderately more motivated to study and use a foreign language than those learning in Finnish, “pupils in CLIL felt much weaker in understanding, reading, writing, and speaking skills of the foreign language than the pupils who were studying in Finnish” and were “generally weaker as language learners” (p. 337) than their Finnish counterparts. Teachers in the study reported their pupils facing difficulties in learning content in a foreign language and described it as challenging and very demanding as pupils are required to resolve communicative problems while operating in a foreign language, which requires them to be extremely active. The author further attributed this situation to CLIL involving a lot of language which is above the pupils’ current language competence, and hence, required a great deal of concentration and can negatively affect learners’ self-esteem and self-concept.

In his comparative study on language education policy and planning, Wannagat (2007) examined the processes and effects of language learning in L2-taught content subjects in two Grade Seven immersion programs – CLIL German and English medium instruction (EMI) Hong Kong. Both programs had their similarities and differences. The latter was more explicit in terms of the theory and objectives underlying learning and teaching content through L2, where the aim of CLIL supports the development of language and subject matter skills. The author concludes that while students in the Hong Kong model were exposed to English in most content subjects when compared to their German counterparts who were only exposed to L2 in two to three content subjects, the latter group of students had increased opportunities to develop their constructive abilities in L2. Students, in the CLIL model, were active and autonomous communicators, who were involved in completing a task, which engaged them in constructing and verbalizing knowledge and meaning through interaction, sharing, and collaboration. This, to

Wannagat (2007), facilitated L2 acquisition and use as it encouraged testing hypotheses and noticing. Wannagat (2007) attributed the CLIL success to the planned code-switching, careful curriculum planning and integration of content and language, better training of teachers, and the strong commitment shown by students and teachers to maintain communication in L2. Wannagat (2007) suggested developing a special methodology for integrating content and language.

It is important to explain that the context under investigation is a combination of CLIL and EMI. The quote below from Morgado and Coelho (2012) perfectly summarizes the context at the College of Law at SQU in terms of students' competence and institutional policy and practice.

CLIL appears essentially as a compensatory strategy to work on the foreign language skills of students in HE when there is no adequate language provision in higher education curricula or when the students' skills in FL are low . . . EMI seems to constitute a practice introduced under internationalization pressures and the presence of many international students, with poor results in terms of effective learning and teaching (p. 134).

There are characteristics drawn from both approaches/models to help achieve one aim and educational policy agenda – help students develop their content and language skills and competence in an ever economically, politically, and socially changing, challenging, and contesting world. This is best achieved through integrating the two together in a HE classroom through acquiring academic skills in an English as a *global lingua franca* context to meet new internationalization demands, expectations, and competitive standards. It is, hence, the purpose of this study to discuss how ICLHE within the Omani context can affect advancing HE policy in the Sultanate of Oman. This will potentially enrich the pertinent literature and can have important implications for other similar contexts especially that this study is the first of its kind and no other similar studies have been published so far.

The Context

The College of Law at Sultan Qaboos University (SQU) was established in 1997 by a Royal Decree and always taught its courses in Arabic, except for two General English Language courses, which were taught in the first two semesters of the eight-semester degree plan and found

insignificant due to their shallow content, insufficient contact time, and lack of print material and technological resources, which had negatively impacted the students' overall development (Al-Issa, 2007).

However, a decision was taken in 2009 to teach 30% of the College courses in English to help serve “Omanization” (nationalization), where expatriate foreign labour, which makes approximately one third of the total population of Oman, is gradually replaced by qualified and skilled Omani personnel. The decision was motivated by the demands of the local job market to better equip university graduates with an operational command of the target language and practical legal knowledge and skills. A decision was therefore taken to teach International Law, Commercial Law, and Economics in English. These are three subfields of Law, which are largely conducted in English in the Omani job market and have been largely dominated and operated by the expatriate labour force, which has an edge over the local work force with respect to English language proficiency and practical legal knowledge and skills. This is particularly the case in the private sector, which has more and better employment chances at present than its public counterpart. The Omani government has stressed on more than one occasion the role of the private sector as a strategic partner in building modern Oman through participating effectively in strategic plans and projects like creating jobs for the qualified and skilled Omanis, for example.

Nonetheless, prior to embarking on any legal credit courses, students are required to take two core courses – English For Law (EFL) (1) and (2) in the first two semesters of their programme respectively, with each being taught for six contact hours per week. These are EAP courses, which aim at integrating the English language with Law to help enhance the students' skills in English and prepare them for their Law courses in the subsequent semesters. EAP courses are taught across SQU's nine colleges and all public and private colleges and universities throughout the Sultanate. This practice is motivated by the significant role of the English language for finding a white-collar job, pursuing HE domestically and abroad, acquiring science and technology, communicating locally and globally and analysing and understanding the different English speaking cultures through different channels of exposure.

Research Questions

In light of the aforementioned discussion, the following questions are asked:

1. What are the perceptions of the EFL (1) students about oral presentations?

2. Is there any statistically significant difference between the males and females' perceptions with regard to the oral presentations?
3. What effects does the implementation of oral presentations in the EFL (1), course have for the new ICLHE plan?
4. What implications do the roles of EFL (1) faculty in the design and implementation of oral presentations have for the implementation of the new ICLHE plan?

Methodology

Participants

The sample for this study consisted of one EFL (1) section taught by the researcher and comprising of 44 College of Law freshmen – 16 male and 28 female students. All students are required to take a placement test designed and written by the Language Centre (LC) at SQU to determine whether their knowledge in English qualifies them to embark on their core courses at their respective colleges, or join the English Foundation Program (EFP) and enroll in language improvement courses at the LC that enhance their language knowledge in the four skills and help them reach an equivalent level to the overall score of Band 5 on the International English Language Testing System (IELTS). It is worth noting that candidates scoring Band 5 on the IELTS are described as “modest users” of English with partial command of the language and ability to handle basic communication in their own field. It is further noteworthy that 16 of the female participants and 4 of the male participants scored over Band 5 on the IELTS prior to joining EFL (1), which makes it less than 50% of the total number of enrolled students. The rest, nonetheless, had to go through the LC courses.

Procedures

A general talk with the study participants revealed that all of them had presented different topics orally either when they were attending their EFP at the LC or at their respective schools prior to joining SQU. However, they were allowed to choose their own topics then, which were presented individually and less subject-specific and more of a general nature (shopping, food, holidays, etc.), which relied more on the students' general everyday knowledge and the textbook content. Additionally, such presentations were not assessed formally, students were not allowed any preparation time, and classrooms included a single gender for most of the time.

Oral presentations within the context of this study, fell somewhere between “guided” and “free” presentations (Al-Issa & Al-Qubtan, 2010) in terms of students’ level and choice of topics. Like the other core courses at SQU, students on EFL (1) are formally assessed out of 100 marks. 50 marks are allocated for the final exam, which tests students’ listening, reading and writing skills. This is while the other 50 marks are left for the faculty’s discretion, provided the students are aware of how these 50 marks are distributed. Most faculty staff at the College opt for a criterion-referenced mode of assessment and for giving their students a series of short tests or quizzes, which are entirely and strictly based on testing the students’ memory about the content of the mandated course textbook and which are assessed out of 40. The remaining 10 marks are almost always allocated for attendance.

In the case of this course, nevertheless, the researcher decided to allocated 10 marks for attendance, 10 marks for classroom participation and 30 marks for the oral presentation. The presentations were assessed “subjectively” depending on the overall impression the instructor forms about the presentation with as much equal attention given to “process” and “product” as possible (Al-Issa & Al-Qubtan, 2010). In other words, the 30 marks were adapted from (Al-Issa & Al-Qubtan, 2010) and distributed as follows: 10 marks for language accuracy and fluency (grammar, vocabulary, pronunciation, and spelling), 10 marks for the organization and delivery of the presentation (use of audio-visual aids, adherence to time, introduction and preview, cohesion and coherence, and summary and conclusion), and 10 marks were for handling the questions of the audience. The students were asked about whether they preferred to take two or three short tests or deliver an oral presentation. All of the female participants and many of the males voted for the latter option.

The 44 participants were then divided into 11 mixed-ability groups of four due to the existing variable levels and time limitations. The males were separated from the females due to cultural reasons. The aim of putting the students into groups was inspired by the Concept-Oriented Reading Instruction (CORI) curricula framework, “which has major implications for second and foreign language content-based curricula” (Stoller, 2004, p. 271). An additional source of inspiration for putting the students into groups is the Collaborative Strategic Reading (CSR) instructional framework (Stoller, 2004), whereby students work collaboratively to comprehend content-area texts and develop their cognitive abilities through social interaction.

Twenty topics were prepared by the researcher – the course instructor (see Appendix A). The topics were based on the content of the 14 chapters the students had studied. Each of the 14 chapters integrates the four skills and introduces and recycles certain grammatical and lexical items relevant to the students' levels and needs. Students in EFL (1) are required to study a textbook that is made of 14 chapters and designed and written by the LC at SQU. The textbook aims at developing students' competencies in the four skills, vocabulary and grammar to help them “listen actively”, “read with understanding”, “speak so others can understand” and “convey ideas in writing” (CELP Course 1 Specifications, p. 7). This is in addition to achieving a number of challenging national standards specified by Oman Accreditation Council on successful completion of the General Foundation Program, which require students to demonstrate academic skills using the target language. Topics covered in the textbook are Employment Law, Financial Regulations, Studying the Law, Contract Law, Tort Law, International Law, Criminal Law, Constitutional Law, Real Property, History of Law, The Legal System, Commercial Law, The Courtroom and Comparative Law. Such a variety of topics, along with the other Arabic-medium legal courses the students take during their first year at the College like Introduction to Law, Basic Law of State, State Theory and Political Systems, Sources of Obligations, Principles of the International Public Law, Principles of Economics, and Arabic Language For Law, can contribute to widening and broadening their technical/legal language and knowledge.

Each of the 20 topics was written on a separate slip of paper and folded and were all put in a bag. Each group was asked to draw a slip out of the bag. The groups were instructed that they would have four weeks to prepare their presentations and that they would be called on in random selection to present. They were further told that each group would be given 15 minutes to present its work and five minutes to answer any questions raised by the audience. Research has shown that longer planning time (Ortega, 1995) positively affects students' inter-language development, significantly impacts fluency, and enhances structural complexity. Planning, according to Ortega (1999), “. . . creates a space for the learner to assess task demands and available linguistic resources and to prioritize strategic allocation of effort and attention accordingly” (p. 138). Ortega (1999) further suggests that pre-task planning promotes a conscious focus on form even at lower levels of language proficiency. Moreover, pre-planners produce more fluent and lexically varied language than their on-line and no planning counterparts, who speak more slowly and monitor and edit their language production more extensively (Yuan & Ellis, 2003).

It was not possible to allocate all groups a specific amount of time in class in order to prepare due to the nature of the topics, which required data collection from sources beyond the literature like topics # 18 and #20, which were drawn by two of the groups. Additionally, time was another obstacle. The semester is made up of 16 weeks and mathematically it was not possible to allocate any time for in class preparation, as it would come at the expense of the formal classes. Nevertheless, when asked about where they had prepared their presentations and the time they had spent on such an activity, different answers were given in relationship to the discussion venue and time. The vast majority of the female participants, who lived on-campus in SQU hostel, prepared their presentations on campus. By contrast, the male participants, who all lived off-campus, had to meet during the evening in coffee shops like Starbucks, Second Cup, and Costa Coffee, which have plenty of branches throughout Muscat to discuss the details of their work. All these coffee shops have a Wi-Fi facility, which made it easy for the students to access the information and data they wanted since most of the topics required online search.

The teacher invited the students to contact him through the e-mail or mobile telephone in case they required any assistance with any aspects of their work during its preparation. Different students in fact raised different questions particularly about the nature of the content to include in their presentations, how it could be approached, and what performance was expected from them. The teacher provided many examples to help answer his students' questions. That was all part of negotiating the instructor's expectations (Morita, 2000).

The participants were also told that while they needed to take turns to present their work, any questions raised by the audience would be answered collectively. They were further instructed to prepare their presentations using PowerPoint slides due to the important implications it has for promoting and student-centeredness, self-directedness, student autonomy, cooperative work, authentic discourse, social interaction, knowledge restructuring, motivation, and stimulating learning environment (Apple & Kikuchi, 2007). All SQU classrooms are equipped with computers, projectors and a screen. They were also told to feel free to prepare any materials to complement their oral presentations with any additional materials that would support their presentations, as use of visuals has been found "effective" in oral presentations (Al-Issa & Al-Qubtan, 2010).

Students were asked to deliver extemporaneous presentations. They had some preparation time and spoke from an outline of pre-determined main points, but were not allowed to script

each word to be spoken to make them take a natural and conversational form. Very few of these presentations were “opinion survey projects” (Apple & Kikuchi, 2007), like #2, for example. On the contrary, most of the presentations were more of conceptual-theoretical projects, like #1, #8, #10, #13, #14, and #17, for example, and were based on systematic and careful investigation of prescribed materials and sources with an aim to synthesize a discussion that allows for deeper levels of thinking and understanding and knowledge and language construction.

The teacher sat at the back of the classroom during the presentations and took notes about the linguistic, content, and organizational aspects of the presentations. Some language errors occurred by more than one student. After each presentation, the teacher approached the front of the classroom and raised different questions to the group members pertinent to those three aspects to help stimulate the students’ memory and enhance their learning of the target language and already-internalized forms and lexis through improving their interactive and cognitive aspects of language acquisition and development (Ellis, 2009). The corrective feedback strategies employed to provide input and promote output (Ellis, 2009) were primarily explicit and included clarification requests, explicit correction, elicitation, paralinguistic signals like signalling the time and nodding to agree with and approve of an idea or shaking the head to express disapproval and disagreement with a particular idea, for example, and metalinguistic feedback like “is that how you would say it (jury) in English?” or “how do you spell “intellectual”?”, for example. Panova and Lyster (2002) and Ellis (2009) stress the importance of using these strategies, as they are more successful at prompting peer and self-repair. The teacher used one or more strategy at the same time, depending on the number and types of mistakes made. On most occasions, students acknowledged the feedback in different ways and repair of errors mainly occurred in the forms of self-repair or peer repair. The latter type occurred as a result of having mixed-ability groups.

It is noteworthy that the instructor assigned to teach the course has a Ph.D. in Education (Teaching English to Speakers of Other Languages – TESOL). He is an Associate Professor and has been teaching for over three decades. He taught EFL in various schools and colleges and worked as a teacher trainer for a number of years. He also taught ICLHE courses to students from the College of Education and College of Science at SQU. He has a wide range of publications about EFL in Oman and the new degree plan at the College of Law and presented about EFL in Oman in numerous local, regional, and international conferences.

Data Collection and Analysis

Data were collected through a survey. Questions were inspired by and derived from the existing literature about the topic of the study. The questionnaire consisted of 36-rating scale questions and aimed at identifying the perceptions of the participants about the use of presentations on the course. To each statement, the participants were asked to indicate whether they “strongly agree”, “agree”, “not sure”, “disagree” or “strongly disagree”, about the different aspects of their oral presentations.

The questionnaire questions were prepared and grouped to elicit information on how the presentations facilitated the development of the students’ academic, cognitive, cultural, linguistic, psychological, and social aspects at a personal and collective level, which are all aims of integrating content and language. The questions further attempted to elicit what strategies were used to prepare and deliver the presentations that have led to implementing an ICLHE policy that would best serve the national aspirations. Seven clusters emerged from analyzing and discussing the 36 items, which have been inspired by the pertinent literature. These are Language Acquisition and Use, Knowledge Acquisition, Decision Making, Presentation Strategies, Content Description, Personality Reinforcement, and Course Design. Analysis and discussion of these clusters has important implications for the roles of EFL (1) faculty in the design and implementation of oral presentations in the new ICLHE plan.

The questionnaire was administered with the students in the classroom to help answer any of their questions and clarify any of their doubts. At the end of the questionnaire, space was provided for the participants to provide any additional relevant comments.

A T-test was also conducted to help answer the second research question about whether there was any statistically significant difference between the males and females’ perceptions with regard to the oral presentations.

To improve questionnaire reliability, the questionnaire was piloted on four students from a different EFL (1) section. The analysis showed that the Cronbach coefficient was .89, which indicated a very satisfying internal cohesion and consistency. The validity, on the other hand, was checked through sending the questionnaire to a panel of two experts from ELT. This process subsequently helped with refining some of the questions and adding and omitting others to meet the aims of the study. The data was then analyzed descriptively in SPSS to compute the participants’ responses.

Results and Discussion

Research Question #1: What are the perceptions of the EFL (1) students about oral presentations?

The grand mean of the 36 items was 4.15, which is high and which reflects the positive perceptions of the students about the oral presentations.

Research Question #2: Is there any statistically significant difference between the males and females' perceptions with regard to the oral presentations?

The T-test result showed a significant difference in Item #6 (see Appendix B) in favour of the male students. This could be attributed to the students' schemas and their relationship with comprehension. In other words, the degree of background knowledge and familiarity of the students with the topics they drew out of the bag and the amount and nature of knowledge involved in those topics could have positively affected their motivation and comprehension and helped them focus on meaning rather than form (Salimi & Fatollahenejad, 2012).

In fact, some female students requested a change of their topics immediately after they drew them out of the bag, which was rejected on the basis of being fair to every group. This is best reflected in the following two comments made by two of the female students about topic choice. The first one wrote . . . *I really want to choose the topics ourselves*, while the other one took a more balanced approach and wrote . . . *I think it will be better if we choose the subject of our presentation, but in spite of that, the subjects that were provided by you were interesting*.

Alternatively, some the female students might have wanted a relatively straightforward topic to ensure a successful oral presentation (Morita, 2004), and hence a higher mark. Examples of the topics found difficult to present were #10 and #15. While the former had a Western cultural component in it (the jury), the latter was complex and had more than one component to consider for discussion. In a strictly and rigidly traditional and criterion-referenced assessment-based education system like the Omani one, for example, results and scores greatly determine students' success and future and drive and shape their motivation and attitudes (Al-Issa, 2010).

The results of the T-test further showed a significant difference in Item #15 in favour of the male students. This item is very closely linked to Item #6, where a significant difference in favour of the males was reported and where knowledge acquisition through the presentation topic was a factor.

While generally speaking pre-task planners give task conceptualization priority over articulation, as suggested by (Yuan & Ellis, 2003), gender, specifically, has a significant role to play in impacting topic processing, knowledge acquisition, and content comprehension, which in turn affects comprehension performance (Martinez, 2013). Martinez used two gender-neutral expository reading comprehension passages from English textbooks and two different reading comprehension assessment tasks on 68 elementary and intermediate male and female undergraduate English language students in Spain. Analysis of the data, which was extracted through a five-point scale questionnaire, showed that male students were superior to female students in comprehension. Martinez (2013) attributes this difference to the types of texts used and the reader's interest in and knowledge about certain topics, which has been found in favor of the males and could have been the case in this study too. In fact, Martinez asserts that females perform even worse, when they are unfamiliar with the texts they read.

Research Question #3: What effects does the implementation of oral presentations in the EFL (1), course have for the new CLIL plan?

Language Acquisition and Use

Table 1
Participants Mean Ratings of Language Acquisition and Use

| No. | Item | Mean | SD |
|-------------------|--|-------------|------------|
| 1 | Presenting gave me a chance to practice my English language. | 4.66 | .53 |
| 2 | Presenting gave me a chance to practice my legal English. | 4.23 | .83 |
| 4 | Presenting helped me acquire new English language. | 4.09 | 1.01 |
| 5 | Presenting helped me acquire new legal English. | 4.25 | 1.04 |
| 7 | Attending my classmates' presentations helped me acquire new English language. | 4.07 | .82 |
| 8 | Attending my classmates' presentations helped me acquire new legal English. | 4.00 | .92 |
| 10 | The post presentations questions helped me acquire new English language. | 4.02 | .95 |
| 11 | The post presentations questions helped me acquire new legal English. | 3.98 | 1.00 |
| 16 | Preparing the presentation in a group helped me acquire new English language from the other members. | 3.82 | 1.04 |
| 17 | Preparing the presentation in a group helped me acquire new legal English from the other members. | 4.02 | 1.09 |
| Grand Mean | | 4.08 | .94 |

This cluster seeks information pertinent to language acquisition and use. The grand mean of this cluster is 4.08. Nevertheless, the first item, obtained the highest mean (4.66) amongst the 10 items. It also obtained a relatively higher score than the second item (4.23).

Also, while Items #4 and Item #5 obtained a high mean (4.09 and 4.25 respectively), the latter item obtained a relatively higher score than the former. In addition, Item #7 and Item #8, obtained a relatively high mean (4.07 and 4.00 respectively). Moreover, Items #10 and #17 obtained the same score (4.02), which is relatively high. By contrast, Item #11 and Item #16 obtained the lowest mean in this cluster (3.98 and 3.82 respectively).

A general note here is that the students acquired more legal and general English and used both versions more when they presented and actively delivered information and knowledge than when they acted as an audience or worked in groups beyond the classroom boundaries. This appears evident in the scores obtained in the first four items quoted in the table, as opposed to the next six items.

As a genuine language development approach and strategy centering round a task, oral presentations empowered students to diversify their sources of exposure to the target language in its general and academic forms, manipulate it to suit the content they wanted to talk about and use it naturally and communicatively with almost no or extremely limited interference from the teacher. CLIL offers a new language focus and teaching methodology (Costa & D'Angelo, 2011) and integrates the four skills (Marsh, 1994). In addition, CLIL stresses the centrality of interaction in the classroom (Coyle, 2006) and its significant implications for contexts like Oman to allow for incidental learning and positive transfer of language (Lorenzo, Casal, & Moore, 2009). CLIL further emphasizes the purposeful linguistic interaction among learners themselves (Xanthou, 2011).

Sherris (2008) further acknowledges that assigning meaningful tasks and activities to students in meaningful contexts pertinent to the content area, as it is the case with oral presentations within the context of this study promote problem solving skills and evaluation of solutions. Sherris (2008) views this kind of powerfully active engagement as leading to language competence improvement, which has important implications for the students' academic and professional future.

Nonetheless, Marsh (1994) draws our attention to the fact that the disciplinary context determines what language is to be used and how. In other words, different contexts have different

genres, which determine what discourse rules and concepts and lexical and structural items should be used. The presentation topics selected within the context of this study tried to guarantee the integration of exposure to and practice of both types of languages – general and legal, as it appears from Appendix A.

On the contrary, Items #2, #4, and #5 (4.23, 4.9, and 4.25 respectively) did not have a high mean as Item #1. This could be due to the sources these participants had consulted. Some sources are linguistically richer than others, and hence, provide a broader and deeper exposure to the target language. One of the participants wrote that most of the data his group obtained . . . *was from the Internet, but a lot of different websites to get as much clear idea and background about the topic as possible.* In addition, Items #7 and #8 had a relatively lower mean than the last three items (4.07 and 4.00 respectively). This is possibly due to the participants' linguistic fluency. I mentioned earlier that the 44 students were mixed-ability and that some of them were linguistically less capable than the others.

Moreover, Items #10 and #11 had even a lower mean than the last two items (4.02 and 3.98 respectively). One can argue that this is based upon the quality of the questions asked by the audience and answers given by the presenters. Some of the questions were very short and superficial, which required very short answers.

Furthermore, Items #16 and #17 had the lowest mean amongst all items in this cluster (3.82 and 3.77 respectively), with Item #18 having a comparatively low mean (4.2) as well, when compared to some items with a higher mean in this cluster. This could be attributed to using Arabic as a means of communication for the groups' discussion sessions, which is a preferred learning strategy, which makes their life easy and can help students save time and energy. Besides, students feel more comfortable and secure using their mother tongue or first language to negotiate, discuss, and check different aspects of the target language like grammar and vocabulary, for example. This is especially the case if the students' target language is not adequate to help them avoid anxiety, frustration, and embarrassment.

Knowledge Acquisition

Table 2

Participants Mean Ratings of Knowledge Acquisition

| No. | Item | Mean | SD |
|-------------------|--|-------------|------------|
| 3 | Presenting gave me a chance to practice my legal knowledge. | 4.43 | .66 |
| 6 | Presenting helped me acquire new legal knowledge. | 4.21 | .79 |
| 9 | Attending my classmates' presentations helped me acquire new legal knowledge. | 4.25 | .69 |
| 12 | The post presentations questions helped me acquire new legal knowledge. | 4.00 | 1.03 |
| 18 | Preparing the presentation in a group helped me acquire new legal knowledge from the other members. | 3.91 | 1.03 |
| 19 | I consulted different legal sources when preparing my presentation. | 3.91 | .94 |
| 23 | The language I acquired from my presentation and my classmates' can help me with other oral presentations I may require to do in the future. | 4.14 | .90 |
| Grand Mean | | 4.12 | .86 |

As Table #2 shows, this cluster has a higher grand mean (4.12) than the previous one. Items #3, #9, #6, #23, and #12 obtained the highest mean respectively. Nevertheless, Item #18 and Item #19 had a comparatively lower mean when compared to the other items in this cluster (3.91). One can note here that the students acquired more legal knowledge from being in the classroom and through presenting and attending their classmates' presentation than consulting human and physical legal sources.

Within the context of this study, participants had options for legal knowledge acquisition, interpretation, and creation as represented in their classmates and sources beyond the classroom and the College boundaries as represented in the Internet, which has redefined the concept of literacy. Other sources of knowledge acquisition were certain knowledgeable individuals as it was the case with topics #18 and #19, which involved data collection from different participants in a form of a mini research conducted by those who presented the topics. One of the students wrote that her group collected their data . . . *from the official websites on the Internet and from doing interviews*. Another student wrote that she obtained her data from . . . *online journals (articles), newspapers, and magazines*. A third student consulted . . . *the Internet and some friends* for knowledge acquisition about her group's topic.

By contrast, oral presentation of topics like #18 and #19, for example, attempted to help students develop as skilled communicators, active learners, critical enquirers, investigators,

reflectors and analysts, problem solvers, knowledge evaluators, explorers, creators, and challenge embracers through collecting data in teams from living sources and presenting them before an audience. This is particularly important when knowing that Research Methods and the Final Project are two bilingual SQU and College elective courses respectively that are available for students to choose from during their study at the College. This way, students can develop and pursue “scholarship” that has positive implications for their future and Oman’s.

Assisting students to create and interpret knowledge in addition to acquiring it using the target language is one of the central aims of CLIL. Stoller (2004) discusses the Cognitive Academic Language Learning Approach (CALLA), which takes three components of academic literacy framework into consideration – Knowledge of the target language, content area, and tasks (LCT), which are structured independently and in interaction with each other. Students within CALLA pull their knowledge of the three domains to participate actively in the academic classroom (Stoller, 2004). This way, students develop their skills and build their knowledge around a certain process-oriented task and within a meaningful academic environment, which furnishes the way for more attachment to learning communities and interactive work (Nordmeyer, 2010).

Furthermore, oral presentations within this study aimed at stressing and encouraging collaborative and collective work. The nature of academic work today has changed and shifted more towards being collaborative and collaborative whereby learning is more interactive and integrative. Scholnik and Kol (1999) perceive pair and group work in oral presentations as positively influencing “social integration” and “negotiation of meaning”, and hence, directly and positively affecting the acquisition of language and knowledge and the quality of the overall outcome. One of the participants wrote that . . . *group work may be the most interesting thing, because you learn from each other. To divide the whole work it can make the search process much easier than doing it individually.* Another participant wrote that . . . *the support I received from my friends in the group made me feel that I can discuss any topic with confidence.* A third participant wrote that she . . . *liked students’ interaction and exchange of knowledge and that some students had background knowledge about the topic which made me feel comfortable presenting it.* The next participant looked at group work as integrating and combining interactive, social, intellectual, and educational aspects.

Different opinions always create better situations. We argued a lot and got mad and so on, but then we tried to take each other's advice and consider the small details to come up with brilliant ideas. It is my first time to deal with my classmates. However, we had a good time and shared good memories. It's always interesting to listen to other points of view.

One can claim that some of the participants learned from observing their fellow students, which reflected positively on their presentation techniques, content choices and decisions, and delivery plans (Morita, 2000). The audience here shared knowledge with the presenters, learned from them, or shared their different insights or interpretations, which are important aspects of cultural knowledge that the oral presentations involved (Morita, 2000).

Many students also tried to make their oral presentations “interesting, engaging, or even memorable” as a “significant aspect of discourse socialization” and a means of communicating their “epistemic stances” (Morita, 2000, p. 291) and demonstrate their “academic apprenticeship” (Morita, 2000, p. 296). This was evident in some of the presenters starting their presentations by asking the audience a question or two as a lead in to the topic.

Nevertheless, Item #19 had a comparatively lower mean than most of the other items in this cluster (3.91). This could be attributed to the amount of contribution some students made to the overall preparation of the topic. The groups were mixed-ability and perhaps the more capable students had a bigger share of the preparation than the less capable ones. Besides, there were complaints from one male group about not learning anything new from their group discussions since certain members put comparatively less effort and time into the project than the rest . . . *Most of them didn't work at all. They didn't attend the meetings and at the end they wanted to get the work and present.* A female group had problems of collaboration and imposition of certain individuals' opinions on the rest of the members . . . *We were in a group that was not cooperating and everyone wanted to impose her opinion on the rest of the members.*

Furthermore, while Item #12 had a high mean (4.00), it remains relatively low when compared to some of the items in this cluster. The underlying reason behind this could be the quality of questions asked and the quality of answers given, which hardly required any higher order thinking skills and analytical and reflective abilities. Some of the audience made basic and

general comments about some topics, rather than asked deep questions. This is while others asked yes/no questions, which lead the presenters to give very short answers.

Decision Making

Table 3
Participants Mean Ratings of Decision Making

| No. | Item | Mean | SD |
|-------------------|--|-------------|------------|
| 13 | Preparing for my presentation helped me think about what English language to use. | 4.23 | 1.18 |
| 14 | Preparing for my presentation helped me think about what legal English to use. | 4.41 | .66 |
| 15 | Preparing for my presentation helped me think about what legal knowledge to include. | 4.36 | .72 |
| Grand Mean | | 4.33 | .85 |

This cluster has the highest mean amongst all the seven clusters (4.33) Nevertheless, decisions about legal knowledge seem to be relatively easier to make than decisions about language. This appears in the mean obtained in Item #13 as compared to the other two items in this cluster.

One of the skills Arab students in general are accused of lacking is critical thinking, due to the nature of education in the region, which has been primarily memory, copying, transmission and delivery-based. Critical thinking is “. . . a very important skill for preparing successful citizens and students for a rapidly changing and challenging world” (Al-Issa, 2010, p. 169) and which involves “. . . asking good and effective questions and seeking knowledge in the age of information explosion, rapid technological change and complex social and political problems” (Al-Issa, 2010, p. 170). Critical thinking, according to Al-Issa (2010), is “. . . a central component for different activities incorporated within different undergraduate program majors including . . . topic presentation” (p. 175).

Nevertheless, decisions about legal English seemed easy to make, as it appears from the mean #14 has (4.41). This is perhaps due to the fact that legal English is more content and subject-specific.

One of the aims of CLIL is thus to help learners develop their intercultural awareness and global citizenship. Students within this context made decisions about what legal knowledge to pick and what to leave out from the wide array of theoretical and practical information available at their disposal. Knowledge adaptation was another aim that students had achieved. This was

evident in those who presented topics #1, #11, #15 and #20, for example, where they discussed the topic in general and then provided a critical account about what they had felt was culturally appropriate for the Omani context. The aim of oral presentations here has been to empower students to be in command of their minds and help them develop as active enquirers and “practicing” critical thinkers, who can gradually develop into “advanced” critical thinkers (Elder & Paul, 1996). One of the students wrote that she . . . *analysed the law itself and narrowed the field of search to the Omani law, because I wanted to know the mistakes in the Omani law before knowing the mistakes of other countries’ laws.* Another student explained how he made a decision about the content of his group’s presentation by saying . . . *I thought about including only what I felt was reasonable to the others and I committed myself to the part I chose for the presentation.*

Another aim of CLIL is to help students develop academic skills. A good number of the participants have already informally expressed their desire and ambitions to pursue their postgraduate education domestically or abroad. Reading for a postgraduate degree requires higher-order thinking skills as in the ability to make informed decisions about what knowledge to argue for and/or against, how and why. In other words, they need to know how knowledge is applied, analyzed, synthesized and evaluated. These are fundamental characteristics of critical thinking, which according to Al-Issa (2010), “. . . is important for academic success and success in life” (p. 172).

Presentation Strategies

Table 4
Participants Mean Ratings of Presentation Strategies

| No. | Item | Mean | SD |
|-------------------|--|-------------|------------|
| 20 | I memorized the content of my presentation. | 4.00 | 1.03 |
| 21 | I recited the content of my presentation. | 3.91 | .83 |
| 22 | I used notes and key words to present my work. | 4.00 | 1.08 |
| Grand Mean | | 3.97 | .98 |

As Table #4 shows, this cluster, which is associated with presentation strategies, has the lowest grand mean (3.97) amongst all seven clusters. Item #20 and Item #22 obtained an identical mean of 4.00, which is relatively high. Nonetheless, Item #21 obtained a relatively lower mean than the other two items (3.91).

Speaking is one of the skills Arab learners struggle to perform (Rabab’ah, 2005), and hence, feel shy or nervous to try to use the language inside the classroom and beyond, which can lead them to memorize. King (2002) is critical of memorizing word-by-word and reciting from different copied sources while presenting orally and sees it as dull. However, Al-Issa and Al-Qubtan (2010) argue that students descending from Asian and African cultures are used to memorization and are in favour of allowing students to memorize in a “good” way. “Good” memorization, according to Duong & Nguyen (2006), is a strategy that helps students learning EFL to improve their effectiveness through learning “. . . by heart with deep understanding and proper application in use for communication” (p. 14). Duong and Nguyen argue that “good” memorization can positively impact students’ confidence through using the target language correctly and effectively and speaking more fluently and accurately. This is particularly the case when considering the mixed-ability participants of this study, where the less capable ones had to exert more effort to construct coherent sentences that would convey the message clearly to the audience.

Oral presentations within the context of this study attempted to help students overcome as many of these language development barriers as possible. It was interesting to see many of the participants mixing and matching between two or more of the three aforementioned strategies, which justifies the somehow overlapping means quoted above.

One of the students compared her EFL (1) oral presentation experience with a previous one by writing that . . . *the previous presentation was all about memorizing, but in this new one I depended on understanding the topic and I gained much knowledge from the information I read.*

Content Description

Table 5
Participants Mean Ratings of Content Choice

| No. | Item | Mean | SD |
|-------------------|---|-------------|------------|
| 24 | The topic I presented was interesting. | 4.09 | 1.07 |
| 25 | The topic I presented was challenging. | 3.84 | 1.06 |
| 26 | The topic I presented added to my legal knowledge. | 4.55 | .63 |
| 27 | The topics presented by the other groups were interesting. | 4.02 | 1.11 |
| 28 | The topics presented by the other groups added to my legal knowledge. | 4.55 | .63 |
| Grand Mean | | 4.21 | .90 |

This cluster has a high grand mean of 4.21. Table #5 shows that Item #26 and Item #28 had the highest mean (4.55) in this cluster. This was complemented by the topics being interesting, as this appears from the mean obtained by Item #24 and Item #27 (4.09 and 4.02 respectively). Nonetheless, Items #25 obtained the lowest mean (3.84).

An important objective of CLIL is to strike the right balance between language and content (Gaballo, 2010). Students need to be directed and guided in this regard to be able to produce natural and innovative language. Students within the context of this study are freshmen with limited academic knowledge about Law and lower intermediate to upper intermediate level in English. Choice of topics that students were most familiar with was important for integrating the right content with language. One of the students described the topics as . . . *interesting . . . talking about laws and rules and regulations*. The topics, as discussed earlier, were based upon the mandated textbook and courses the students were studying then. However, they were adapted and manipulated to provide a different perspective, while meantime stimulate the students' interest and challenge their knowledge and skills. This appears evident in topics #2, #18, and #19, which are opinion survey-oriented and which were presented by some groups. Gaballo (2010) advises course planners, curriculum writers and teachers “. . . not to overwhelm students with excessive amounts of content that may lead to overlooking the language teaching and learning dimension of instruction” (p. 6).

Thus, Items #25, #24, and #27 had a relatively low mean (3.84, 4.02, and 4.09 respectively), when compared to the other two items in this cluster. The participants' limited knowledge about presentation techniques, Law, and English have probably put a limit to their discussion of the different topics they presented. Additionally, 30 marks were allocated for the oral presentations within the context of this study, which could have distracted the students' attention and turned it more towards thinking about their work than the other groups'. One of the students compared the oral presentations he did in the past with this one by saying that . . . *This had lots of marks, so we had to work hard for it. The presentations we had before were on easy subjects, but this had specific information*.

Presenting a technical topic in a foreign language is a complex, demanding, and challenging task that requires good knowledge, skills, and experience. Experienced presenters have the ability and skills to reach their audience and engage them in different ways and stimulate their interest through the different approaches, strategies, and techniques they adopt. Nonetheless, some of the

techniques that some students used while presenting their topics were reading the script/slides, using a small size of the font, speaking in a monotonous voice, overloading some slides, not establishing eye contact, and standing behind the front desk.

Some students thus considered their topics unchallenging. This was particularly the case with those who presented topics #5 and #11, where the former discusses a famous bank robbery, while the latter discusses the work of two famous lawyers. In addition, some students requested changing their topics as soon as they drew them out of the bag because they judged them as boring due to academic and cultural reasons like topics #1 and #10. While the former dealt with discussing a theory of Law, the latter was more about a non-existent legal system in the Islamic and Arab world and culture. These students' request was turned down, as each group was given one chance to draw one topic and prepare to present it. It is interesting to know that all four topics (#1, #5, #10, and #11) are of a theoretical-conceptual nature. They all discuss legal theories and concepts. One can argue that the excessive, strict, rigid, and boring theoretical teaching of Law at the College, as repeatedly voiced by several students at the College since its opening, has led the students to feel saturated with handling theoretical legal aspects. The excessive theoretical input and exposure at the College, as found by Al-Issa (2014), has been a dominant feature of its education and one which has been criticized for strongly affecting the students' attainment, attitudes, learning behaviour, and motivation, as it has favoured faculty staff's and textbooks' knowledge over other equally or more important knowledge, gave quantitative measurement education an edge over the qualitative mode, and gave product-based education priority and legitimacy over process-based (Al-Issa, 2014). This has negatively affected the students' image in the job market, reduced their chances of finding the right jobs, and disturbed Omanization (Al-Issa, 2014).

Conversely, many students thought that topics #2, #12, and #19 were particularly interesting, as they gave students an opportunity to communicate their epistemic stances (Morita, 2000) and demonstrate their sense of belonging to the legal community through the depth of knowledge and expertise they possessed about their potential professional field. In other words, such topics included examples, which helped students mainly and largely relate the topics to their schemas about the Omani context and nationwide trending issues of breaking the law and corruption, which are being discussed openly almost on a daily basis through the social media and other free communication technologies like WhatsApp, for example. Such topics allowed the students to

create space for their voices and show their discontent with the current legal *status quo* in the Sultanate. Interesting examples deriving from these three topics mainly centred round the Arab Spring or Arab Awakening or Arab Uprising, which witnessed its first spark in Tunisia in December 2010 and soon stretched to cover all Arab countries. The Arab nations have protested against a set of interrelated economic, political, and social issues such as the widespread poverty, increasing numbers of unemployment, almost entire absence of democracy, noticeably deteriorating level of human rights, and unfair distribution of wealth in the region. These are issues found in Oman too and the new generation is aware of their existence and their negative implications for nation building and citizenship development.

Personality Reinforcement

Table 6
Participants Mean Ratings of Personality Reinforcement

| No. | Item | Mean | SD |
|-------------------|---|-------------|------------|
| 29 | Presenting helped me play a more dynamic and interactive role inside the classroom. | 4.18 | .97 |
| 32 | Presenting in front of my classmates and teacher helped me gain courage. | 4.30 | .90 |
| 33 | Presenting in front of my classmates and teacher helped me gain self-esteem. | 4.21 | .88 |
| 34 | Presenting in front of my classmates and teacher helped me gain self-confidence. | 4.48 | .70 |
| Grand Mean | | 4.29 | .86 |

This cluster has a high mean of 4.29, which suggests the significant contribution oral presentations have made to the students’ personality reinforcement. The biggest advantage of presenting in front of one’s classmates has been developing self-confidence, as it obtained a mean of 4.48. In the second and third place came gaining courage and self esteem (4.30 and 4.21) respectively. While Item #29, obtained a relatively lower mean than the other three items in this cluster, it still obtained a relatively high mean of 4.18.

While men and women work side by side in all walks of life in a Muslim and Arabic culture like Oman, co-education is only found in a very small number of international private schools and at the college level. For the female participants to stand in front of class and be watched and heard by their male classmates is not a very usual scenario and one that does not happen on a daily basis. Such a scenario requires courage and self-confidence on the part of the female students. It was noticed that some of the female participants were nervous and shy during

presenting their topics. However, they were determined to accept the challenge and go through the experience, as it is a genuine preparation for the future.

Nevertheless, this should not give the impression that the male participants were all self-confident and courageous speakers. Many of the male participants across the Omani HE system are not used to standing in front of their female classmates to speak about a technical topic in a foreign tongue and perhaps do not feel comfortable about it for different reasons. Some of the least linguistically capable students in the class were thus males who could have felt anxious about making mistakes that would cause them some sort of embarrassment. This could be attributed to these students' personalities (Williams & Andrade, 2008), speaking in front of others (Williams & Andrade, 2008), lacking language fluency and accuracy (Morita, 2000; Williams & Andrade, 2008; Woodrow, 2006), performing in English in front of classmates and giving oral presentations (Horwitz, Horwitz, & Cope, 1986; Williams & Andrade, 2008; Woodrow, 2006), lacking critical thinking skills (Morita, 2000), answering questions about their presentations (Williams & Andrade, 2008), lacking experience in presenting orally (Williams & Andrade, 2008), and calling on students to present in random selection (Williams & Andrade, 2008). One of the participants wanted the teacher to . . . *Inform the groups about the date of presentation to be well-prepared.*

Nonetheless, one of the participants wrote that . . . *The oral presentations helped us to strengthen our personalities and we became confident about ourselves. We also have the skills of discussion by asking questions and engaging in answering.* Another student liked oral presentations because they gave him . . . *The opportunity to talk in a comfortable way and continuous way with self-confidence and without fear.* A third participant linked language and content knowledge acquisition and construction with self-confidence development and wrote that . . . *Presentations helped me to know more words and more legal information and have self-confidence.* A fourth participant referred to the implications of oral presentations to performance in other English-medium core courses. She wrote that . . . *Presenting was a good idea and it gave us self-confidence to present in other English courses like International Law.* The next student described the impact of oral presentations as an effective tool for achieving self-reflection leading to self-confidence development by saying that . . .

It always revives your self-confidence and takes you to a brand new experience with new people. It's hard and complicated, but its positive points exceed its negative points. You get to know yourself better and your weaknesses and to try to improve and take yourself to a higher level than before.

Confident and courageous speakers are also competent speakers, who know what to say, how, when, where and why. This is a central aim of CLIL, which pays equal and simultaneous attention to implicit foreign/second language acquisition, social interaction, creation of knowledge and academic skills (Xanthou, 2011) through fusing language education and subject education in order to achieve fluency (Gaballo, 2010).

Course Design

Table 7

Participants Mean Ratings of Course Design

| No. | Item | Mean | SD |
|-------------------|---|-------------|------------|
| 30 | Presenting added a positive change to the course. | 4.02 | .85 |
| 31 | I would like presentations to be a part of the overall assessment system of the other courses I am taking at the college. | 3.91 | 1.05 |
| 35 | The time given to prepare for the presentation was enough. | 4.50 | .95 |
| 36 | The time allocated to present our work was enough. | 4.43 | .106 |
| Grand Mean | | 4.21 | .97 |

This cluster seeks information pertinent to the role and place of oral presentations in the EFL One course. Items #35 and #36 obtained the highest mean (4.50 and 4.34 respectively) in this cluster. Item #30 obtained a relatively high mean (4.02). On the other hand, Item #31 obtained a relatively low mean (3.91).

One can argue that the reason behind this is possibly related to the theory and practice underlying education in the Arab World in general, which has always been textbook and exam-oriented and teacher-dominated and one which has considered students as empty vessels, and hence, excluded them from any decision making. Students in such a rigidly and strictly controlled and authoritarian system are spoon-fed certain types of knowledge by their teachers that are deemed to have an edge over all other types of knowledge, as they derive from the textbook and other specific traditions, and hence, serve the different interests of certain individuals (Al-Issa, 2010). This has subsequently affected the students' motivation, learning behaviour, attainment,

and attitudes regarding the nature of knowledge as a powerful and an infinite source, its renewability, perpetual construction and deconstruction, and openness to criticism, acceptance, and rejection.

Moreover, Pistorio (2009) reminds us that CLIL emphasizes language learning and not language teaching in order to help students receive the adequate amount of exposure to the target language (medium of communication) to achieve high levels of language proficiency and content mastery. Knowles (1975, 1980) argues that unlike children, adults find it important to know the reason(s) behind their learning. They further like to experience learning through steering and directing to discover new things. Knowles (1975, 1980) further stresses that it is important to delegate responsibility to adults and consider their voices through involving them in decisions relevant to planning, instruction and evaluation.

It is interesting to know that around four of the male participants objected to having an oral presentation and expressed their preference to take a test instead. This can be attributed to the fact that oral presentations require skills beyond memorization. In other words, they entail a cognitive challenge and require students to step out of their comfort zones and play multiple roles, which they are not used to playing and have not been trained to play at school and College of Law. Exams, to those participants, is an easier option and way to assess their knowledge, as they are less time and energy consuming and perhaps less stressful when compared to a four-week preparation for the sake of a 20-minute presentation.

Thus, this could have perhaps influenced the attitudes of some of the participants to express uncertainty about the oral presentations as helping them to play a dynamic and interactive role inside the classroom. Furthermore, about 36 of the participants did not involve the audience in their presentations due to lack of experience about oral presentations. However, they communicated “a sense of novelty” (Morita, 2000) to attract their audience’s attention. This took the shape of using certain powerful and evocative words, changing the standard oral presentations format, manipulating the original content of the presented topic, using support items such as videos and newspapers articles, and taking an overt stance, rather than a neutral one, on a debateable issue to engage the audience. This was evident in including pictures and videos that conveyed and emphasized their points of argument and using examples from real life to support their discussion, as it was the case with topic #2, #12, and #19, for example.

Presenting to peers, according to Morita (2000), often puts pressure on the students to stimulate or satisfy the audience intellectually. Peers here offered different opinions and perspectives, which posed challenge to the presenters. A few students made some intelligent post-presentation comments about some of the examples quoted in topic #7 and cases and data presented in topic #20, for example. A few others further asked some intelligent questions about the complex relationship between the five fields in topic #15, for instance, which required the presenters to think about sound elaborations to justify their earlier answers.

Knowledge in the case of the context in question, nonetheless, mostly went straight from the presenter to the audience. This is despite the fact that almost all topics presented could have easily had an interactive component integrated within them, which could have positively influenced the course of things for the presenters and the audience.

Research Question #4: What implications do the roles of EFL (1) faculty in the design and implementation of oral presentations have for the implementation of the new CLIL plan?

First, EFL (1) faculties at the College need to be aware of the “best practice” (Mehisto & Asser, 2007) that can benefit students in achieving high levels of language and content in an ICLHE context in specific and in a world largely controlled by economic interests and benefits and driven by knowledge, information, and skills. Choice of an adequate process-oriented task like oral presentations, for example, which have been found to promote humanistic, progressive, and liberal education and reflect positively on policy implementation, need to be a priority in the Omani EFL (1) and (2) College of Law classrooms in particular and the other English-medium legal courses in general to contribute to human capacity building.

Second, it was found that there are several aspects pertinent to the oral presentations that need to be addressed in order to help the students in the future make the most of the experience. Such aspects should reflect positively on the implementation and success of the new degree plan. Important issues like choice of presentation topics, distribution of groups, asking good and deep questions, embracing good memorization, reducing students’ anxiety, raising students’ awareness about some presentation techniques like moving around while presenting, avoiding overloading slides, establishing eye contact with the audience, avoiding speaking in a monotonous voice, and avoiding verbatim need to be addressed in the future.

Third, the new ICLHE plan, and as discussed above, aims at preparing graduates who are academically, cognitively, culturally, linguistically, psychologically, and socially adequate in order to cater for the needs of the local job market and facilitate Omanization, modernization, and nation building (Al-Issa, 2014). Attention to these aspects can help equip graduates with marketable skills, competence, and productivity level compatible with today's employers' aspirations, expectations, requirements, and demands in Oman and beyond. This is particularly the case in the age of globalization and internationalization, where English has become a focal point and a tool for acquiring and advancing different scientific and academic fields and disciplines like Law, for example. Perhaps one way of achieving this, as the experience has shown, is through giving students the opportunity to present more than once during the semester.

Conclusion and Recommendations

This paper investigated the perceptions of 44 EFL (1) course students about their oral presentations at an Omani ICLHE classroom. The data analysis and discussion have revealed that the participants found the experience relevant for their academic and professional lives. This is evident in the grand mean (4.15) the seven clusters have had. Such experience has its positive implications for contexts similar to the Omani one, other disciplines, and other educational levels.

The data analysis and discussion has revealed that oral presentations can play a significant role in policy implementation in the Omani ICLHE or other similar contexts worldwide. This is especially true as SQU – the only state-owned university in Oman, which has been making relentless effort for the past two decades or so to improve the quality of its academic programs and have them accredited internationally to help improve its world ranking.

This study had a number of limitations. First, the study examined students' perceptions about oral presentations through using a questionnaire. A future study can perhaps investigate the same topic through collecting data qualitatively to investigate language and content benefits. Furthermore, teachers' perceptions can be investigated about the effects of oral presentations, as the literature has stressed their role in impacting students' attainment and policy implementation. Second, the study was restricted to a small number of participants (n=44), which does not allow for generalization of results. Next, this study was restricted to reporting only one presentation, which reduces chances of understanding the contribution of the presentations to the students' language development. In other words, it is important to use oral presentations for a longer time

span to obtain more reliable results. Last but not least, this study focused on one EFL (1) section/class. Data from more than one section/class could be collected and analyzed in the future to enhance the validity of the outcomes, as variables like teacher's personality and certain strategic issues pertinent to the planning and execution of oral presentations could have influenced the outcome.

References

- Al-Issa, A. (2007). English language teaching at the College of Law—Muscat, Sultanate of Oman: Analyzing needs and understanding problems. *Asian Journal of English Language Teaching, 17*, 65–86.
- Al-Issa, A. (2010). Factors influencing critical thinking in the ELT context. In R. Al Mahrooqi & V. Tuzlukova, editors, *The Omani ELT symphony: Maintaining linguistic and socio-cultural equilibrium* (pp. 169-202). Sultan Qaboos University, Oman.
- Al-Issa, A. (2014). Researching the uses of English language in the law job market in the Sultanate of Oman: Implications for policy-practice improvement. *Educational Research for Policy and Practice, 13*(1), 25-44.
- Al-Issa, A. & Al-Qubtan, R. (2010). Taking the floor: Oral presentations in the TEFL classroom. *TESOL Journal 1*(2), 227-246. Retrieved from http://www.tesol.org/s_tesol/login.asp?TP=view.asp&Param=@CID=1997@DID=1035.
- Apple, M. & Kikuchi, K. (2007). Practical PowerPoint group projects for the EFL classroom. *The JALT CALL Journal, 3*(3), 110-122.
- Banegas, D. (2012). Integrating content and language in English language teaching in secondary education: Models, benefits, and challenges. *Studies in Second Language Learning and Teaching, 2*(1), 111-136.
- Coleman, J. (2006). English-medium teaching in European higher education. *Language Teaching 39*, 1–14.
- Costa, F. & D'Angelo, L. (2011). CLIL: A suit for all seasons. *Latin American Journal of Content & Language Integrated Learning, 4*(1), 1-13.
- Coyle, D. 2006. Content and language integrated learning motivating learners and teachers. *Scottish Languages Review, 13*. Retrieved from <http://www.scilt.stir.ac.uk/SLR/Issue%2013/SLR13%20Coyle.pdf>.
- de Graaff, R., Koopman, G., Anikina, Y., & Westhoff, G. (2007). An observation tool for effective L2 pedagogy in content and language integrated learning (CLIL). *The International Journal of Bilingual Education and Bilingualism, 10*(5), 603-624.
- de Graaff, R., Koopman, G., & Westhoff, G. (2007). Identifying Effective L2 Pedagogy in Content and Language Integrated Learning. *Vienna English Working Papers, 16*(3), 12-19.

- Duong, O. & Nguyen, H. (2006). Memorization in EFL students' strategies at university level in Vietnam. *TESL-EJ 10* (2). Retrieved from <http://teslel.org/ej38/a4.pdf>.
- Elder, L., & Paul, L. (1996). *Critical thinking development: A stage theory*. Retrieved from <http://www.criticalthinking.org/page.cfm?PageID=483&CategoryID=68>.
- Ellis, R. (2009). Corrective feedback and teacher development. *L2 Journal*, 1(1), 3-18.
- Eyjólfssdóttir, T. (2011). Content-based instruction: A closer look at CBI in Iceland. B.Ed. degree, School of Education.
- Gaballo, V. (2010, November). *Integrating content and language in specialized language teaching and learning with the help of ICT*. In ICT for Language Conference Proceedings. Milan: Simonellie Editor, Vol. 167, pp. 29-34.
- Horwitz, E., Horwitz, M., & Cope, J. (1986). Foreign language classroom anxiety. *Modern Language Journal*, 70, 125-132.
- Kasper, L. (1997). The impact of content-based instructional program on the academic progress of ESL students. *English for Specific Purposes*, 16(4), 309-320.
- King, J. (2002). Preparing EFL learners for oral presentations. The Internet TESL Journal VIII (3). Retrieved from <http://iteslj.org/Lessons/King-PublicSpeaking.html>.
- Knowles, M. (1975). *Self-directed learning: A guide for learners and teachers*. Chicago, IL: Follett Publishing Company.
- Knowles, M. (1980). *The modern practice of adult education: Andragogy versus pedagogy*. Englewood Cliffs, Cambridge: Prentice Hall.
- Lorenzo, F., Casal, S., & Moore, P. (2010). The effects of content and language integrated learning in European education: Key findings from the Andalusian bilingual sections evaluation project. *Applied Linguistics*, 31(3), 1-25.
- Marsh, D. (1994). *Bilingual Education and Content and Language Integrated Learning*. International Association for Cross-cultural Communication, Language Teaching in the Member States of the European Union (Lingua) University of Sorbonne. Paris.
- Martinez, A. (2013). Analysis of the effect of content familiarity and gender on English as a foreign language reading comprehension by Spanish university students. *Porta Linguarum* 21, 69-84.
- Mehisto, P. & Asser, H. (2007). Stakeholder perspectives: CLIL programme management in Estonia. *International Journal of Bilingual Education and Bilingualism*, 10(5), 683- 701.

- Miles, R. (2009). Oral presentations for English proficiency purposes. *Reflections on English Language Teaching*, 8(2), 103–110.
- Morgado, M. & Coelho, M. (2012). CLIL VS English as a medium of instruction: The Portuguese higher education polytechnic context. *Revista Egipcia Sciencia*, 12, 123-145.
- Morita, N. (2000). Discourse socialization through oral classroom activities in a TESL graduate program. *TESOL Quarterly*, 34(2), 279-310.
- Nordmeyer, J. (2010). *At the intersection of language and content*. Retrieved from <http://www.tesolmedia.com/docs/bookmail/660/Chap1.pdf>.
- Ortega, L. (1995). The effect of planning in L2Spanish oral narratives. Honolulu, HI: University of Hawaii Second Language Teaching and Curriculum Centre.
- Ortega, L. (1999). Planning and focus on form in L2 oral performance. *Studies in Second Language Acquisition*, 21, 109-148.
- Panova, I. & Lyster, R. (2002). Patterns of corrective feedback and uptake in the adult ESL classroom. *TESOL Quarterly*, 36(4), 573-595.
- Pinkley, D. (2012). *Children learning English as a foreign language*. Pearson Longman. Retrieved from <http://www.pearsonlongman.com/primaryplace/pdf/clilmonograph.pdf>
- Pistorio, M. (2009). Teacher training and competences for effective CLIL teaching in Argentina. *Latin American Journal of Content & Language Integrated Learning* 2(2), 37-43.
- Pragorbsuk, W. & Moore, I. (2002). The evaluation of EAP science new materials. *PASAA*, 33, 70-80.
- Radzuan, N. & Kaur, S. (2011). Technical oral presentations in English: Qualitative analysis of Malaysian engineering graduates' sources of anxiety. *Procedia - Social and Behavioral Sciences* 29 (2011) 1436 – 1445.
- Rabab'ah, G. (2005). Communication problems facing Arab learners of English. *Journal of Language and Learning*, 3(1), 180-197.
- Salimi, A. & Fatollahenejad, S. (2012). The effects of strategic planning and topic familiarity on Iranian intermediate EFL learners' written performance in TBLT. *Theory and Practice in Language Studies*, 2(11), 2308-2315.
- Scholnik, M. & Kol, S. (1999). *Using presentation software to enhance language learning*. The Internet TESL Journal V (3). Retrieved from <http://iteslj.org/Techniques?Scholnik-PresSoft.html>.

- Seikkula-Leino, J. (2007). CLIL learning: Achievement levels and affective factors. *Language and Education*, 21(4), 328-341.
- Sherris, A. (2008). *Integrated language and content instruction*. *CAL Digest*. Centre of Applied Linguistics. Retrieved from <http://www.cal.org/resources/digest/integratedcontent.html>.
- Siriphotchanakorn, C. (2005). Students' opinions toward oral presentations in EFL classrooms. Unpublished master dissertation. Srinakharinwirot University, Thailand.
- Soureshjani, K. & Ghanbari, H. (2011-12). Factors leading to an effective oral presentation in EFL classrooms. *The TFLTA Journal*, 3, 34-48.
- Stoller, F. (2004). Content-based instruction: Perspectives on curriculum planning. *Annual Review of Applied Linguistics*, 24, 261-283.
- Wannagat, U. (2007). Learning through L2: Content and language integrated learning (CLIL) and English as a medium of instruction (EMI). *International Journal of Bilingual Education and Bilingualism*, 10(5), 663-682.
- Wiesemes, R. (2005) *The Final Report for the Content and Language Integrated Project*. London: The Chartered Institute of Logistics and Transport.
- Williams, K. & Andrade, M. (2008). Foreign language learning anxiety in Japanese EFL university classes: Causes, coping and locus of control. *Electronic Journal of Foreign Language Teaching*, 5(2), 181-191.
- Woodrow, L. (2006). Anxiety and speaking English as a second language. *RELC Journal*, 37(3), 308-328.
- Xanthou, M. (2011). "The Impact of CLIL on L2 Vocabulary Development and Content Knowledge", *English Teaching: Practice and Critique*, 10(4): 116-126.
- Yuan, F. & Ellis, R. (2003). The effects of pre-task planning and on-line planning on fluency, complexity and accuracy in L2 monologic oral production. *Applied Linguistics*, 24(1), 1-27.

Appendix A

1. Pick a theory about Law. Describe it. Discuss its applications today. Tell us what you think about it.
2. Law makers are Law breakers! Discuss this statement. Support your discussion with real examples.
3. Search for & pick a famous murder court case/lawsuit. Describe it. Tell us what you think about it.
4. Search for & pick a famous tort law court case/lawsuit. Describe it. Tell us what you think about it.
5. Search for & pick a famous bank robbery court case/lawsuit. Describe it. Tell us what you think about it.
6. Search for and pick a famous intellectual property court case/lawsuit. Describe it. Tell us what you think about it.
7. Outline certain things that you think are forbidden in the law, but which you disagree with. Justify your point of view.
8. Discuss the principles and practices of the Islamic banks and say why they should/should not replace the other banks in Oman.
9. Find one law in the Holy Quran. Discuss its uses today. Compare it to its counterpart/equivalent in the human-made law.
10. Describe the “jury” system as implemented in the West. Give your opinion about it. Discuss its suitability for our legal & social system.
11. Find two famous lawyers. Discuss their work. Discuss why & how they became famous. Discuss how their work can inspire you in the future.
12. Many questions have been raised about human rights in the Arab World. Give some examples and comment on them in light of Islam and the developed world.
13. Describe and discuss Articles 5 and 6 of the Basic Law of State from an Islamic perspective and compare them to a few other contexts around the world.
14. Discuss a few aspects about how similar or different the Omani legal system is to the Islamic one. Quote examples where necessary to support your discussion.
15. Law, Politics, Economics, Arabic Language & English Language. How do they interrelate? How do you plan to improve your level at all those in the future? Why?

16. There are two major schools of Law – Latin (Civil) & Anglo-American (Common). Discuss what each one means. Which one is used at our College? Which one do you prefer, why?
17. Discuss the Omanization Law in light of the Article 50 of the Omani Labour Law. Conduct some interviews with a number of different Omani employees occupying different jobs and report the results.
18. Interview Dr. Adil A-Miqdadi in Arabic and Dr. Amal Abdullah in English. Ask each of them SEVEN intelligent questions about the Commercial Law. Report them to the class with a comparative commentary.
19. Think of four weird/strange/dumb law from around the world. Collect answers from 10 different people (students, faculties, administrators, family members, relatives, friends, etc.) about the suitability of implementing these laws for Oman. Report your findings to the class.
20. The Public Authority for Consumer Protection was recently established in Oman. Describe its functions & responsibilities. Compare it to its equivalent/counterparts in other countries. Report on some of the cases it has pursued. Give your opinion about some of the sentences imposed.

Appendix B

| Items | Gender | M | SD | T | Sig. |
|--------------|---------------|----------|-----------|----------|-------------|
| 1 | Male | 4.71 | .46 | .920 | .363 |
| | Female | 4.56 | .63 | | |
| 2 | Male | 4.29 | .81 | .612 | .544 |
| | Female | 4.13 | .89 | | |
| 3 | Male | 4.46 | .64 | .427 | .672 |
| | Female | 4.38 | .72 | | |
| 4 | Male | 4.04 | 1.04 | -.476 | .636 |
| | Female | 4.19 | .98 | | |
| 5 | Male | 4.11 | 1.17 | -1.215 | .231 |
| | Female | 4.50 | .73 | | |
| 6 | Male | 4.39 | .63 | 2.167 | .036 |
| | Female | 3.88 | .96 | | |
| 7 | Male | 4.04 | .88 | -.345 | .732 |
| | Female | 4.13 | .72 | | |
| 8 | Male | 3.93 | 1.05 | -.681 | .500 |
| | Female | 4.13 | .62 | | |
| 9 | Male | 4.32 | .67 | .912 | .367 |
| | Female | 4.13 | .72 | | |
| 10 | Male | 3.96 | 1.10 | -.534 | .596 |
| | Female | 4.13 | .62 | | |
| 11 | Male | 4.00 | 1.09 | .197 | .845 |
| | Female | 3.94 | .85 | | |
| 12 | Male | 3.96 | 1.10 | -.300 | .766 |
| | Female | 4.06 | .93 | | |
| 13 | Male | 4.11 | 1.31 | -.892 | .377 |
| | Female | 4.44 | .89 | | |
| 14 | Male | 4.32 | .72 | -1.174 | .247 |
| | Female | 4.56 | .51 | | |
| 15 | Male | 4.57 | .50 | 2.721 | .009 |
| | Female | 4.00 | .89 | | |
| 16 | Male | 3.82 | 1.02 | .027 | .979 |
| | Female | 3.81 | 1.11 | | |
| 17 | Male | 4.04 | 1.10 | .103 | .918 |
| | Female | 4.00 | 1.10 | | |
| 18 | Male | 3.86 | 1.04 | -.438 | .663 |
| | Female | 4.00 | 1.03 | | |
| 19 | Male | 4.11 | .74 | 1.914 | .062 |
| | Female | 3.56 | 1.15 | | |

| Items | Gender | M | SD | T | Sig. |
|--------------|---------------|-------------|------------|-------------|-------------|
| 20 | Male | 4.04 | 1.14 | .300 | .766 |
| | Female | 3.94 | .85 | | |
| 21 | Male | 3.86 | .93 | -.545 | .589 |
| | Female | 4.00 | .63 | | |
| 22 | Male | 4.21 | 1.07 | 1.788 | .081 |
| | Female | 3.63 | 1.02 | | |
| 23 | Male | 4.21 | .99 | .752 | .456 |
| | Female | 4.00 | .73 | | |
| 24 | Male | 4.18 | 1.09 | .712 | .480 |
| | Female | 3.94 | 1.06 | | |
| 25 | Male | 4.04 | .88 | 1.652 | .106 |
| | Female | 3.50 | 1.26 | | |
| 26 | Male | 4.50 | .64 | -.632 | .531 |
| | Female | 4.63 | .62 | | |
| 27 | Male | 4.14 | .85 | .949 | .348 |
| | Female | 3.81 | 1.47 | | |
| 28 | Male | 4.50 | .64 | -.632 | .531 |
| | Female | 4.63 | .62 | | |
| 29 | Male | 4.04 | .74 | .133 | .895 |
| | Female | 4.00 | 1.03 | | |
| 30 | Male | 4.07 | 1.09 | -.997 | .324 |
| | Female | 4.38 | .72 | | |
| 31 | Male | 3.96 | .96 | .456 | .651 |
| | Female | 3.81 | 1.22 | | |
| 32 | Male | 4.29 | .98 | -.093 | .926 |
| | Female | 4.31 | .79 | | |
| 33 | Male | 4.18 | .94 | -.257 | .799 |
| | Female | 4.25 | .77 | | |
| 34 | Male | 4.46 | .69 | -.161 | .873 |
| | Female | 4.50 | .73 | | |
| 35 | Male | 4.43 | 1.10 | -.654 | .517 |
| | Female | 4.63 | .62 | | |
| 36 | Male | 4.11 | 1.23 | -2.011 | .051 |
| | Female | 4.75 | .45 | | |
| Mean | Male | 4.16 | .87 | .072 | .943 |
| | Female | 4.14 | .79 | | |

Examining Prewriting Strategies in L2 Writing: Do They Really Work?

Anna Dina L. Joaquin

California State University, Northridge

Stephanie Hyeri Kim

California State University, Northridge

Sun-Young Shin

Indiana University

Bio data

Dr. Anna Dina L. Joaquin is Assistant Professor in the Department of Linguistics/TESL at California State University, Northridge. Her interests include first and second language acquisition, and TESL/TEFL pedagogy. She is the author of *Enculturation Processes in Primary Language Acquisition* (Equinox, 2013) and co-editor of *Exploring the Interactional Instinct* (Oxford, 2013). anna.joaquin@csun.edu

Dr. Stephanie Hyeri Kim is Assistant Professor in the Department of Linguistics/TESL at California State University, Northridge. Her research interests include conversation analysis and its applications to language teaching/learning, and TESL/TEFL pedagogy. She has published her articles in *Text & Talk*, *Journal of Pragmatics*, and *Research on Language and Social Interaction*. stephanie.kim@csun.edu

Dr. Sun-Young Shin is Assistant Professor of Second Language Studies at Indiana University. He received his Ph.D. in Applied Linguistics at UCLA. His research focuses on language assessment and program evaluations. His work has been published in several journals including *Language Testing*, *Language Teaching Research*, and *ReCALL*. shin36@indiana.edu

Address correspondence:
18111 Nordhoff Street,
Northridge, CA 91330

Abstract

Prewriting strategies are often taught to L2 learners to help them find a way into the writing topic and organize their ideas before writing an essay. However, little attention has been directed toward understanding whether prewriting and what type of prewriting actually benefits L2 writers. This paper examines the use and efficacy of prewriting strategies among 513 university level English language learners who have taken a timed essay placement exam. Our results show that more than half of the students chose to prewrite, and the majority of those that prewrote chose outlining, listing, or a combination of strategies in the L2. The students that prewrote outperformed those that did not, which confirms previous research findings that prewriting facilitates the writing process for L2 writers. However, in contrast to other studies, among those that prewrote, outlining and listing were not the more effective strategies, and there was no significant difference in test scores among those that used combined strategies. Another finding is that students who elaborated their prewriting scored higher than those who prewrote minimally or in a standard manner. This study provides ESL/EFL researchers and teachers with useful insights into understanding the use and efficacy of prewriting strategies in L2 writers.

Introduction

Prewriting is defined as “structured activities to provide motivation, content, fluency, and language” (Weigle, 2014, p. 227), which ultimately serves to help students find a way into the writing topic and facilitates production. Typical prewriting activities include pre-discussions, freewriting, drafting, outlining, listing, cubing, questioning, looping and webbing/clustering (Kroll, 2002; Spack, 1996; Weigle, 2014), and can be done silently, individually, in small groups, or as a class (Weigle, 2014). Such strategies are often promoted in textbooks for teacher-trainers and learners, as well as in ESL/EFL classrooms, where writing, particularly developing academic writing skills, is a core objective. Students and teachers are often encouraged to use such strategies because research suggests that they are an important part of the writer’s planning and can facilitate the writing process, and therefore contribute to a better product. However, much of the research on the efficacy of prewriting has been with writer’s writing in their first language (L1), and surprisingly few studies have examined the effects of prewriting in L2 writing. Though studies in L1 writing can inform L2 writing, as ESL or EFL composition research has pointed out similarities in the processes used by experienced L1 and L2 writers (Raimes, 1985), research has also shown that L2 writing is distinct from L1 writing (Ferris & Hedgcock, 2009). For example, learners’ knowledge and experience in writing in their L1 may facilitate or hamper writing in their L2. Thus, given the pedagogical push to use prewriting, it is necessary to explore the question: Does prewriting really work for L2 writers? Is it a facilitative strategy for ESL/EFL writers?

This paper explores this question by examining the use of prewriting strategies among 513 university level English language learners who have taken a timed essay placement exam. Specifically, we examine the relationship between the type of prewriting strategy employed (freewriting, outlining, listing, webbing, a combination of strategies, or no prewriting), the use of the L1 or L2, and the degree of elaborateness of the prewriting strategy with performance.

Though our data is from a US ESL context, such beliefs and practices have been influencing teaching practices in English as a foreign language (EFL) contexts more generally (Yang & Gao, 2013). Furthermore, most of the students who took the exam were entering international students who learned English in an EFL context, Asia in particular. Therefore, this study is relevant for understanding the efficacy of prewriting in both ESL and EFL contexts, and has pedagogical

import as it provides ESL/EFL researchers and teachers with useful insights into the understanding of the use and efficacy of prewriting strategies by L2 writers.

Literature Review

In L2 writing research, a dominant model is the Process Approach (Ferris & Hedgcock, 2009; Hyland, 2003; Weigle, 2014; Zamel, 1983). Such an approach emphasizes the individual student as an independent producer of texts, while the teacher supports the students through various stages of the writing process. Though there are many variations of the Process Approach, a common model is the planning-writing-reviewing framework (Flower, 1989; Flower & Hayes, 1981). This framework views writing as a “non-linear, exploratory, and generative process, whereby writers discover and reformulate their ideas” (Zamel, 1983, p. 165) throughout the process. Though composing, revising, and proofreading may be recursive and even simultaneous, the overall framework begins with the selection of a topic, followed by an essential prewriting stage.

Though prewriting is often promoted as an essential component of the writing process, theoretical perspectives have not always been certain of prewriting’s facilitative effects. In L1 research, from which L2 writing processes are often derived, the Interaction Hypothesis (Elbow, 1981) suggested that planning strategies, such as outlining, deteriorated text quality, as it may be too restrictive and may prevent writers from making use of opportunities that arise during the interaction of planning, translating and reviewing. Therefore, proponents of the Interaction Hypothesis promoted freewriting instead, which is to write immediately and without restriction on form and content upon reading a task prompt. In contrast, the Overload Hypothesis considered attentional resources and proposed that prewriting activities, including both mental and written outlining, allowed writers to focus on translating ideas into writing (Kellogg, 1988). Thus, planning improves text quality because planning reduces cognitive demands placed on writers “by freeing spaces in their limited working memory during the composing process” (Ong, 2014, p. 19).

Indeed, several studies have supported the Overload Hypothesis. In one study, Gillis and Olson (1990) examined the difference in the essay exam scores of 100 English-speaking undergraduate and 37 graduate students, who either did no planning, some planning, or extensive

planning before writing essays. The results indicated that students who did some planning scored better than students who did no planning, and students who did extensive planning scored better than students who did no or some planning. Another widely referenced study by Kellogg (1990) examined the role of prewriting among 207 college students. The students wrote a short informative essay within 30 minutes and were randomly assigned into four conditions: a control group that did not receive instructions to prewrite; a group that was asked to 'list' ideas in the order they should appear in the text; a group that was asked to 'cluster' ideas; and a group that was told to create a structured 'outline'. The results demonstrated that there was no difference between the essays produced after making a cluster and those with no planning at all. However, the students that listed performed significantly better, while the students that outlined performed the best overall. Thus, Kellogg's study suggests that not only does prewriting improve the final product, but that some prewriting techniques (particularly listing and outlining) may be more effective.

Another study suggests that the medium of prewriting has an effect on the process toward the final product (Rau & Sebrechts, 1996). Researchers asked participants to plan either silently or through thinking aloud. In the silent condition, participants were told to think silently about their composition for five minutes, while those in the silent outline condition were told to write an outline for the same amount of time. The think- aloud participants were divided into two groups: think aloud while planning and think aloud while outlining. Both groups were given five minutes. The prewriting activities were followed by approximately 30 minutes of writing the final product. The results showed that participants who silently planned and thought aloud without a written outline made more content revisions and had more pauses during the writing process, although there was no significant difference in the final text quality (Rau & Sebrechts, 1996). Such a study demonstrates a relationship between pre-writing and real-time writing behavior, especially for increasing fluency.

Prewriting, specifically outlining, in electronic form has also been demonstrated to improve the quality of final products and decrease mental effort (Smet, Broekkamp, Gruwel, & Kirschner, 2011). Other studies demonstrate that prewriting strategies result in more creative and sophisticated compositions, and are judged to be of higher quality (Perl, 1979; 1996). Thus, research on planning among L1 writers suggests a facilitative effect on the writing process.

While research concerning the L1 writing process suggests that prewriting has a positive effect on the final product, little is known about the effects of prewriting in L2 writing, though such strategies are often promoted in textbooks for teacher-trainers and learners. Among L2 writers, two theories have been relevant: Skehan's Limited Attentional Capacity Model (Skehan, 1998) and Robinson's Cognition Hypothesis (Robinson, 2005). Though different in their explanation, both suggest that planning decreases the complexity and attentional demands of a task and therefore should positively impact the writing to be more fluent and complex. While some researchers have suggested that planning has no effect (e.g., Johnson, Mercado, & Acevedo, 2012), several studies have shown that prewriting activities do in fact improve the final product in L2. Shi (1998) has found that prewriting discussions facilitated the writing process among adult ESL students. In comparison to students who did not engage in prewriting discussions, students that had pre-discussions wrote shorter drafts with a greater variety of verbs. In addition, Ellis and Yuan (2004) found that written pre-task planning, in comparison to no planning, resulted in greater accuracy, fluency, and syntactical variety. Mohseniasl (2014) demonstrated that prewriting activities, namely brainstorming, concept mapping, and freewriting, alleviate writing apprehension and optimize writing achievement. Ong and Zhang (2013) specifically found that among Chinese EFL learners, freewriting, compared to other planning conditions (i.e. prolonged planning), "enhanced the quality of the learner's writing" (p. 375). Such studies seem to support Skehan (1998) and Robinson's (2005) hypotheses that planning before writing alleviates attentional and cognitive demands, which improves L2 writing. Indeed, learners may also have the perception that prewriting is facilitative. As one learner from a qualitative study on the perceptions of the usefulness of pre-writing reports:

Good planning is the cornerstone of successful writing. An unplanned piece of writing is like an unplanned life; it is often disorganized, ununified, and incoherent. Such writing can be flooded by many mistakes as well as faults in reasoning, in presentation of material, and in grammatical structure (Adegbija, 1991, p. 230).

However, Hyland (2008) has noted that though some L2 learners acknowledge that planning may be beneficial, they also feel that it may be counterproductive. For example, one Japanese learner of English stated, "I know that my writing is quite bad. Even I write in Japanese is still does not

make sense sometime. I think because I don't make a plan. When I try to make a plan my ideas disappear" (Hyland, 1998, p. 275).

This study aims to fill a gap in understanding the use and efficacy of prewriting strategies by L2 writers. Though studies suggest, and learners perceive, that prewriting strategies are effective, most studies have been done in the L1 and less attention has been directed toward understanding whether such prewriting actually benefits L2 writers. In addition, previous studies on L2 have often been in experimental conditions in which participants were asked or prompted to prewrite and to use a particular strategy (e.g., Kellogg, 1990; Rau & Sebrechts, 1996). In this study, L2 writers are responding to a prompt in a timed exam and are not given any explicit instructions as to whether they should prewrite or not. Thus, in the current study, the prewriting that occurs is self-initiated. Furthermore, to our knowledge, none of the previous L2 writing studies have done a comparison of the specific strategies and their effects on the final product with L2 learners. Specifically, this study compares the performance of learners who chose to outline, freewrite, list, web, or not prewrite at all. In addition, none of the previous studies examined other characteristics such as the language used (L1 or L2) and degree of elaborateness of the planning in a timed placement exam. Thus, this study aimed to examine the relationship between the use of specific prewriting strategies and their characteristics with the final product of L2 learners.

Data and Methods

Participants

Our data are drawn from the English as a Second Language placement examination (ESLPE) administered at a major four-year research university in the U.S. The exam is required of all entering graduate students and selected transferring undergraduate students whose first language is not English, except for students holding a bachelor's or higher degree from a university in a country where English is both the language of instruction and the primary language of daily life.¹ Most students taking the exam are recently-arrived students and in their first year of studies at the university. While it should be noted that neither the Test of English as a Foreign Language (TOEFL) nor any other English proficiency test can replace the English placement exam, the university requires its applicants to have a minimum IELTS score of 7.0 and a minimum TOEFL

¹ For reasons that cannot be explained here, international students entering as freshmen are given an exam that is not specifically tailored to students whose first language is not English.

IBT score of 83 for undergraduate students and 87 for graduate students. Some departments may require higher scores of their graduate applicants. The exam is a high-stakes test because based on their performance, students can be required to take one or several English as a Second Language courses (up to 2 for graduate students and 3 for undergraduate students) during their studies at the university, which could delay their time to degree.

Research Questions

Given the gap in previous research, we asked the following four questions aimed at understanding prewriting behavior in L2 writing and its effects on student performance.

1. *Do students prewrite during a timed essay exam?*
 - a. *If so, which strategies do they employ and to what degree?*
 - b. *Do they use their L1 or L2?*
2. *Do students who prewrite produce better essays than those who do not prewrite?*
 - a. *Do better performing students use a specific type of prewriting strategy?*
3. *Does the language used, L1 or L2, during the prewriting activity affect students' performance?*
4. *Does the degree of prewriting elaborateness affect students' performance?*

Data

Our data comes from 513 written essays administered throughout the 2012-2013 academic year. Students are given one hour to complete the examination, which only consists of a written component. Students are given a choice to respond to one of two provided prompts in an essay format (see Appendix A). Having the ability to choose the topic they prefer reduces topic effect on performance. Students are asked to bring their own pens or pencils and are given 5 pieces of paper in total; instructions are provided on the first page and two prompt choices on the last page, and three lined pages (with blank space on the back) are provided in the middle of the packet for writing the essay. All exams are hand-written. The instructions given to students on the first page of the test are:

You will have 60 minutes to plan, write, and revise a formal academic composition on one of the two topics on the last page. Choose only one of the topics on page 5 for your composition. Use examples from your own experience or from any reading you have done to support and explain your ideas. Your composition will be evaluated on linguistic control (grammar, expression, and word choice) and your ability to express your ideas clearly and accurately in academic English. You may use the back of this page for making notes and planning your answer and the lined pages for your composition.

It is worth noting that students are given the option to use the back of the first page to make notes and plan (i.e. prewrite). Indeed, students who prewrote made notes on the back of the first page, and also on the last page where the prompts are given. See Appendix A for the two essay prompts used during the 2013 exam administration.

Raters were full-time lecturers who teach ESL courses at the university. Each test was scored by two raters based on the rubric found in Appendix B. They were all trained and calibrated for essay rating, and inter-rater reliability for the data was also high ($\alpha=.90$). The average of the two ratings was used as the final score.

Data Analysis

The averaged essay scores obtained from two raters were used as a dependent variable, along with the following variables including different prewriting strategies and combinations, the degree of elaborateness of prewriting, and the type of languages they used for prewriting were systematically coded and used as independent variables. These data were analyzed for independent samples t-test and an Analysis of Variance (ANOVA) for mean comparisons of final essay scores across different independent variables using SPSS 21 (2012).

Coding and Definitions

In order to develop a coding scheme for the prewriting strategies, we examined a random set of tests and decided collectively how each test would be coded. Then, the tests were divided so that each author had the same number of tests to code. While coding was done separately, whenever

any uncertainty arose as to how a certain test should be coded, the authors consulted each other to achieve consistency.

We first coded each prewriting sample for the language used: L1, L2, or mixed. Then prewriting strategies were categorized into seven types based on demonstrated student behavior on the exam.² The types are: drafting, freewriting, outlining, listing, webbing/clustering, no prewriting, and other (see Table 1 and Appendix C for examples).

We define ‘drafting’ as continuous prose in which the writer has transferred much of what they have written to the final product. The draft need not be a full draft, but the portions drafted must be similar in wording and structure to the final product. ‘Freewriting’ is also continuous prose; however, the writer has not transferred much of what they have written to the final product. The ideas may appear in the final product, but the freewriting is significantly different in wording and structure from the final product.

We define an ‘outline’ as an organized list of main ideas in which subordination often occurs, representing the supporting examples or ideas that the writer may have intended to follow for the final piece of writing. Walvoord, Anderson, Breihan, McCarthy, Robison, and Sherman (1995) defined an outline as a written, vertical list of ideas or information in the sequence that the writer intends for the final piece of writing. While this may be a traditional perspective of an outline, we found that participants’ outlines were often not vertical, but did correspond to some degree to the structure of the final product.

We have generally followed the definitions for listing and webbing from Hyland (2003, p. 130). ‘Listing’ is defined as the listing of details for an essay topic (people, place, actions, feelings, objects, etc.). ‘Webbing’ or ‘clustering’ is a pattern of ideas that are joined by lines showing connections between them. The forms of prewriting that were unique to the test taker or that could not be categorized under the present scheme were categorized as “other”.

² If the prewriting was written in a language unfamiliar to us, we consulted native speakers of the language who also had extensive experience teaching ESL and were familiar with such strategies.

Table 1

Summary of prewriting strategies and definitions

| Types | Definitions |
|--------------------|--|
| Drafting | Continuous prose; writer has transferred near exact wording and structure to the final product |
| Freewriting | Continuous prose; wording and structure differ significantly from the final product |
| Outlining | Organized list of main ideas, in which subordination occurs representing supporting examples or ideas that the writer may intend to follow for the final product |
| Listing | Listing of details for an essay topic (people, place, actions, feelings, objects, etc.) |
| Webbing/Clustering | A pattern of ideas that are joined by lines showing connections between them |

In addition, participants varied in the extent to which they prewrote. For example, some lists were longer and more elaborate than others; some wrote words while others wrote phrases or sentences. For each prewriting sample, the degree of detail was categorized as minimal, standard, or elaborate (Benton & Blohm, 1988; Chai, 2006). If the prewriting was identifiable as a particular type of strategy, but only had words, or seemed incomplete, it was considered “minimal”. If the prewriting was identifiable as a type and seemed complete in that it contained most of the main characteristics of the strategy, it was categorized as “standard”. For example, a standard outline would have a list of the main ideas with supporting examples. If the prewriting had more than the main characteristics and had extensively developed ideas and included sentences, then it was considered “elaborate”. An outline with a developed introduction and conclusion, as well as the standard topic ideas and supporting topics, would be categorized as “elaborate” (see Appendix D for examples). As to the webbing/clustering strategy, it was considered “minimal” if only 1-2 words or phrases were connected to the key word(s). A “standard” webbing would exhibit more associated words linked to the key words, typically 3-4 words joined together with lines or arrows. The prewriting that went beyond this standard type, typically with a more complex connection of words/phrases that were developed in “all directions”, was considered “elaborate” (See Appendix E for examples).

Results

Our results section is organized according to the research questions, with the parts of each question answered in order.

Research question 1: *Do students prewrite during a timed essay exam?*

a. If so, which strategies do they employ and to what degree?

b. Do they use their L1 or L2?

Results showed that 57% of students ($n=291$) prewrote, and 43% of them ($n=222$) did not prewrite during the timed essay exam. We identified six different prewriting strategies: *Outlining* (52%, $n =152$), *Listing* (29%, $n =83$), *Drafting* (5%, $n =14$), *Webbing* (3%, $n =10$), *Freewriting* (2%, $n =6$), and other unspecified prewriting types (9%, $n =26$). The majority of students (80%, $n =232$) used a single prewriting type, whereas 20% of them ($n =59$) used more than one type in their prewriting. Four of them used three different styles of prewriting. See Table 2 for more information about various mixed prewriting types used by the students. It is important to note that such prewriting strategies are not mutually exclusive. Indeed, some participants combined strategies. Other forms of prewriting were also observed, such as organizing positive and negative supporting ideas into a quadrant form or writing only few words (sometimes illegible words) on the margins. As can be seen in Table 2, the most common combined prewriting types were outlining and listing (36%) and *webbing* was the most common additional prewriting strategy type. It should also be noted that the majority of the students (87%, $n =253$) used English, and some students used both English and their L1s (10%, $n =28$), and a small number of students used only their L1 for prewriting (3%, $n =10$).

Table 2

Combinations of prewriting strategies

| | Frequency | Percent |
|---------------------------------|-----------|---------|
| Listing & Outlining | 21 | 36% |
| Listing & Freewriting | 1 | 2% |
| Listing & Drafting | 4 | 7% |
| Listing & Other | 3 | 5% |
| Listing & Webbing | 6 | 10% |
| Outline & Webbing | 12 | 20% |
| Outline & Draft | 8 | 14% |
| Listing, Outline & Webbing | 2 | 3% |
| Outline, Draft & Webbing | 1 | 2% |
| Webbing, Listing, & Freewriting | 1 | 2% |
| Total | 59 | 100% |

Research question 2: *Do students who prewrite produce better essays than those who do not prewrite?*

a. Do better performing students use a specific type of prewriting strategy?

Table 3 below shows the descriptive statistics (i.e., the means [*M*], standard deviation [*SD*], Skewness, and Kurtosis) for the timed essay exam scores, which were normally distributed.

Table 3

Descriptive statistics for the timed essay test scores

| | |
|-----------------|-------|
| <i>N</i> | 513 |
| <i>M</i> | 4.75 |
| <i>SD</i> | 0.84 |
| <i>Skewness</i> | -0.11 |
| <i>Kurtosis</i> | -0.70 |

Table 4 below reports Independent Samples T-test results, showing that students with prewriting activities ($n = 291$, $M = 4.84$) outperformed those who did not use them ($n = 222$, $M = 4.63$), $t(511) = 2.87$, $p < .00$ on the timed essay exam.

Table 4

Independent Samples T-test for comparing means of test scores between prewriters and non-prewriters

| <i>t</i> | <i>df</i> | <i>p</i> | Mean difference | Standard error difference | 95% confidence interval of the difference | |
|----------|-----------|----------|-----------------|---------------------------|---|----------------|
| 2.87 | 511 | .004 | .22 | .07 | .07 (Lower) | .36 (Upper) |

Nevertheless, as can be seen in Figure 1, there was a minimal overall score difference among the different prewriting strategies. A one-way ANOVA was conducted to see if better performing students used a specific type of prewriting strategy. ANOVA results, as in Table 5, reveal that any difference in essay exam scores among different prewriting types was not statistically significant, $F(5, 285) = 1.42, p = .22$.

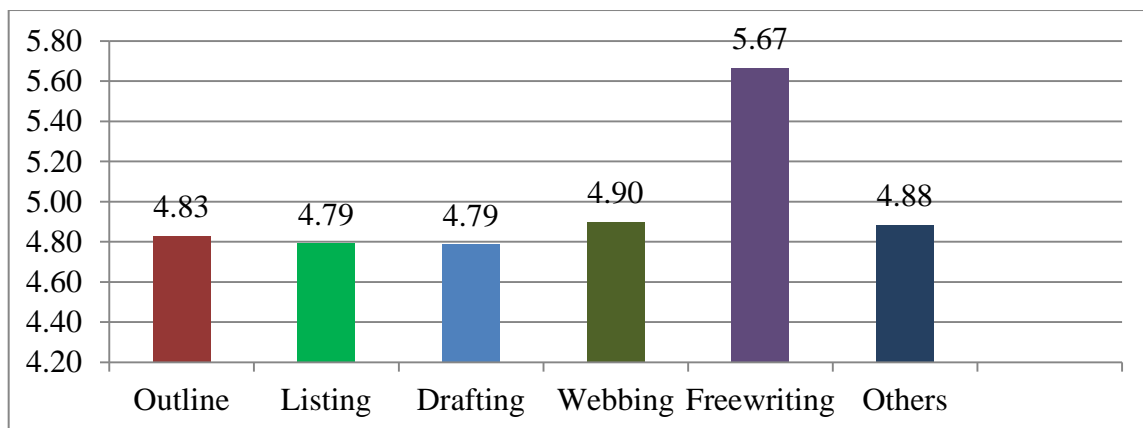


Figure 1. Mean comparisons among different prewriting strategies

Table 5

One-Way ANOVA comparing means of test scores obtained from each different prewriting type

| Source | <i>SS</i> | <i>df</i> | <i>MS</i> | <i>F</i> | <i>p</i> |
|------------------|-----------|-----------|-----------|----------|----------|
| Prewriting types | 4.45 | 5 | .89 | 1.42 | .22 |
| Error | 179.32 | 285 | .63 | | |

As can be seen in Table 2, some students mixed different types of prewriting strategies; however, Figure 2 below shows that students who used a single prewriting type performed very similarly to those who combined strategies in their prewritings. Independent samples t-test was conducted to examine if there was any statistically significant mean difference between two groups. As shown

in Table 6 below, the mean difference was not significant, $t(289) = .10, p = .92$. As addressed above in Table 2, there were 10 different combinations of prewriting types, but a one-way ANOVA, displayed in Table 7, reveals that no meaningful score difference was found among them, $F(9, 49) = 1.43, p = .20$.

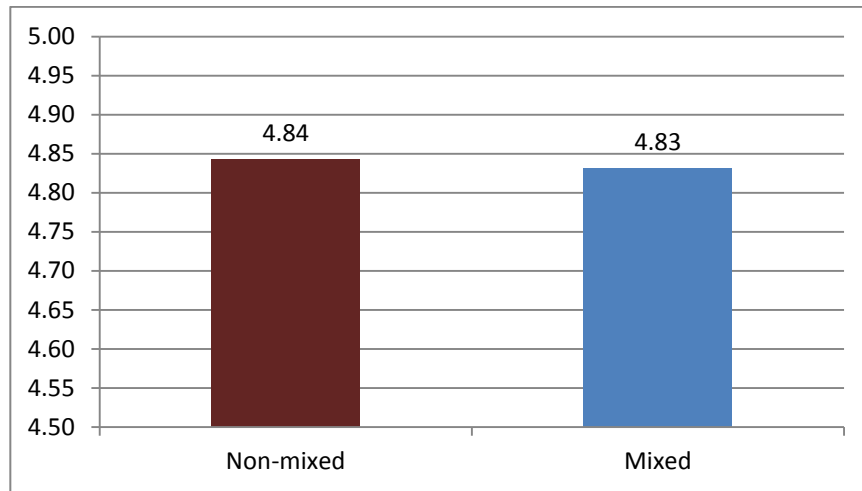


Figure 2. Mean comparison between single and mixed prewriting groups

Table 6

Independent Samples T-test for comparing score means between students who used a single prewriting strategy and those who mixed strategies

| <i>t</i> | <i>df</i> | <i>p</i> | Mean difference | Standard error difference | 95% confidence interval of the difference | |
|----------|-----------|----------|-----------------|---------------------------|---|---------|
| .10 | 289 | .92 | .012 | .12 | -.22 | .24 |
| | | | | | (Lower) | (Upper) |

Table 7

A one-way ANOVA for comparing score means of 10 different combinations of prewriting types

| Source | <i>SS</i> | <i>df</i> | <i>MS</i> | <i>F</i> | <i>p</i> |
|-------------|-----------|-----------|-----------|----------|----------|
| Mixed types | 8.39 | 9 | .93 | 1.43 | .20 |
| Error | 31.91 | 49 | .65 | | |

Research question 3: *Does the language used during the prewriting activity affect students' performance?*

As can be seen from Figure 3, there was a small mean difference among students who used different languages for their prewriting, although students who used both their L1s and English scored slightly lower than the other two groups. A one-way ANOVA was conducted, as reported in Table 8 below, to examine if there are any significant mean differences in their writing scores among students who used their L1s, English, or both. The results show that the language they used for prewriting did not lead to a statistical difference in score on the timed essay exam, $F(2, 288) = 2.02, p = .13$.

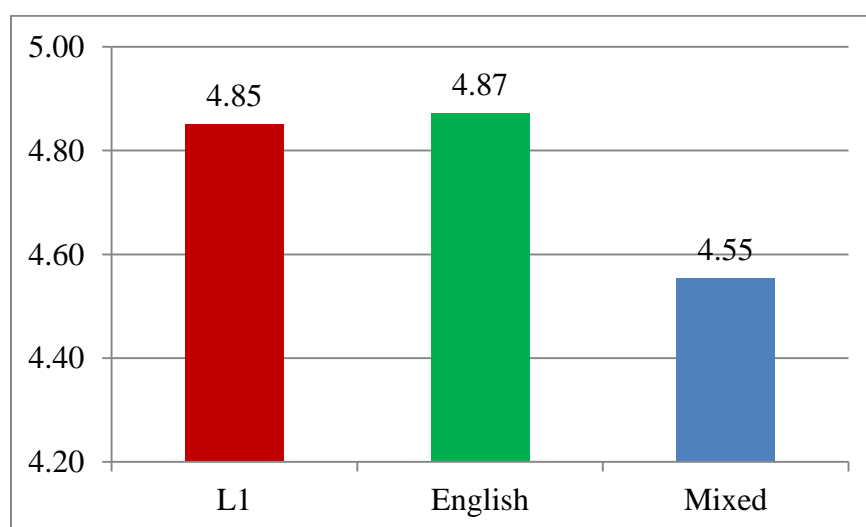


Figure 3. Mean comparisons among groups who used different languages for prewriting

Table 8

A one-way ANOVA for comparing score means of languages used in prewriting activity

| Source | SS | df | MS | F | p |
|-----------|--------|-----|------|------|-----|
| Languages | 2.54 | 2 | 1.27 | 2.02 | .13 |
| Error | 181.23 | 288 | .63 | | |

Research question 4: *Does the degree of prewriting elaborateness affect students' performance?*

With regard to the degree of development in prewriting, 36% of students ($n = 105$) wrote minimally, 41% of them ($n = 119$) in a standard manner, and 23% of them ($n = 67$) elaborated in their prewriting. A one-way ANOVA was carried out to see if there were statistically significant mean differences among the students who used the different degrees of prewriting elaborateness.

As can be seen in Table 9 below, the ANOVA results reveal that the degree to which students elaborated on their prewriting significantly affected their essay scores, $F(2, 288) = 6.50, p < .002$. Figure 4 below shows that students who elaborated on their prewritings scored higher than the other two groups who prewrote minimally or in a standard manner. Tukey's Post hoc comparisons further found that there was a statistical difference in essay exam scores between the *elaborate* and *minimal* groups (mean difference = .44, $p < .00$), whereas the mean difference between the *standard* and *elaborate* groups was not statistically significant but was very close to being statistically significant (mean difference = .28, $p = .06$).

Table 9

A one-way ANOVA for comparing score means of writings used the different degrees of prewriting elaborateness

| Source | <i>SS</i> | <i>df</i> | <i>MS</i> | <i>F</i> | <i>p</i> |
|---------------|-----------|-----------|-----------|----------|----------|
| Elaborateness | 7.94 | 2 | 3.97 | 6.50 | .002 |
| Error | 175.83 | 288 | .61 | | |

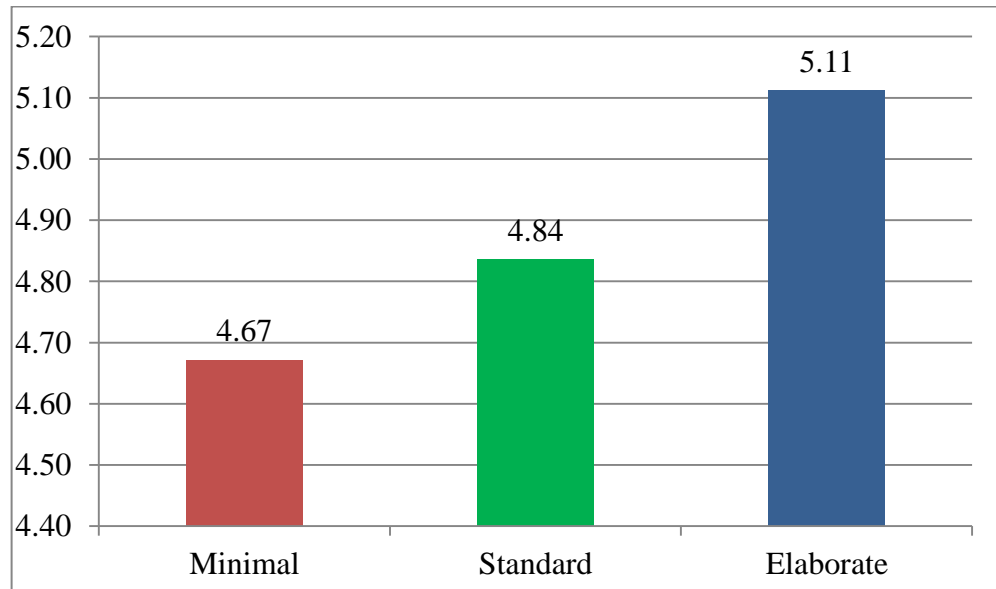


Figure 4. Mean comparisons among the different degrees of elaborateness of prewriting.

Discussion

Based on the results of previous L1 and L2 writing research, we hypothesized that those who performed well would be more likely to employ a prewriting strategy. In addition, as previous research noted that outlining is the most effective strategy, we expected that students who used outlining would outperform those who used listing, webbing and other strategies. Research has also shown that though L2 writers still depend heavily on their L1 in some way, more advanced learners tend to rely less on their L1 during their writing process than the less advanced (Cumming, 1989; Jones & Tetroe, 1987; Raimes, 1985). The assumption is that advanced learners have a better grasp of the language and therefore can prewrite with more confidence in their L2. For example, Van Weijen (2009) found that L1 use during L2 writing is negatively correlated to L2 text quality. Logically, we hypothesized that the students who prewrote in their L1, or not predominantly in their L2, are more likely to have lower scores. Finally, given the time constraints, we posited that students would vary in the degree of their prewriting, but few would be extensively developed.

Our results show that more than half of the students chose to prewrite and a majority of the students chose to write using a single prewriting type, but many also mixed prewriting strategies. Furthermore, the majority of those that prewrote chose either outlining or listing in the L2. The students that prewrote outperformed those that did not, which confirms previous research findings (Ellis & Yuan, 2004; Gillis & Olson, 1990; Kellogg, 1990; Ong, 2014; Ong & Zhang, 2013; Rau & Sebrechts, 1996) that prewriting facilitates the writing process for L2 writers. However, in contrast to other studies, among those that prewrote, outlining and listing were not the more effective strategies. There was no significant difference in test scores among the essays which utilized varying prewriting strategies.

We also found it interesting that many students combined prewriting strategies. Some of the previous research had categorized participants into groups that were asked to employ a particular strategy or not to employ a strategy. For example, participants were asked to outline or not to outline. Students in our data, however, were not told to prewrite. Instead, as noted earlier, they were given the option to use the back of the first page to make notes and plan for their essay. With such instructions, we were able to observe their individual preferences - prewriting or no prewriting, language and strategy preferences. Moreover, there was not a significant difference in

student performance between students that used a single strategy and those that combined. Such results are a reminder for English language teachers, who have a tendency to emphasize a particular strategy or redirect learners' planning preferences, to consider individual differences (Dornyei, 2005; Kormos, 2012; Manchon, 2009; Schumann, 2004), and to expose students to a variety of prewriting strategies. Teachers should cultivate the use of prewriting, but perhaps not constrain a student to employ a specific strategy.

Though we hypothesized that the students who prewrote in their L1 or not only in their L2 would be more likely to have lower scores, as writing in the L1 would be an indication of their lack of the grasp of the L2, our results suggest that the students' use of their L1, L2, or mixed did not lead to a statistically significant difference in scores. Therefore, prewriting in the L1, L2 or mixed is not necessarily an indicator of the proficiency or ability of the L2 writer, at least among the university-level students. Indeed, research results have been conflicting in this aspect. Lally (2000) found that prewriting tasks done in the L1 slightly improved vocabulary in comparison to those that prewrote in their L2, while Lay (1992) also found that students who used both their L1 and L2 during the prewriting and composing stages produced essays that were better organized with more complex ideas and details. Our study adds to the increasing body of literature that explores the effects of prewriting in the L1, L2 or both on the final product. Although prewriting in the L2 may be good for practicing writing in the target language, the implication of our study for teachers is that students should be allowed to prewrite in the language of their choice if the end goal is the production of a good essay or final product, and not to practice the L2 via the prewriting strategy.

Another finding is that students who elaborated their prewriting scored higher than those who prewrote minimally or in a standard manner. Therefore according to the results of our study, the two factors that seem to facilitate performance are not only employing a prewriting strategy, but the degree of development of the prewriting. This finding seems to support Skehan's (1998) Limited Attentional Capacity Model and Robinson's (2005) Cognition Hypothesis, as prewriting decreases the cognitive demands while writing the final essay because ideas, details, examples, key phrases, keywords etc. have been thought of a priori. Therefore, presumably, the more elaborate the prewriting, the more the student has planned the essay, which frees cognitive resources and facilitates an otherwise overwhelming and complicated task. In addition, both theories suggest that planning can lead to writing that is more fluent and complex. Though we did

not do a formal analysis or comparison of the complexity of the writing, one component that raters are trained to pay attention to is the complexity of the language. For example, in addition to attending to grammar, punctuation, and meaning, raters score the essays according to students' use of subordination, embedded clauses, repetition of key words, transitional words and phrases, and sophistication of vocabulary. If students who elaborated more in their prewriting also scored higher, then their higher score can be a reflection of increased fluency and complexity. An empirical study examining the specifics of the essays is needed to find out whether a higher score is indeed a reflection of more complex language, which would further support Skehan's Limited Attentional Capacity Model and Robinson's Cognition Hypothesis.

These findings have implications for the way that timed language placement essay exams are administered. The ESLPE exam, from which our sample is taken from, is a 60 minute exam, which is longer than other assessment exams with a writing component. Thus, though students are not obligated to prewrite, the time allotted gives students the freedom to allocate some time to prewriting. Though it is not known why some students did not prewrite, one reason could be their consciousness of the time constraints. For the students that did not prewrite because of their perception of their lack of time, their score may not be an accurate reflection of their writing ability. Furthermore, if the process approach is a framework for the writing process, which begins with prewriting, our study supports the necessity of such a step. Studies have shown that the process-oriented approach has not successfully been incorporated into assessment contexts (Cho, 2003; Freedman, 1993). Our study therefore supports providing more time for prewriting, if not integrating allotted time for prewriting during timed essay placement exams for the test to be a better indicator of their ability, although some research does caution against giving too much of an extended time to plan (Ong, 2013).

Though this study was conducted on university essay placement exam in an ESL context, the findings on the use and efficacy of prewriting activities on English timed essays can be applied to Asian EFL contexts, where English writing tests are developed and used as a part of overall English proficiency exams. For example, the timed essay test format is included in the General English Proficiency Exam (GEPT) developed and administered in Taiwan and the Test of English Proficiency (TEPS) created and used in South Korea. It should also be noted that the International English Language Testing System (IELTS) and the Test of English as a Foreign Language (TOEFL) are also widely used in Asia and both have timed essays.

Thus, our results would provide useful insights into the extent to which different prewriting activities are utilized and result in enhanced English essay products in those EFL English proficiency exam contexts.

Limitations

This study aimed to fill a gap in understanding the use and efficacy of prewriting strategies by L2 writers, as many studies have been with L1 writers. We found that while a majority of the students did prewrite, a large percentage also did not prewrite. Of course, it is unknown whether students prewrote in their minds and to what extent. Therefore, our research only applies to the efficacy of written prewriting strategies and does not include pre-discussions or think-aloud strategies. It should also be noted that although there is a significant score difference between the groups using and not using prewriting strategies, the difference between the means is relatively small. Thus, this issue needs to be replicated with other language groups, L2 writing ability levels, and prompts in future studies. Moreover, this study does not account for whether the performance was moderated by general L2 proficiency and knowledge of the targeted writing genre (Johnson, 2014).

Further investigation of these issues is necessary. For example, does the degree of other forms of prewriting (i.e. pre-discussions) make a difference in writing performance? What is the amount of time needed for prewriting to have optimal effects? Such research can provide ESL researchers, teachers, and test developers with a better understanding of the use and efficacy of prewriting strategies in L2 writers.

Conclusion

As the Process Approach dominates pedagogical writing practices, the first step - prewriting - is taught to be a necessary and significant step that facilitates the production of a well-written paper. While research in L1 writing seems to support the use of prewriting, we found that surprisingly few studies have examined the efficacy of prewriting strategies considering the significant differences between L1 and L2 writing. Therefore, this study aimed to contribute to the small, but growing body of research regarding the facilitative effects of prewriting in L2 writing. Our results confirm the significance and efficacy of prewriting for L2 writers during a timed exam, but also

demonstrate that it may not be the particular type of strategy employed that facilitates, nor the language(s) in which one plans, but rather the degree to which one develops his/her prewriting.

References

- Adegbija, E. (1991). A survey of students' prewriting activities and their implications for teaching. *English for Specific Purposes, 10*, 227-235.
- Benton, S. L., & Blohm, P. J. (1988). Elaboration of ideas in prose production. *Journal of Experimental Education, 56*(2), 60-66.
- Chai, C. (2006). Writing plan quality: Relevance to writing scores. *Assessing Writing, 11*, 198-223.
- Cho, Y. (2003). Assessing writing: Are we bound by only one method? *Assessing Writing, 8*, 165-191. Cumming, A. (1989). Writing expertise and second language proficiency. *Language Learning, 39*, 81-141.
- Dornyei, Z. (2005). *The psychology of the language learner: Individual differences in second language acquisition*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Elbow, P. (1973). *Writing without teachers*. London: Oxford University Press.
- Elbow, P. (1981). *Writing with power: Techniques for mastering the writing process*. New York: Oxford University Press.
- Ellis, R., & Yuan, F. (2004). The Effects of planning on fluency, complexity, and accuracy in second language narrative writing. *Studies in Second Language Acquisition, 26*, 59-84.
- Ferris, D., & Hedgcock, J. (2009). *Teaching ESL composition: Purpose, process, and practice*. New York: Routledge Taylor & Francis.
- Flower, L. (1989). Cognition, context, and theory building. *College composition and communication, 40*, 282-311.
- Flower, L., & Hayes, J. (1981). A cognitive process theory of writing. *College Composition and Communication, 32*, 365-387.
- Freedman, S. W. (1993). Linking large-scale testing and classroom portfolio assessments of student writing. *Educational Assessment, 1*(1), 27-52.
- Friedlander, A. (1990). Composing in English: Effects of a first language on writing in English as a second language. In B. Kroll (Ed.), *Second language writing: Research insights for the classroom* (pp. 109-125). Cambridge: Cambridge University Press.
- Gillis, M. K., & Olson, M. (1990). Do college students who plan before writing score better on essay exams? In T. V. Rasinski, N. D. Padak, & J. Logan (Eds.), *Reading is knowledge* (pp. 7-9). Pittsburg, KS: College Reading Association, Kansas City.

- Hyland, K. (1998). *Hedging in scientific research articles*. Amsterdam: John Benjamins.
- Hyland, K. (2003). *Second language writing*. Cambridge: Cambridge University Press.
- Johnson, M. (2014). Does planning really help?: Effectiveness of planning in L2 writing. *Journal of Second Language Teaching and Research*, 3(1), 107-118.
- Johnson, M., Mercado, L., & Acevedo, A. (2012). The effect of planning sub-processes on L2 writing fluency grammatical complexity, and lexical complexity. *Journal of Second Language Writing*, 21, 264-282.
- Jones, S., & Tetroe, J. (1987). Composing in a second language. In A. Matsushashi (Ed.), *Writing in real time* (pp. 34-57). Norwood, NJ: Ablex Publishing.
- Kellogg, R. (1988). Attentional overload and writing performance: Effects for rough draft and outline strategies. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 14(2), 355-365.
- Kellogg, R. (1990). Effectiveness of prewriting strategies as a function of task demands. *American Journal of Psychology*, 103(3), 327-342.
- Kormos, J. (2012). The role of individual differences in L2 writing. *Journal of Second Language Writing*, 21(4), 390-403.
- Kroll, B. (2002). Considerations for teaching an ESL/EFL writing course. In M. Celce-Mucia (Ed.), *Teaching English as a Second or Foreign Language*. Boston: Heinle & Heinle.
- Lally, C. G. (2000). First language influence in second language composition: The effect of prewriting. *Foreign Language Annals*, 33(4), 428-432.
- Lay, N. (1992). Composing processes of adult ESL learners: A case study. *TESOL Quarterly*, 16(3), 406.
- Manchon, R. M. (2009). Individual differences in foreign language learning: The dynamics of beliefs about L2 writing. *RESLA*, 22, 245-268.
- Mohseniasl, F. (2014). Examining the effect of strategy instruction on writing apprehension and writing achievement of EFL Learners. *Theory and Practice in Language Studies*. 4(4), 811-817
- Ong, J., & Zhang, J. (2013). Effects of the manipulation of cognitive processes on EFL Writer's texts quality. *TESOL Quarterly*, 47(2), 375-398.
- Ong, J. (2013) Discovery of ideas in second language writing task environment. *SYSTEM*, 41, 529-542.

- Ong, J. (2014). How do planning time and task conditions affect metacognitive processes of L2 writers. *Journal of Second Language Writing, 23*, 17-30.
- Perl, D. (1979). The composing process of unskilled college writers. *Research in the Teaching of English, 13*, 317-336.
- Raimes, A. (1985). What unskilled ESL students do as they write: A classroom study of composing. *TESOL Quarterly, 19*(2), 229-258.
- Rau, P., & Sebrechts, M. (1996). How initial plans mediate the expansion and resolution of options in writing. *The Quarterly Journal of Experimental Psychology, 49A*(3), 616-638.
- Robinson, P. (2005). Cognitive complexity and task sequencing: Studies in a componential framework for second language task design. *International Review of Applied Linguistics, 43*, 1-33.
- Rohman, D. (1965). Pre-writing: The stage of discovery in the writing process. *College Composition and Communication, 16*, 106-112.
- Schumann, J. H. (2004). The neurobiology of aptitude. In J. Schumann, Wood, LA., Lee, N., Jones, NE., Crowell, SE. (Ed.), *The neurobiology of learning: Perspectives from second language acquisition* (pp. 7-22). Mahwah, NJ: Lawrence Erlbaum.
- Sharples, M. (1999). *How we write: Writing as Creative Design*. eBook: Routledge.
- Shi, L. (1998). Effects of prewriting discussions on adult ESL students' compositions. *Journal of Second Language Writing, 7*(3), 319-345.
- Skehan, P. (1998). *A cognitive approach to language learning*. Oxford: Oxford University Press.
- Smet, M., Broekkamp, H., Gruwel, S., & Kirschner, P. A. (2011). Effects of electronic outlining on students' argumentative writing performance. *Journal of Computer Assisted Learning, 27*, 557-574.
- Spack, R. (1996). *Guidelines*. New York: St. Martin's Press.
- Stallard, C. (1974). An analysis of the writing behavior of good student writers. *Research in the Teaching of English, 8*(2), 206-218.
- van Weijen, D., van de Bergh, H., Rijlaarsdam, G., & Sanders, T. (2009). L1 use during L2 writing: An empirical study of a complex phenomenon. *Journal of Second Language Writing, 18*, 235-250.

- Walvoord, B., Anderson, V., Breihan, J., McCarthy, L., Robison, S., & Sherman, A. K. (1995). Functions of outlining among college students in four disciplines. *Research in the Teaching of English*, 29(4), 390-421.
- Weigle, S. C. (2014). Considerations for teaching second language writing. In M. Celce-Murcia, D. Brinton, M. A. Snow (Eds.), *Teaching English as a second or foreign language*. Boston: Heinle Cengage Learning.
- Yang, L., & Gao, S. (2013) Beliefs and practices of Chinese univeristy teachers in EFL writing instruction. *Language, Culture and Curriculum*, 26(2), 128-145.
- Zamel, V. (1983). The composing process of advanced ESL students: Six case studies. *TESOL Quarterly*, 17, 165-187.

Appendices

APPENDIX A: Essay Prompts

Topic A

For thousands of years human beings have first communicated with one another in the language of clothes. Long before I am near enough to talk to you, you announce your age, sex, and class to me through what you are wearing, as well as information – or miscommunication – about your occupation, origins, personality, opinions, tastes, and current mood. I register this information unconsciously, and you simultaneously do the same for me. By the time we meet, we have already spoken to each other in the language of clothes.

adapted from Alison Lurie, “The Language of Clothes”

Assignment: What is your view on the idea that people communicate important information about themselves through their clothes? In your essay, support your position by discussing an example (or examples) from literature, the arts, science and technology, politics, sports, current events, or your experience or observation.

Topic B

Competition is a remnant of a primitive past, but it does not come to use through the genes. It is passed onto us through training in our society, training that starts early. I do not believe that rivalry among children of the same family is instinctual; I believe that it is endangered by parents who were themselves victims of victims. As a result many children are pressured into Little League or similar activities and organizations, which really satisfy their parents’ craving for competitive success.

adapted from Theodore Isaac Rubin, “Reconciliations”

Assignment: What is your view on the idea that competition is a negative value imposed on children by their parents? In your essay, support your position by discussing an example (or examples) from literature, the arts, science and technology, politics, sports, current events, or your experience or observation.

APPENDIX B: ESLPE Composition Rating Scale

ESLPE Composition Rating Scale

Revised Fall 2000

- 6 Exempt from ESL Service Courses**
- Grammar is near native-like with little or no evidence of ESL errors. There may be basic writer developmental errors (e.g., spelling, sentence fragments & run-ons, interference from oral language)
 - The writing exhibits a near native-like grasp of appropriate academic vocabulary and register and there are few problems with word choice OR the writing is fluent and native-like but lacks appropriate academic register or sophisticated vocabulary.
 - Cohesion between paragraphs, sentences, and ideas is successfully achieved through a variety of methods (transitional words & phrases, a controlling theme, repetition of key words, etc.)
- 5 ESL 35**
- The number and type of grammatical errors are limited and usually follow a discernible pattern; these include errors in article usage, noun number, subject/verb agreement, verb form (tense/aspect) and word form.
 - Meaning is never obscured by grammar or lexical choices that are not native-like.
 - The writing exhibits fluency, which is achieved through a variety of simple and complex sentence structures. These are usually constructed appropriately, although there may be some problems with more complex grammatical structures (e.g., subordination or embedding relative clauses).
 - Register and vocabulary are generally appropriate to academic writing.
 - Cohesion is adequate and achieved through the use of transitional words and phrases.
- 4 ESL 33C**
- Grammar errors may occur in article usage and noun number, subject/verb agreement, verb form (tense/aspect), word form/choice, relative clause formation, passive voice, and coordination and subordination.
 - Errors are noticeable, but rarely obscure meaning.
 - The writing is less fluent than the 5 paper. It may sound choppy because there are too many simple sentences. Or there may be too many complex sentences that are too long to process. Or there may be some non-native-like sentence fragments and run-ons.
 - Vocabulary may be repetitive or inaccurate, and the register may exhibit a lack of academic sophistication.
 - There may be a limited lack of cohesion and difficulty with paragraphing.
- 3 ESL 33B**
- Patterns of errors occur in article usage and noun number, subject/verb agreement, verb form (tense/aspect), and/or word form.
 - Errors are noticeable and occasionally obscure meaning.
 - Although there is a good basic knowledge of sentence structure, the writing lacks fluency because of errors in or avoidance of relative clauses, passive voice, and/or coordination and subordination. There may be some non-native-like sentence fragments and run-ons.
 - Vocabulary is inaccurate in places and may rely on repeating words and expressions from the prompt.
 - The writing exhibits a basic knowledge of cohesive devices but these may be misapplied, or the devices used may not create cohesion.
- 2 ESL 33A**
- Frequent patterns of errors occur in article usage and noun number, subject/verb agreement, verb form (tense/aspect), and/or word form.
 - Errors are noticeable and obscure meaning.
 - Although there is a basic knowledge of sentence structure, this is not consistently applied. The writing exhibits errors in and avoidance/absence of relative clauses, passive voice, and/or coordination and subordination. There are non-native-like sentence fragments and run-ons.
 - Vocabulary is generally basic and word choice is inaccurate. The writer may rely on repeating words or expressions from the prompt. The register can often resemble either a conversational narrative or a stilted, confusing attempt at academic discourse.
 - Although there is some use of cohesive devices, it is neither consistent nor always effective, and may be simple and repetitive in many cases.
- 1 Pre-University**
- Pervasive patterns of errors occur in article usage and noun number, subject/verb agreement, verb form (tense/aspect), and word form.
 - Except in very simple sentences, meaning is frequently obscured.
 - A basic knowledge of sentence structure is lacking, and there are frequent errors in and/or avoidance of relative clauses, passive voice, and/or coordination and subordination. When sentences are complete, they are often simple or are expressions learned as “chunks.”
 - Vocabulary is quite basic, and more sophisticated attempts at word choice are often inaccurate or inappropriate. The register is often too conversational for academic purposes or, if an academic tone is attempted, it is incomprehensible.
 - There may be attempts to use cohesive devices but they are either quite mechanical or so inaccurate that they mislead the reader.
- 0 No Response**

APPENDIX C: Prewriting examples

1. Outlining example (TT290)

Approve.

① reflect culture

- China
- western
- mid Eastern

ie

2. ① tell personality

- formal
- sloppy
- like sport

IS

① politics

→ voting
→ lecture

→ dress persuade

Madam
primiver.

2. Listing example (TT464)

- UCLA Tees, sweatshirts & other clothing
- class / society
- India lower class, upper class etc.
- revealing judged as slutty
- formal meetings
- UCLA store
- kings in France → Hot culture
- traditional clothing

3. Webbing example (TT251)

dress code
current moods - light color
dark
bright
costume party WRITE
style, cowboy jeans
practical = warm
sweat (dry-tech)

4. Combined strategies (outline and draft) (TT271)

DRAFT

- 1) More Entertainment → Sport Competition
- 2) More Efficiency → Technological development.
- 3) More positive work → Politicians.
- 4) Better boundaries.

~~Introduction~~ ¶

As far as anyone can recall, competition is a value human beings have always had. Tribes fighting for a piece of meat ~~can be a~~ clearly shows how ancient this value is. ^{members} in the Paleolithic & Neolithic eras.

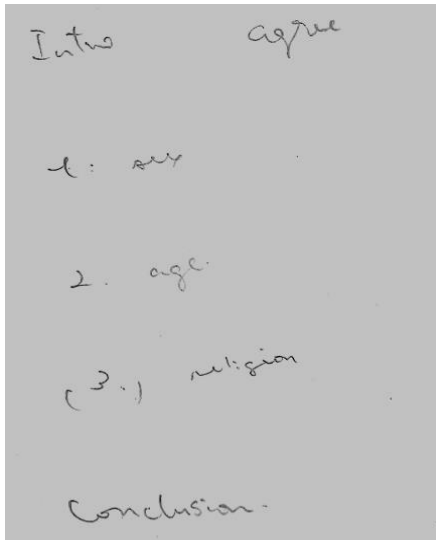
Surprisingly, it is not a value specific to human beings but can also be found in animal behavior: two males competing to seduce a female.

Thus, the origin of this value is a controversial issue, ~~is it~~ Does it come to us through the genes? Or is it a value unconsciously imposed by society on us?

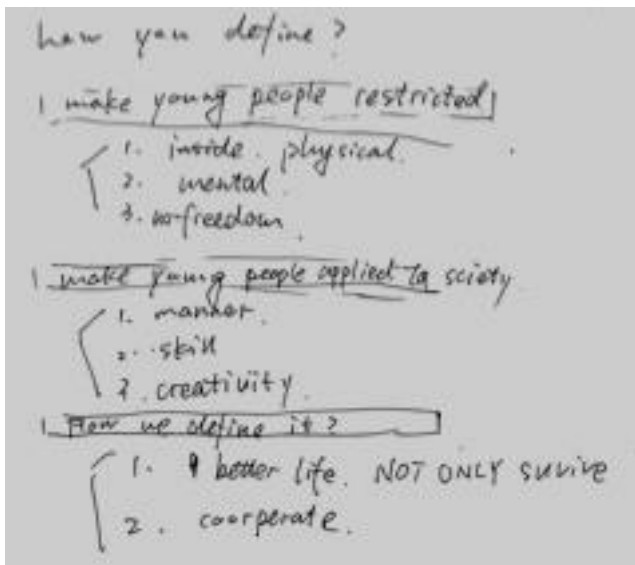
In "Reconciliations", Theodor Isaac Rubin argues ¶

APPENDIX D: Degree of development in outlining

1. Example of a “minimal” outline (TT453)



2. Example of a “standard” outline (TT129)



3. Example of an "elaborate" outline (TT481):

Competition is a negative value imposed on children by their parents?

- ① Opening
Intended for a good purpose.
- ② Bad sides - Ex. 1 gymnastic athletes.
- ③ Bad sides - Ex. 2 personal experience.
- ④ Good sides - Ex. 3 high productivity (academic research)
- ⑤ Good side - Ex. 4 Olympics bonding ~~and~~ establish friendship.
- ⑥ Conclusion -
 - ① not necessarily imposed by parents for their own desire → success
 - ② Can be utilized → better the world

APPENDIX E: Degree of development in webbing/clustering

1. Example of a "minimal" webbing (TT 401):

competition is part of society

games → one winner

motivation

→ games → not bad

→ school → useful

→

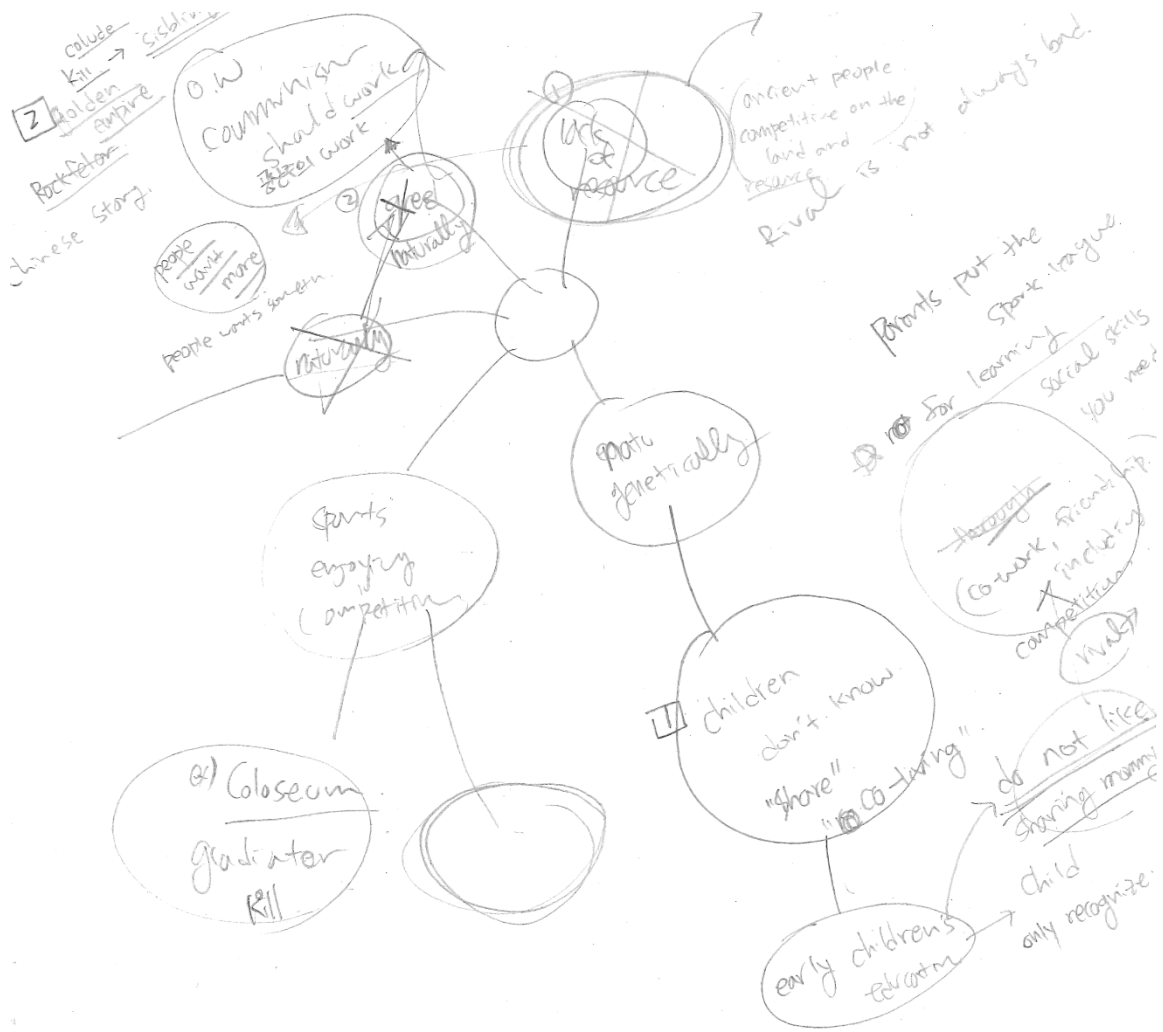
2. Example of a "standard" webbing (TT 237):

same - different - global standard. international brand.
Primitive - Contemporary.

Dress Code

Communicate information — fashionable. taste.
— materiality.

3. Example of an "elaborate" webbing (TT406):



**No Point in Talking about What I Want to the Teachers:
A Call for a Dialogic Needs Assessment**

Soyeon Kim

Korea Institute for Curriculum and Evaluation (KICE)

Bio data

Soyeon Kim is an Associate Research Fellow at the Korea Institute for Curriculum and Evaluation in Korea. Her research interests include second language writing, multicultural education, teacher education, and critical pedagogy. Currently she is exploring ways to support socio-economically underprivileged English language learners in Korea using a Bourdieusian framework. soynell3@kice.re.kr

Abstract

A student needs assessment is one of the crucial stages in designing or revising a syllabus for EFL writing classes. However, teachers often experience difficulties in finding out what their students want to learn from English writing classes. As a way to understand students' wants and needs, teachers often use a questionnaire or open-ended question survey. However, it is problematic if one-shot, non-formative, and non-dialogical types of needs assessment are utilized as standards. These types of needs assessment also have a limitation in that students do not know what they need or want in a specific learning situation until they actually engage in learning the content. For this reason, this qualitative study is designed to trace the transformations of three Korean university students' needs during a semester using a dialogic framework. Data, including three rounds of in-depth interviews, field notes, and class observation, was collected and analyzed according to themes and types of needs. The results showed that student needs were transformative from vague to specific needs about the writing class. It was also found that dialogic needs assessment provided a formative learning experience for students and teachers by empowering students. However, the limited type of

Address correspondence:
Jeongdong Bldg, 15-5, Jeong-
dong, Jung-gu, Rm #614 Seoul,
100-784, Korea

evaluative teacher feedback cultivated misunderstanding about the teacher's level of interest in students' learning. Pedagogical implications for researchers and writing

teachers in EFL countries will be provided for them to be more aware of on-going dialogic needs assessment and ways to implement dialogic needs assessment in large scale writing classes will be introduced.

Keywords: needs assessment, dialogism, needs, college writing courses, Korean university students

Introduction

Understanding language learners' needs is one critical element in designing an effective English language course. Many studies have been conducted to examine student needs through a range of quantitative research techniques, such as surveys. However, teachers still fail to understand what their students need or want because students themselves do not know of what they need or want unless they have undergone the learning content of the class. Student needs become more difficult to examine when the students have little learning experience with the subject, such as writing in English for EFL learners. For this reason, the strength of this qualitative study is that it explores three Korean university students' perceived needs and wants and discusses how those stated needs transformed as a composition course progressed during one semester using an on-going process of data collection.

The theoretical framework of this study is drawn from dialogism of Bakhtin (1982) and practiced by Freire (2000) and Benesch (2001) in education and second language writing fields respectively. In order to understand Korean university students' perceived needs and wants for an English writing class, where communication between the teacher and students plays a critical role, qualitative data collecting techniques were used, such as class observation, ethnographic interviews, and field notes. In this study, the student needs and wants were not considered to be concrete knowledge or a truth that the teacher could discover from the students. Instead, student needs were treated as a kind of verb, which transforms, changes, and thus can be negotiated between the teacher and the students. Four types of needs among the six categories of needs (Hutchinson & Waters, 1987) were used as tools to analyze and explore the students' stated needs in three different phases of the interviews. Unlike studies which focus on developing a data collection framework for student needs in the process of designing language courses, the current study suggests an on-going process needs analysis as a part of students' language learning experience, extending its commonly expected role of curriculum development.

The focus of this article is one thread of a larger ethnographic style study exploring the negotiation process and strategies between a Korean writing instructor and seven Korean students about their changing needs. For this article, the data has been reanalyzed and presented as a case study focusing on three university students. Through the data analysis, the emerging themes of stated student needs and how these needs were transformed will be mainly discussed. The reasons why some of the students felt reluctant to share what they felt they needed for their writing class will be highlighted, discussing the impact of customary student needs assessment on the students' understanding of student needs assessment.

Feedback is considered necessary to improve students' writing and thus has become an

important part of the process of teaching writing (Hyland, 2000, Wingate, 2010). Many studies have been conducted to understand the role of feedback (Hyland & Hyland, 2006), the effectiveness of the quality of teacher feedback on students' understanding (Invanic, Clark, & Rimmershaw, 2000; Lea & Street, 1998; Lizzio & Wilson, 2008; Mutch, 2003), students' perceptions of formative feedback (Carless, 2006; Walker, 2009), and types of feedback and students' motivation (Ashwell, 2000; McGarrell & Verbeem, 2007). Unfortunately, in spite of the importance of feedback, writing instructors often face a challenge in using formative teacher feedback, particularly when they deal with large class sizes (Denton, Madden, Roberts, & Rowe, 2008). Some of these teachers replace formative feedback with simple evaluative terms or a numerical score. However, few studies have investigated how students perceive a writing class where these limited types of feedback are mainly used. The results of using limited types of feedback will be presented, discussing the dialogic functions of feedback in writing classes. Finally, the suggestion that employing dialogic needs assessment in writing classes will be discussed with pedagogical implications for English language educators and particularly for English literacy educators.

Literature Review

Dialogism in Education

Dialogism lies at the heart of the theoretical and conceptual framework of this study. Discussing *heteroglossia*, the multiple voicedness in speech, Bakhtin (1982) explains the two forces at work. One is the centripetal forces that move linguistic elements and rhetorical modes toward the center and unity of language to make a standard or official language. These forces aim at ideological unification and centralization in linguistic life. Centripetal forces have a positive function in education in that they help people to overcome the heteroglossia in language, maximizing mutual understandings and agreements. The second are the centrifugal forces which work as processes of "decentralization and disunification" of elements of language and produce various languages, dialects, or genres (Bakhtin, 1982, p. 272). Centrifugal forces work to produce the dynamics, uniqueness, and multivocalness in nature, and thus, education as well. As the two forces work in speech, the uninterrupted dialogic processes of both centralization and decentralization are enacted at the same time.

Dialogic relations between the centripetal and centrifugal forces are both descriptive and idealized relationships by virtue of struggle and negotiation. In other words, the dialogic relations do not represent only a natural logic but also an idealized goal of nature. Lillis (2003) discusses the two levels of Bakhtin's dialogic notion. On level one, dialogic relations are a

given in the nature of language. On level two, the dialogic relation between the two kinds of force is “something to struggle for,” (p. 197) and thus, is an idealized condition in human nature and communication that we need to achieve as a goal. Dialogism in this study will be understood as the goal that the teacher and students need to struggle to achieve so as to have a mutual understanding about their different expectations in writing in English.

Many educators focus on the dialogized heteroglossic power in human nature and the language classroom, and attempt to seek solutions from this dialogic notion for the problems caused by unbalanced power and monological ideas in education. In this case the centrifugal forces represent the multiple voices of students as well as the dynamics that the students bring into a classroom. Freire (2000) sheds light on practicing this dialogism in the education field. He criticized the monologic “banking” (p. 72) image of education and “narrative” (p. 71) patterns of teacher-centered classrooms. He asserted that it is problematic if teachers see themselves as banks that are full of the right knowledge and students as passive depositories. In this relationship, students become objects which need to be filled with given knowledge from teachers, and thus, students are not required to think, feel, make decisions, negotiate, or be critical.

Students who have different cultural and socioeconomic backgrounds bring diverse voices about their wants and needs into composition classes. In order to accommodate these varied needs, it is important for teachers to dialogue with these multi-voices of their students and decide what would work best for them to assist their learning to write. However, if the negotiations between the two directions of centripetal and centrifugal forces among students as well as between students and teachers are forbidden, this works against human nature and will eventually have a negative impact on students’ learning as well. Wong (2006) warns about the authoritative discourse in second language classes which is “univocal” (p. 134) and does not permit dialogue. She explains that this is opposed to the unfinished, open-ended, and multi-voiced dialogism that the teachers strive to achieve. Unfortunately, such authoritative discourse is often observed in many parts of Korean education and other locations, being represented by an authoritative teacher figure that solely makes decisions for a class.

Dialogic interaction is more necessary in large English composition classes, which are common in classrooms in the Asian EFL context, such as in Korea, China, Taiwan and Japan. EFL teachers often face challenges when they attempt to use teaching approaches which have been developed in English speaking countries (Shamim, 1996). Unfortunately, this is the case for most EFL practitioners since many approaches have been developed in ESL teaching contexts where language classes are relatively smaller than in expanding circle countries

(Kachru, 1992). Acknowledging that large classes are the reality in many parts in the world, Wong (2006) suggests that dialogic interaction is possible even in these teaching contexts, if the teacher is “aware of students’ individual perspectives, backgrounds, and needs” (p. 69). For this reason, large classes in EFL are a good example of sites of struggle, where students’ multi-voices and interests are in conflict but rarely negotiated. Thus, this study aims to explore students’ needs, addressing the issues involved in teaching a large composition class in Korea.

Dialogic Needs Analysis

Even though dialogism is a critical issue in TESOL and education in general, the dialogized heteroglossic function has not been emphasized in needs assessment. The general definition of needs refers to the information from a group of students about their expectations of a particular class (Songhori, 2008). The term ‘needs’ is frequently used as an “umbrella term” (West, 1994, p. 3) encompassing issues from broad student expectations about a specific course to linguistic information gathered from teachers to the design of a language class. Hutchinson and Waters (1987) proposed the six categories of necessities, lacks, wants, learning strategies, constraints, and the language audit. Four of these were modified and utilized in this study: wants, lacks, learning strategies, and constraints (see Table 1).

Table 1
The Four Categories of Needs

| Categories | Descriptions |
|---------------------|---|
| Wants | “What the learners want or feel they need” (Hutchinson & Waters, 1987, p. 57): “Subjective needs” (West, 1994, p. 4) |
| Lacks | The perceived gap between the necessities and what the learner knows already |
| Learning Strategies | Learners’ preferred learning strategies for their language learning and development |
| Constraints | Perceived constraints of the learning situation such as the available facilities (staff, time), attitudes, culture, and materials |

Note: Adapted from Hutchinson, T. & Waters, A. (1987). *English for specific purposes*. Cambridge. Cambridge University Press.

In this study, wants refers to subjective needs which are derived from students’ statements about what they want from a course. Hutchinson and Waters (1987) define lacks as the gaps between what the learners already know and what they need to learn. Since the study did not involve lacks identified by a teacher, however, the definition of this term will indicate the perceived

lacks from the students' side. Learning strategies mean the students' preferred strategies used when they are learning English. Constraints refer to the restrictions in language classes from attitudes, culture, and availability of facilities. In terms of these four categories, the transformation of the students' understanding about what they want from their writing class will be explored.

Three main needs assessment approaches have been proposed to understand students' needs and develop an effective curriculum for language learners. These are Target situation analysis, Present situation analysis, and Means analysis. Munby (1978) developed an extensive framework for student needs assessment to develop a communicative language syllabus, which later became known as a target situation analysis (Chambers, 1980). From the understanding that "the systematic processing of the profile of communication needs for the particular participant input is a prerequisite to the syllabus specification output" (Munby, 1978, p. 218), he attempted to collect various data about language learners to identify linguistic forms that language learners could be anticipated to use in certain situations. However, Munby's model has four limitations. First, it was too complicated and time consuming to collect intensive information about learners (Hutchinson & Waters, 1987; West, 1994). Second, it failed to tap into learners' sociopolitical and affective concerns (Dudely-Evans & St. John, 1998; West, 1994). Third, this approach fails to specify how to develop a communicative syllabus using the collected information (Dudely-Evans & St. John, 1998; Songhori, 2008) and, finally, it is unable to provide information about the language ability of the students in the present situation.

In order to compensate for the drawbacks of Target situation analysis, Present situation analysis was proposed by Richterich and Chancerel (1977/1980) which collects data about "learners' current proficiencies, perceptions, and ambitions" (Hyland, 2007, p. 155). Researchers (e.g. Benesch, 2001; Johns, 1990; Prior, 1991, 1995; Ramani, Chacko, Singh, & Glendinning, 1998) conducted studies using present situation needs analysis to design syllabuses which better accommodate students' needs. The fast development of ESL and EFL courses throughout the world, however, has brought another problem, that of implementing language syllabuses mainly developed in English speaking countries into countries where English is a foreign or second language.

Means analysis (or the ecological approach) has been proposed in response to the fast growing demand for language courses in these outer and expanding circle countries (Holliday & Cooke, 1982; Holliday, 1994). Means analysis has its value in that it helps educators understand the significance of being sensitive to local context and prevents the imposition of teaching curriculums alien to the local situation (Holliday, 1994). Acknowledging that "what works well

in one situation may not work in another” (Dudley-Evans & St. John, 1998, p. 124), the advocates of Means analysis suggest language teachers identify information about their teaching contexts (the ecosystem) which includes “the teachers, methods, materials, facilities, and the relationship of the course to its immediate environment” (Hyland, 2007, p. 155). In this sense, Means analysis succeeds in recognizing the wider socio-cultural contexts in needs assessment.

However, all three main research paradigms have two common deficiencies in responding to students’ needs. First, quantification of dynamic student needs is problematic in that these approaches treat student needs as measurable objects (Hyland, 2007) or as a collectivity (Benesch, 2001). Second, these approaches are monologic as they imply only researchers’ involvement in collecting, analyzing, and discussing students’ needs. Students are the original beneficiaries of the needs assessments because these are the records of the journeys of searching for students’ wants and needs to create effective language classes. However, ironically, these approaches have neglected to invite the agents of their learning to this journey. Consequently, there has been a demand for a new approach focusing on dynamic students’ needs with the invitation of the students who are the agents of learning.

This study proposes a dialogic needs assessment framework as an alternative form of needs assessment to delve into three Korean university students’ dynamic needs. In this study, needs are not treated as measurable objects (Hyland, 2007) but transforming qualities that should be discussed through a negotiated process as a part of student’s learning experience. Dialogic needs assessment sheds new light on the rights analysis of Benesch (1999a, 1999b, 2001). Benesch (2001) replaces “rights” (p. 62) with the term needs and explains that rights is a more appropriate term in describing tensions between students and the teacher, as it indicates students’ right to negotiate their needs with their teachers. Researchers (i.e. Dudley-Evans, 2001), suggest language educators focus on rights analysis because, while needs analysis is a starting point for discovering institutional requirements and expectations, rights analysis provides the momentum for change (Hyland, 2006, 2007). Through her series of empirical studies (Benesch, 1999b, 2001), Benesch explored ways to assist ESL students’ learning using rights analysis and the results of other empirical studies (Flowerdew, 2005; Molle & Prior, 2008) also emphasize the importance of raising awareness of students’ rights to share their needs and participate in changing their classes.

In this study, even though the three students were not able to contribute to modifying their writing class, rights analysis (Benesch, 2001) played an important role in helping to examine their changing needs. The conceptual framework is similar to Benesch’s rights analysis in that needs analysis is treated as a practice situated in students’ learning. However, the term right will

not take the place of needs. Instead, the researcher proposes to understand the students' situated needs in the on-going dialogic framework (Bakhtin, 1982). For these reasons, two questions have driven the study: first, what expectations and needs are brought by the three Korean university students about their writing class? Second, how and why did their needs transform as the class progressed throughout the semester?

Research Design

Site of Study and Participants

The research site was a large English writing class at a university located in Seoul, Korea. The class, English Writing through Movies, was an elective course available to any student in this school. Four classes were offered, each with a maximum enrollment of 80 students. A male adjunct professor taught the four sections and data collection was carried out with his consent. The observed class was two hours long, which included a one hour lecture and another hour of in-class writing. It had 40 students, which is regarded as a large English writing class.

Three students, Adam, Christine, and Dong-hwan (pseudonyms) participated in this study. Adam, a 21 year old freshman in the School of Communication, had an early education background in the United States and Ireland. He learned to write in English in his secondary school years in Ireland, which involved writing only 300 to 1000 word summaries and reports, and thus, he stated that he needed to learn to write academic essays. Adam was studying English to prepare for the Test of English for International Communication (TOEIC) for future job applications. Christine, a double major in the Department of Public Administration and Public Policy and the School of Business Administration, was a female participant. As many Korean students do, Christine began to learn English in elementary school. Most of her previous English courses focused on training students for school tests and the university entrance exam. Since these English tests did not assess writing skill, she neither practiced nor learned writing in English until she studied English for the Test of English as a Foreign Language (TOEFL) a year ago. Last semester, in her disciplinary course, she learned how to write a cover letter and wrote a narrative essay of what she had done during *Chusuk*, the Korean Thanksgiving day. However, she did not receive any teacher feedback on her paper. Her papers were simply graded and the scores were included in her final grade. Dong-hwan was a 23 year old student in the School of Business Administration. Since he had a relatively clear goal for his future, receiving a certain range of score in TOEIC was an important goal for him. Dong-hwan was taking a speaking course in a private language school in order to prepare for the TOEIC speaking test. In terms of experience of learning writing in English, he did not take any formal classes before taking this

class.

One of the advantages of a qualitative study is being able to do in-depth research into a particular phenomenon in a natural teaching context through field work. In order to examine multifaceted students' needs, this study employed *emic* (insider's) and *epic* (outsider's) perspectives. In order to do this, the researcher played two different roles in the writing classes: a researcher and a doctoral student. Creating close relationships and becoming a member of the writing classroom was a critical key point in being able to successfully conduct this study. Maxwell (2005) suggests that the best way for a researcher to position herself is "to put [herself] in their position, and ask how [she] would feel if someone did to [her] what [she is] thinking of doing, while making allowances for differences in culture and norms" (p. 85). Thus, by taking the sides of both the professor and students, the researcher attempted to create rapport with the writing instructor as well as gain membership in the students' community.

Data Collection

Multiple qualitative data was collected through participant observation, field notes, ethnographic interviews, casual conversation, artifacts (such as exchanged emails, copies of text messages with the participants, and outcomes from classroom writing activity), and weekly field notes for one semester. The researcher observed every meeting of the writing class and recorded basic information (such as number of students, topics discussed between the students and the teachers, communication methods, and topic of the day) required for participant observation (Angrosino, 2007). In the field notes, as Street, Baker, & Tomlin (2005) suggest, two columns were used to separate observations from interpretation or comments. Three interviews of around an hour were also conducted with each student participant. All interviews were recorded and transcribed later with the permission of the students. In the initial stage, grand-tour questions were employed (Crang & Cook, 2007; Spradely, 1979) and, in the later interviews, more focused follow up questions were used. Casual conversations with or among the participants were recorded through the researcher's recollection and the accuracy of the notes were checked with the participants.

Data Analysis

The collected data was analyzed through three different phases. In phase 1, the data was analyzed through open coding with certain codes developed. In the second phase, patterns were found and labeled into themes. In the last phase, the researcher reviewed all of the data and confirmed the themes by "re-ordering, re-contextualizing, and re-assembling the data" (Crang &

Cook, 2007, p. 133). To manage the massive amounts of collected data, NVivo 2 was used as a tool for assisting building codes and managing their hierarchy. It is worth noting that NVivo 2 is not an automatic theory or code builder software but an efficient assistant tool for sorting coded data and retrieving it (Park, 2004, 2009). With it, four categories of needs were used to track the three students' needs in three different phases.

Results and Discussion

Transformation of the Themes of the Students' Needs

Table 2 shows the results of the three Korean students' perceived needs based on the three interview phases. Various themes emerged according to the three different phases of the interviews. In the first interview, there was a particular theme which was irrelevant to a writing course. This indicates that the Korean students did not have a clear understanding about writing in English classes. In the first and second interviews, two students (Christine and Dong-hwan) expressed that they had an emotional block about discussing their needs. As the interviews progressed, a participant was able to articulate the reasons why she felt reluctant to discuss her needs about the class.

Table 2
Types of Student Needs and Developed Themes

| Interview | Time period | Types of Needs | Themes |
|-----------|---|---------------------|--|
| First | 2nd~3rd week of the class | Wants | Practical conversation |
| | | Wants | A logical essay |
| | | Wants | Organization of essays |
| | | Wants | Resume, TOEFL, business writing and diary |
| | | Wants | Peer discussion |
| | | Constraints | Emotional barriers to discuss needs |
| Second | 8th~9th week of the class | Lacks & constraints | Peer discussion |
| | | Lacks & constraints | Teacher and peer feedback |
| | | Wants | Sample writing models |
| | | Learning strategies | Formulaic writing |
| | | Constraints | Emotional barriers to discuss needs |
| Third | 14th week ~ after the semester | Wants | Communication between student and teacher or student and student |
| | | Wants | Teacher feedback |
| | | Learning strategies | Writing examples (Writing formula) |
| | | Wants | Dividing classes into different levels |
| | | Lacks | Punctuation/basic forms |
| | | Lacks | Different writing genres |
| Wants | Ask my needs in the midst of the semester | | |

In the first interview, six themes evolved: practical conversation; a logical essay; organization of essays; resume, TOEFL, business writing and diary writing; peer discussion; and emotional barriers to discuss needs. First, students expressed their need to learn “practical expressions” (Christine, first interview) and “English which I can use right away” (Dong-hwan, first interview). These students were not able to clearly articulate what they meant by ‘practical.’ However, Christine explained, “I mean, rather than grammar, focusing on English speaking, which I really can use in my daily life” (first interview). According to this statement, she revealed her wish to learn daily conversation from this composition class. Second, Adam and Christine expressed their desire to learn how to write a logical essay. Third, Adam said that he wanted to learn the structure of English writing. In the same vein, Christine mentioned logical flow as an important element that she wanted to learn in this class. Fourth, Adam and Dong-hwan talked about other writing genres that they wanted to learn in a composition class, such as writing preparation for the TOEFL test and business writing. Fifth, Adam suggested that he wanted to have an opportunity to discuss his paper with his classmates. Finally, as a constraint, Christine stated her emotional barriers to discuss her needs about the writing class due to her previous needs assessment experiences. Being reluctant to share her needs about the class, she explained that she thought that “it’s useless to talk about what I want” (first interview).

As the results show, these needs were generally very vague and students had a hard time articulating their needs about their writing class. In other words, the students did not have a clear idea about writing in English classes. For example, Christen and Dong-hwan wanted to learn speaking in the writing class. Also, the students generally used vague and uncertain terms such as practical English, the logical flow, and the structure. When they were asked to explain more about these terms, they were neither able to explain further nor could they provide precise examples.

In the second interview, five themes were found: peer discussion; teacher and peer feedback; sample writing models; formulaic writing; emotional barriers to discuss needs. The students started discussing their needs that were closely relevant to the particular writing class. First, Adam expressed his need to have a peer review so as to discuss his writing topics and develop his ideas. Acknowledging his awareness of the problems of studying in large classes, he proposed an alternative way to have peer discussion, which he believed to be an important part of his learning. He said, “if there are too many students in the classroom, [my professor] can choose a day and hold the discussion, making groups of certain sizes. Students would then gather and talk to each other. Discussion is pretty [important]” (Adam, second interview). The second theme was detailed teacher or peer feedback on the writing. Christine said, “I hope he

can give us feedback, once I submit the drafts” (Christine, second interview). However, she also admitted that it was difficult to have detailed feedback due to the large number of students in the class. The third theme, sample writings, were proposed as an alternative solution to teacher and peer feedback. Students assumed that they could understand their professor’s expectations through the feedback process. Christine and Dong-hwan explained that due to the lack of teacher and peer feedback, they were having a hard time figuring out what the writing instructor wanted to see from their papers. Christine spoke of her suggestion, saying, “[if I were the teacher,] I would show my students many kinds of examples. [I believe] the more students read the more they can learn. They can write something by imitating. That’s pretty helpful” (Christine, second interview). Dong-hwan also wanted to look at writing examples. He said “well, for example, such as providing someone’s essay. Then, I will think, ‘aha!’ That’s a good idea too” (second interview). Fourth, Dong-hwan consistently expressed his desire to learn formulaic writing. Dong-hwan attributed his expectation to learn formulaic writing to his TOEIC speaking class where an instructor had him memorize certain sentences in order to answer questions in the test. He claimed that the writing class was difficult because the professor did not provide “efficient” writing information such as “In the introduction, you can use this sentence, and then write your own argument, like that” (Dong-hwan, second interview). Fifth, again, Christine mentioned the discomfort that she felt about discussing her needs.

The third interview results revealed seven different needs which were very clearly targeted to the current writing class. These needs were communication between student and teacher or student and student, teacher feedback, writing examples, dividing classes into different levels, punctuation/basic forms, different writing genres, and ‘Ask my needs in the midst of the semester.’ First, continuing from the second interview, Adam found dialogues between classmates or between the students and the writing professor critical. He conveyed his strong desire, saying, “this is the class where communication is really needed” because “the most important part in writing is developing ideas” (Adam, third interview). Adam stated that he wanted at least a casual conversation with his classmates about his drafts. Second, the students still conveyed their desire to have teacher feedback. However, students were well aware of the fact that the professor could not provide precise teacher feedback due to the large class size. Dong-hwan especially had a strong point of view about having formative teacher feedback. Expressing disappointment about his final grade, Dong-hwan said that he “did everything wrong” (third interview) in his essay and “if he [the writing professor] could provide a chance to have a better understanding” about essay writing, he could have written a better essay. He further explained that what he needed was to have “teacher feedback so that I could check if I

am on the right track or not” (Dong-hwan, third interview). In the case of Adam, he received the best grade in this writing class. However, he still considered his writing to be “okay” but not a “good essay” since he “wrote it without any kind of feedback” (Adam, third interview). Third, Christine and Dong-hwan once more claimed that the writing professor needed to provide them with writing samples. The two students explained that they had had a rough time figuring out the professor’s expectation of their essays because they did not have a chance to look at a full example that the writing professor might like. Christine found that the examples in her textbook were “unclear for us as we read only one short example in each section” (third interview) of the introduction, body, and conclusion. Her recollection was that “I practiced *non-sul* (a Korean essay) by imitating the sample essays. It’s like learning from imitating at the basic level of writing.” Dong-hwan wanted a writing model that he could modify and use for his essay as he learned Korean writing in this way. Thus, while Christine wanted to verify the genre of the essay from writing models, Dong-hwan delineated English writing as learning and applying writing formulas from writing samples.

Fourth, Adam, who had the longest and most varied English writing experience, emphasized that, if he were the writing professor, he would “divide the writing class into different levels” (third interview). Adam believed that if there were many writing classes targeting the students at a certain level, he could learn more about writing in English. Fifth, classes teaching different writing genres were needed for the students. Christine referred to her need to learn diverse writing genres. According to this, she seemed to develop genre awareness of writing through the semester. Sixth, for the first time, Christine mentioned her need to learn writing mechanics and punctuation. Christine identified “a format” (third interview) as what she needed to learn from a composition class. To be more specific, she explained by saying, “such as Korean people write the title here [pointing to the top center of the paper] and name there [pointing at the right top].” Due to her limited knowledge of the writing format, she had simply guessed what the essay should look like and submitted her paper. Finally, Christine argued that needs assessment should be conducted in the midst of the semester. This will be discussed more in relation to her resistance to talk about her needs in the previous two interviews.

Transformations of Types of Student Needs

The results show that the students’ unclear wants of the first interview evolved into clearer lacks, constraints, and learning strategies about their current writing class in the second and third interviews (Table 2). In the first interview, their stated needs were categorized as five wants with one constraint that were very vague statements. In other words, the students were

uncertain about what they wanted from the class. However, from the second interview, the three students began to clearly identify what they wanted from the particular writing class. On every occasion the researcher rephrased the questions and asked about their needs in a general composition class, they usually went back to refer to their current English writing class. This reflects that the students' sensitivity to their immediate situation made them more concerned about discussing their immediate needs in their current writing class (Benesch, 1999a, 1999b, 2001; Flowerdew, 2005). This supports Benesch's argument that students' needs are something "not pre-established but must be discovered" (2001, p. 109) because the students could clarify what they wanted after they experienced learning in the writing class.

Thus, student needs must be understood as a situated practice of "discover[ing] what is possible, desirable, and beneficial at a certain moment with a particular group of students" while "experimenting with ways to modify them" (Benesch, 2001, p. 109). The immediateness of response to the needs assessment results and the issue of who were the beneficial subjects of the needs analysis were a big concern for the three participating students. Even though they knew that there was little opportunity to alter their writing class, they wished to talk about what they wanted to change because the current class situation was the matter of interest in their needs assessment. In the last interview, the types of student stated needs were very focused on wants, lacks, and constraints in the current writing class. This result emphasizes again that a needs assessment must be able to delve into students' situated needs in their present learning context.

Three reasons for the students' unclear needs were revealed. First, the students had limited writing experience in learning writing in both English and Korean. This made them unable to anticipate what elements would be important for learning writing. For instance, for Dong-hwan, this was the first formal class he had taken to learn writing in either English or Korean. Even though he learned how to write essays in Korean indirectly through his various subject classes, he had not taken a formal Korean composition class. Adam also learned writing in English in his non-composition classes during his K-12 education, and thus, did not have a clear idea about what he needed to learn in a composition class. Christine was the only student who had previous experience in a formal learning context. However, since the class was a test preparation class, she also had limited experience in learning to write in English.

Second, Korean university students are rarely required to write in English outside of class (Lee, 2006). Due to the lack of need to write in English, they could not imagine what they needed to learn and the situations for writing in English. For example, though Adam and Dong-hwan mentioned that they wanted to learn business writing, they could neither explain what they meant nor provide specific examples. They merely assumed that they might be writing

some documents in English when they worked in a company in the future.

Lastly, the three students rarely had an opportunity to talk about what they wanted in previous classes. This was why they took a relatively longer time to respond to these questions and why their answers were unclear. However, even though the three students found it difficult to think and talk about their needs, they soon actively engaged in the conversation and expressed their wish for their professor to listen to their opinions and respond to them in his class.

Having No Feedback Means Having No Communication

The professor, in fact, provided minimal written feedback on his students' papers, marking them with evaluative terms, such as great or weak, due to the large number of students in four classes. According to the professor, these minimal comments were meant to show his enthusiasm for and interest in his students' writing practice within the limited teaching environment. However, the participant students did not perceive this minimal feedback as a formative feedback, even stating that they were not able to receive teacher feedback.

Moreover, students equated communication with feedback which ultimately caused discomfort due to its absence. Associating peer discussion as a form of feedback, Adam identified the lack of communication between himself and the professor or among classmates as the greatest disadvantage. Christine and Dong-hwan attributed their frustration at having difficulties in understanding what their professor wanted from their essays to lack of reader feedback as a written form of communication. These students revealed their awareness of the dialogic nature of writing as a socially engaged practice. Since the dialogic nature of writing (Killoran, 2005) was restricted in the class, students felt anxiety in this writing class.

An interesting issue is that the three students complained particularly about the failure to have any dialogue to enhance their ability to learn about writing. The students asserted that they had needed help to expand their ideas in their writing and wanted to learn from their writing mistakes through written or spoken conversation with others. For example, Christine and Dong-hwan suggested gaining sample writing models as alternatives for written communication on their drafts. These results are supported by Strauss and Xiang (2006) who explain that "language in collaborative interchange has the capacity to advance, broaden, and clarify our understandings" (p. 359). Unfortunately, some students even criticized the professor's indifference to the lack of feedback or communication and felt that he did not take this issue seriously. Describing his feeling of rejection, Adam said, "I perceived his [the professor's] feedback as a communication between him and myself" (third interview) but "it didn't happen

[...] and he said he could not revise student's drafts."

The students misinterpreted the professor's efforts to communicate with them through the minimal feedback he gave. He did the best he could do in the face of a lack of preparation time but the students rather found it monologic, ineffective, and indifferent. Due to the monologic communication structure of the writing class, which starts from the teacher and ends at the student, the students' misunderstanding was inevitable. Lillis (2003) explains that, whether it is the person's intention or not, if a teacher "denies students' contributions to [...] meaning making" (p. 196) and, then, the teacher holds his position of authority for students posing problems, this turns him into an authoritative "interpreter of the world" (p. 196). The limited communication fed into student perception that the writing instructor was authoritative and indifferent, causing an emotional block between the students and the teacher.

Students' transformations result from dialogism

The dialogic and on-going nature of the interview process affected their understanding of needs assessment and provided a formative learning experience. The following two cases of Christine and Dong-hwan illustrate their transformations throughout the interviews. One of the two participants who were reluctant to talk about their needs in the first interview, Christine in particular kept maintaining her skepticism about discussing her needs until the second interview. Two reasons were found for this reluctance: the monologic nature of needs assessment and the Korean hierarchical relationship between teacher and students. First, Christine raised the issue of the importance of dialogic needs analysis which endorses students' rights to participate in modifying their classes. The following excerpt represents Christine's skepticism when she was asked to discuss her needs about the writing class.

There is no point in talking about what I want from a class, like in a teacher evaluation from students [because] it would be useless to talk about it at the end of a semester [...] nothing would be changed. (Christine, first interview)

Christine considered that her participation would not be appreciated since no efforts were made in changing her previous classes and she regarded needs assessment as a mechanical routine that professors usually conduct at the end of each semester. Due to this fact, she was not enthusiastic about expressing her needs in the interview.

Second, the Korean hierarchical relationship between teacher and student seemed to have a negative impact on her perception about needs assessment. Christine explained that she felt

uncomfortable discussing her wants with a teacher or a professor. She was afraid that if she did, the professor would view her as a “brat” (second interview). Her fear was not based on any of her prior experiences of being rejected or punished for proposing her ideas for a class. However, she found her Korean professors strict and imagined that the professor would not like it if she made suggestions to them.

However, Christine was willing to discuss her needs and wants with her professors as an on-going process, if the professor allowed it. She did not want needs assessment as a onetime event conducted in the early phase of a semester because she generally “had no idea” (third interview) about her needs by then. Likewise, she also was not in favor of the needs assessment made after the class was over because she was not able to benefit from this. “Rather,” Christine suggested that, “I would appreciate it if a professor asked me in the middle of the semester” (third interview) so that she could make suggestions to her professor. Christine assumed that she could change her class through on-going needs assessment, changing her position from discomfort to willingness to sharing her needs. This change appears to be related to the nature of the interview being dialogic and on-going. The dialogic interviews became the opportunity for her to be empowered as an evaluator of her learning and the writing class, which traditionally was the role of the teacher (Freire, 2000). Through these opportunities, she became conscious of her power and the need to have the right to transform the class.

The dialogic nature of the interview process provided a formative learning experience for students. For example, Dong-hwan held a misleading view about being critical in his essay during the second and the first part of the third interview. He explained that writing a critical essay for him meant criticizing unethical issues in the movies he had learned, as his English-Korean dictionary indicated. However, while he was discussing his experience of writing the final draft in the third interview, he refined and corrected his understanding.

Rather than criticizing, it's different from criticizing. There are some issues that many people have taken for granted. But, [a critical essay] tries to think about these again from a different viewpoint. I think that means being critical in essay writing. But, THIS [pointing at his draft] is not a critical essay. I just realized that after I listened to what I was saying. (Dong-hwan, third interview)

Dong-hwan further explained that he was not able to reflect on what he had learned from the professor, because he was preoccupied with finishing his drafts. However, in the dialogic interview, Dong-hwan redefined and corrected his understanding about the meaning of being

critical in his essay writing. These results emphasize that dialogism in needs assessment is not only a research method but also a method in which both teachers and students can learn from each other.

Conclusion and Implications

This study explored the dynamic transformations of three Korean university students' needs in an English composition class as well as the misunderstandings created due to the limited communication, which could have been solved by conversation among the teacher and students. Based on these findings, suggestions to examine needs assessment and writing feedback in a dialogic frame will be discussed as well as pedagogical implications for EFL writing classes.

Revisiting Needs Assessment in an On-going Dialogic Frame

Various researchers (Benesch, 1996, 1999a, 2001; Dudley-Evans, 2001; Flowerdw, 2005; Hyland, 2006, 2007; McLaren, 2009; Molle& Prior, 2008) have emphasized the value of an on-going dialogic needs assessment process that attempts to understand students' needs as variable and situated practice and responds to their immediate learning experience by embracing their voices and modifying a class. The results of the exploration of the transforming needs of the three Korean university students support the importance of conducting this form of dialogic needs assessment.

Student needs are developmental and transformative as students are constantly changing through varied kinds of knowledge from their diverse experiences (Wong, 2006). Even though the students could not clarify what they wanted in the beginning of the class, they became able to articulate their needs through the nurtured understanding about writing in English. In this sense, conducting a single student needs assessment in the form of collecting student needs as quantifiable objects is misguided. The collected needs will be transformed into something different through the varied learning experiences that students will undergo in a writing class.

Valuing the quality of making change through dialogue, the three students pointed out that the starting point of the transformation of the classroom must take place "here and now" (Freire, 2009, p. 59). However, since most needs assessments that they had experienced targeted future students, not themselves, the Korean students assumed that these assessments merely reflected transient interest on the professor's part. Inviting students to participate in class modification through on-going student needs assessment for their present class is more likely to produce meaningful and valuable results.

Last but not least, it should be emphasized that dialogic assessment expands students'

learning experience, providing them with a retrospective and evaluative moment for learning writing. In this sense, dialogism in needs assessment becomes a learning process for both teachers and students.

Reconsidering the Dialogic Functions of Feedback

The students drew attention to the dialogic function of feedback as a critical element in an English composition class. The results revealed that lack of communication between student and student or student and teacher was a serious issue among the participants and it also became the major hindrance in their learning. The students proposed various kinds of solutions to this problem, such as having sample writings, precise written teacher feedback, peer feedback, and casual peer discussion. In a larger framework, they related these feedback activities to having a conversation with others to improve their writing. In other words, the students perceived having feedback as communicating with others to enhance their writing proficiency.

The students suggested three functions of feedback in the light of dialogism. First, the students identified that dialogic feedback with either the teacher or classmates will enable them to learn about their mistakes and errors. Second, feedback can help students understand their teacher's expectations of their assignments. According to Thaiss and Zawacki (2006), university students often rely on teacher feedback to determine the expectations of their teachers. Thus, the frustration Christine and Dong-hwan experienced was inevitable when they were deprived of their opportunity to have a dialogic interaction through feedback (Lillis, 2003). Finally, dialogism in feedback leads students to greater understanding about their writing assignments (Thasis & Zawacki, 2006). The three students desired to benefit even from a casual peer conversation when they were situated in a limited communicative environment. Strauss and Xiang (2006) support the three students' account in that students need to have encounters with either their peers or a teacher in order to understand, expand upon, and express their ideas.

Implications for EFL Writing Classes

The results of the study strongly suggest that conducting on-going dialogic needs assessments will open wider opportunities for teachers to understand and accommodate students' needs. Above all, once they examine their students' dynamic needs, without losing the initial goals, teachers need to negotiate the different expectations with the students to include the students' voices and modify their class. Examples of a series of studies conducted by Benesch (1996, 1999a, 2001) give good pictures of how to practice dialogic needs assessment in a writing class. Also on-going dialogic needs assessment can assist in designing a syllabus or curriculum as

Flowerdew (2005) describes in her study of her English for occupational purposes writing class. Through posing and solving problems about the topics related to their own disciplines, the students affected changes in their academic subject areas which helped them feel empowered. These processes will create a flexible and effective writing course based on what students need here and now.

Writing educators who are suffering from environmental limitations such as large student numbers can consider using varied types of teacher or peer feedback as a means of promoting interaction and student learning. For example, if some technology aids are available, the teacher can share students' writing drafts using PowerPoint slides or in a Word file. This will enable students to see the teacher's expectations of their drafts as well as promoting discussion on writing in English. Sharing writing models can be controversial, however. Considering the case of Dong-hwan, who wanted to copy sample writing, with careful advice and guidance, teachers can make the most of using writing samples to help students understand teachers' expectations for writing assignments.

Small group writing can promote peer feedback and discussion in large classes. With writing groups, teachers can practice various kinds of feedback such as written teacher feedback, individual conferencing in the midst of group work, oral discussion with groups, oral and written pair feedback, whole class discussion, and mini lessons using students' writing drafts. Gebhard (2006) addresses practical issues and solutions that writing teachers can use in their writing classes. For example, "student-to-student writing conference" (p.242) procedures that he suggested can be a good model for teachers with large writing classes.

Adapting a reading and writing workshop (Wong, 2006, p.70-73) may be another effective choice for teachers with 50 students or more. In her original reading and writing workshop, students were allowed to have time to read and write, choose books, give feedback, have dialogue, and create a learning community within highly structure workshops. Teachers may add some of these activities or procedures as a part of their writing instruction. In this case, teachers can consider various group sizes for group work and different types of group activities at the same time to promote student feedback and dialogue.

When teachers and students have easy access to internet connected computers, web-based social networking platforms can provide the learning space through group writing and delayed peer and teacher feedback. For instance, wikis promote student-teacher communication and peer learning (Bubas, Kovacic, & Zlatovic, 2007; Zorko, 2007) and collaborative learning in particular can enhance writing competencies and interest in language learning (Wang, 2014). Thus, a blended writing course utilizing web-based networking platforms will assist teachers

who want to practice dialogic needs assessment in their writing classes.

The results present two more valuable suggestions for administrators of higher education in EFL countries. Most of all, smaller class sizes for English composition courses will promote more sophisticated teacher feedback and active peer discussion that leads to a better quality of university student writing in return. Limited communication with the professor and classmates about their writing practice turned out to be the critical obstacle to their literacy development in English. Running smaller class sizes can solve many problems. Also, various types of English writing classes should be provided to accommodate students' different language proficiencies and writing goals. Participants with limited experience in learning writing in both Korean and English require different teaching approaches and content compared to more advanced student writers.

In making these suggestions, it is important to note that I am well aware that time and economic constraints might influence the possibility of them being able to be implemented in EFL classrooms. However, these suggestions call for serious attention to be given to the development of university students' writing proficiency in EFL countries. I hope the present study will encourage more research to investigate understanding students' needs in a dialogic framework and cultivate insightful discussions about designing effective and meaningful English writing composition courses.

References

- Angrosino, M. (2007). *Doing ethnographic and observational research*. Thousand Oaks, CA: Sage.
- Ashwell, T. (2000). Patterns of teacher response to student writing in a multi-draft composition classroom: Is content feedback followed by form feedback the best method? *Journal of Second Language Writing*, 9(3), 227-257.
- Bakhtin, M. M. (1982). *The dialogic imagination by M. M. Bakhtin: Four essays*. Austin, TX: University of Texas Press.
- Benesch, S. (1996). Needs analysis and curriculum development in EAP: An example of critical approach. *TESOL Quarterly*, 30(4), 723-738.
- Benesch, S. (1999a). Rights analysis: Studying power relations in an academic setting. *English for Specific Purposes*, 18, 313-327.
- Benesch, S. (1999b). Thinking critically, thinking dialogically. *TESOL Quarterly*, 33, 573-580.
- Benesch, S. (2001). *Critical English for academic purposes: Theory, politics, and practice*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Bubas, G., Kovacic, A., & Zlatovic, M. (2007). Evaluation of activity with a wiki system in teaching English as a second language. Proceedings of the International Conference ICT for Language Learning (pp. 201-21). Florence, Italy.
- Chambers, F. (1980). A re-evaluation of needs analysis. *ESP Journal*, 1(1), 25-33.
- Carless, D. (2006). Differing perceptions in the feedback process. *Studies in Higher Education*, 31(2), 219-233.
- Crang, M. & Cook, I. (2007). *Doing ethnographies*. Thousand Oaks, CA: Sage.
- Denton, P., Madden, J., Roberts, M., & Rowe, P. (2008). Students' response to traditional and computer-assisted formative feedback: A comparative case study. *British Journal of Educational Technology*, 39(3), 486-500.
- Dudley-Evans, T. (2001). English for specific purposes. In R. Carter & D. Nunan (Eds.), *The Cambridge guide to teaching English to speakers of other languages* (pp.131-136). Cambridge, England: Cambridge University Press.
- Dudley-Evans, T. & St. John, M. (1998). *Developments in ESP: A multi-disciplinary approach*. Cambridge, England: Cambridge University Press.
- Flowerdew, L. (2005). Integrating traditional and critical approaches to syllabus design: The 'what,' the 'how' and the 'why'? *Journal of English for Academic Purposes*, 4, 135-147.
- Freire, P. (2000). *Pedagogy of the oppressed*. (M. B. Ramos, Trans.). New York, NY: Continuum International. (Original work published 1974).

- Freire, P. (2009). From pedagogy of the oppressed. In A. Darder, M. P. Baltodano & R. D. Torres (Eds.), *Critical pedagogy reader* (2nd ed., pp. 52-60). New York, NY: Routledge.
- Gebhard, J. G. (2006). *Teaching English as a foreign or second language: A self-development and methodology guide* (2nd ed.). Ann Arbor: The University of Michigan press.
- Holliday, A. (1994). *Appropriate methodology and social context*. Cambridge, England: Cambridge University Press.
- Holliday, A. & Cooke, T. (1982). An ecological approach to ESP. *Lancaster Practical Papers in English Language Education*, 5, 123-143.
- Hutchinson, T. & Waters, A. (1987). *English for specific purposes*. Cambridge, England: Cambridge University Press.
- Hyland, K. (2000). *Disciplinary discourses: social interactions in academic writing*. London: Longman.
- Hyland, K. (2006). *English for academic purposes: An advanced coursebook*. London, England: Routledge.
- Hyland, K. (2007). Genre pedagogy: Language, literacy and L2 writing instruction. *Journal of Second Language Writing*, 16, 148-164.
- Hyland, K. & Hyland, F. (2006). Feedback on second language students' writing, *Language Teaching*, 39(2), 83-101.
- Ivanic, R., Clark, R., & Rimmershaw, R. (2000). What am I supposed to make of this? The messages conveyed to students by tutors' written comments. In M. Lea and B. Stierer (Eds.), *Student writing in higher education: New contexts*, (pp. 28-39). Buckingham: Open University Press.
- Johns, A. M. (1990). Coherence as a cultural phenomenon: Employing ethnographic principles in the academic milieu. In U. Connor & A. M. Johns (Eds.), *Coherence in writing: Research and pedagogical perspectives* (pp. 209-226). Alexandria, VA: TESOL.
- Kachru, B. (1992). World Englishes: approaches, issues and resources. *Language Teaching*, 25(1), 1 – 14.
- Killoran, J. B. (2005). ePluribus unum? Dialogism and monologism in organizational web discourse. *Journal of Technical Writing and Communication*, 35(2), 129-153.
- Lee, Y. (2006). We know what the useful thing is: An examination of students' perceptions on classroom-based teaching for EFL writing. *Foreign Language Education*, 13(2), 309-330.
- Lea, M. & Street, B. (1998). Student writing in higher education: An academic literacies approach. *Studies in Higher Education*, 23(2), 157-172.

- Lillis, T. (2003). Student writing as 'academic literacies': Drawing on Bakhtin to move from critique to design. *Language and Education*, 17(3), 192-207.
- Lizzio, A. & Wilson, K. (2008). Feedback on assessment: Students' perceptions on quality and effectiveness. *Assessment & Evaluation in Higher Education*, 33(3), 263-275.
- Maxwell, J. A. (2005). *Qualitative research design: An interactive approach* (2nd ed.). Thousand Oaks, CA: Sage.
- McGarrell, H. & Verbeem, J. (2007). Motivating revision of drafts through formative feedback. *ELT Journal*, 61(3), 228-236.
- McLaren, P. (2009). A look at the major concepts. In A. Darder, M. P. Baltodano & R. D. Torres (Eds.), *Critical pedagogy reader* (2nd ed., pp. 61-83). New York, NY: Routledge.
- Molle, D. & Prior, P. (2008). Multimodal genre systems in EAP writing pedagogy: Reflecting on a needs analysis. *TESOL Quarterly*, 42(4), 541-566.
- Munby, J. (1978). *Communicative syllabus design*. Cambridge, England: Cambridge University Press.
- Mutch, A. (2003). Exploring the practice of feedback to students. *Active Learning in Higher Education*. 4(2), 24-38.
- Park, C. W. (2004). From qualitative to quantitative continuum: Exploring the possibility with the aid of CAQDAS (NVivo 2). *English Teaching*, 59(4), 71-91.
- Park, C. W. (2009). *Systematic management and analysis of qualitative data for the field researchers: NVivo 8 Applications*. Seoul: Hyungsul.
- Prior, P. (1991). Contextualizing writing and response in a graduate seminar. *Written Communication*, 8, 267-310.
- Prior, P. (1995). Redefining the task: An ethnographic examination of writing and response in graduate seminars. In D. Belcher & G. Braine (Eds.), *Academic writing in a second language: Essays on research and pedagogy* (pp. 47-82). Norwood, NJ: Ablex.
- Ramani, E., Chacko, T., Singh, S. J., & Glendinning, E. H. (1998). An ethnographic approach to syllabus design: A case study of the Indian Institute of Science, Bangalore. *English for Specific Purposes*, 7, 81-90.
- Richterich, R. & Chancerel, J. L. (1977/ 1980). *Identifying the needs of adults learning a foreign language*. Strasbourg, France: Council of Europe.
- Shamim, F. (1996). In or out of the action zone: Location as a feature of interaction in large ESL classes in Pakistan. In K. M. Bailey & D. Nunan (Eds.), *Voices from the language classroom* (pp. 123-144). Cambridge, England: Cambridge University Press.
- Songhori, M. H. (2008). Introduction to needs analysis. *English for Specific Purposes World*, 4,

1-25.

- Spradley, J. P. (1979). *The ethnographic interview*. London: Holt, Rinehart & Winston.
- Street, B., Baker, D., & Tomlin, A. (2005). *Navigating numeracies: Home/school numeracy practices*. Leverhulme Numeracy Research Programme. Dordrecht, Netherlands: Kluwen.
- Strauss, S. & Xiang, X. (2006). The writing conference as a locus of emergent agency. *Written Communication*, 23(4), 355-396.
- Thaiss, C. J. & Zawacki, T. M. (2006). *Engaged writers and dynamic disciplines: Research on the academic writing life*. Portsmouth, England: Boynton/Cook.
- Walker, M. (2009). An investigation into written comments on assignments: Do students find them usable? *Assessment & Evaluation in Higher Education*, 34(1), 67-78.
- Wang, Y. C. (2014) Promoting collaborative writing through wikis: a new approach for advancing innovative and active learning in an ESP context. *Computer Assisted Language Learning*, (ahead-of-print) , 1-14.
- West, R. (1994). Needs analysis in language teaching. *Language Teaching*, 27(1), 1-19.
- Wingate, U. (2010). The impact of formative feedback on the development of academic writing. *Assessment & Evaluation in Higher Education*. 35(5), 519-533.
- Wong, S. (2006). *Dialogic approaches to TESOL: Where the ginkgo tree grows*. New York, NY: Lawrence Erlbaum.
- Zorko, V. (2007). A rationale for introducing a wiki and a blog in a blended-learning context. *CALL-EJ online*, 8(2).

Evaluating genre-based writing instruction: Materials, instructional mode and student learning styles

Diane Johnson

The University of Waikato, New Zealand

Hsiu-Chen (Antonia) Lin

Wenzao Ursuline University of Languages, Kaohsiung, Taiwan ROC

Bio data

Dr Diane Johnson is a senior lecturer in applied linguistics at Waikato University in New Zealand. She conducts research and publishes in the areas of discourse analysis, language teaching methodology, curriculum and syllabus design and language revitalisation.

Dr Hsiu-chen Antonia Lin is an associate professor in the Department of English at Wenzao Ursuline University of Languages in Kaohsiung, Taiwan. She received her PhD in Applied Linguistics from the University of Waikato in New Zealand. Her research interests lie in genre-based writing instruction and computer applications in ELT.

Abstract

The study reported on here was undertaken in Taiwan with a group of intermediate-level tertiary students of English. At its core, the project sought to measure the efficacy of a purpose-designed, genre-based writing course but, in order to make the data collection and reporting more nuanced, the project also sought to measure any potential impact on engagement and overall writing development that the mode of delivery and the students' learning styles might have. The content of the writing course was based on 'cognitive' genres (e.g. *argument*, *recount* etc.) and included instruction in the overall rhetorical structure, internal discourse structure and typical language features of a number of model texts. The course lasted for 10 days (50 hours) and involved 28 participants divided into three groups who were taught using one of three delivery modes—fully online, face-to-face or blended.

Address correspondence:

Te Whare Wananga o Waikato
Gate 1 Knighton Road, Private
Bag 3105, Hamilton 3240, New
Zealand

The participants completed both a pre-test and post-test involving a genre-centered writing task which was graded using a criterion-referenced approach. Participants also

completed a pre-course learning styles inventory and, following the course, responded to a course questionnaire. A number of students also participated in a follow-up focus-group discussion. A comparison of pre-test and post-test writing task scores indicated improvement in all assessed areas and questionnaire responses and informal feedback indicated a high level of satisfaction with the course. Face-to-face and blended mode group participants were the most positive about the course but the writing performance of the online mode group showed the greatest improvement. There was no detectable relationship between learning styles and learning mode preferences.

Keywords: EFL academic writing, computer-mediated writing instruction, genre-centered writing instruction, cognitive genre

Introduction

In the design, development and delivery of teaching and learning materials, the choice of instructional mode, as Sadler-Smith and Riding (1999) suggest, is now being acknowledged as an increasingly significant consideration. (See also Gagne & Briggs 1979; Rowntree 1982; Reigeluth 1983; Gough 1996; Hayes & Allison 1996). This consideration, alongside a greater focus on the potential effects of individual learning and cognitive styles on performance (see for example Cronbach & Snow 1977; Entwistle 1988; Schmeck 1988; Riding & Cheema 1991), suggests that in developing a writing programme, successful outcomes for students may depend not just on the overall orientation and quality of the course materials but also on the selection of an appropriate mode of delivery and on the potential impact of the learning style of the students. In order to test the potential interplay of these three factors on engagement and writing development, this study, conducted in Taiwan, involved a writing course offered on a voluntary basis over a ten-day period (50 hours) to students studying English as a major or minor subject who were attending a tertiary-level educational institution. Underlying the study was the following research question:

In terms of attitudes and performance, how do groups of intermediate level students of English at tertiary level respond to a genre-centered writing course delivered in three different modes (face-to-face; fully online; blended) and is there any evidence that different learning styles have any relationship with learning mode preferences and/or writing performance?

All of the participants completed a learning styles questionnaire, a course response questionnaire and a pre-test and post-test involving genre-centered writing. The grading of pre-test and post-test writing tasks was criterion-referenced with the criteria relating directly to course objectives. A number of the students also participated in focus group discussions following the course and others gave informal feedback through email correspondence.

Selecting an approach to the teaching of writing: A critical review of selected literature

It is commonly assumed that the teaching of writing in the early 20th century, particularly in the 1940s and 1950s, was ‘product-oriented’, with an emphasis, irrespective of topic, on the production of five paragraph essays (Nystrand, Greene, & Wiemelt, 1993, p. 275). This has been labelled as ‘current-traditional rhetoric’: a term first was used by Fogarty (1959) and

subsequently popularized by Young (1978, p. 31). However, as Miller (1991, p. 110) has indicated, this label was actually created at the same time as what is sometimes referred to as ‘process theory’ as a way of promoting the latter by presenting it as being “pitted against old practices”. Thus, it involved the discursive creation of “a daemon for the sake of expelling it” (Pullman, 1999, p. 23), a “caricature against which the process movement developed” (Matsuda, 2003, p. 671).

The concept of ‘a process approach’ creates “a misleading image of unity and coherence” (Tobin, 1994, p.4). As Bizzell (1986) and Faigley (1986) have observed, a multiplicity of different approaches have all claimed to be process-oriented. A similar point could be made about the use of the term ‘post process’ in the context of writing pedagogy and about using the singular (the genre ‘approach’) when referring to what has been described as “the main institutionalized alternative to process pedagogy currently on offer” (Atkinson, 2003, p. 11). As Matsuda (2003, p. 73) has observed, “the post-process movement does not represent a unified theoretical front”. Application of the term ‘post process’ in the context of writing pedagogy is actually little more than an “heuristic for expanding the scope of the field of second language writing” and “needs to be understood not as the rejection of process but as the recognition of the multiplicity of L2 writing theories and pedagogies” (p. 65).

Focusing on processes

The emphasis on processes in the teaching of writing that is often now evident in many first and second language contexts (L1 & L2) is often traced back to a work by Emig, first published in 1971. In fact, however, as observed by Nystrand (2006), there is evidence of a similar type of emphasis in a number of studies in the area of writing that were reviewed in the early 1960s by Braddock, Lloyd-Jones and Schoer (1963). Studies such as these had a considerable influence on the establishment of composition and rhetoric courses in United States universities. However, they would have been likely to have had less impact had it not been for an Anglo-American seminar held at Dartmouth in New Hampshire in 1966. That seminar aimed to improve the teaching of English through collaboration among scholars in different countries, particularly the USA, the UK and Canada. It led to the publication of *Growth through English: A report based on the Dartmouth seminar* by John Dixon in 1967. In that book, Dixon stressed that language is learned *through the experience of using it*. This simple observation was to have a major influence on the ways in which English was taught in schools and colleges. So far as writing is concerned, it led to a reduction of emphasis on mechanical aspects of writing (such as punctuation) and sentence level grammar and an

increased emphasis on attempts to replicate the processes thought to be involved in writing. All of this was part of a more general shift in educational philosophy, a shift towards learner-centered education, which inevitably, over time, has had an impact on the teaching of second/foreign languages.

At the heart of process-centered approaches to the teaching of writing is the perception of writing as problem solving (Flower & Hayes, 1981, p. 370; Hayes & Flower, 1980, p. 3) or, in the words of Odell (1980, p. 140), as “a process of discovery, a process of exploring . . . of creating, testing, and refining hypotheses”. In such a context, the teacher does not dominate but provides, along with students’ peers, feedback and a sense of audience (Tangpermpoon, 2008, p. 5). Within this context, writing has generally been seen as an activity requiring an encouraging, positive, and cooperative environment, one in which there should be minimal interference (Emig, 1983; Gould, 1980; Odell, 1980; Raimes, 1983; 1985; Zamel, 1983). Particularly in the late 1970s and 1980s, emphasis was often placed on the students’ search for an authentic voice (Tobin, 2001, p.4). Indeed, process-centered approaches have often been associated with what Elbow (1973) called ‘free writing’.

Typically, writing came to be conceptualized as involving a number of (often recursive) stages: *prewriting*, *drafting*, *revising*, *editing* and *publishing*. It has sometimes been claimed that the processes involved are “cognitive or internal” (Atkinson, 2003, p. 10). However, Susser (1993, p. 33) has noted that what is generally conceived of as a process is actually a set of pedagogical practices (see also North (1987)). Furthermore, although Matsuda (2003, p. 69) has noted that process-centered approaches are often described as “the most successful . . . pedagogical reform in the teaching of writing”, there is actually little hard evidence that process-writing techniques “lead to significantly better writing” (Hyland, 2002, p. 29).

By the 1980s, process-centred approaches had come to “dominate the professional literature on the teaching of writing” (Applebee, 1986, p.97) and had begun to “serve as a kind of disciplinary shibboleth” (Tobin, 1994, p. 7). They were initially largely confined to L1 contexts (Caudery, 1995, ¶1; Gao, 2007, ¶8) although Johns (1990, p.26) has observed that their influence in the L2 context “cannot be exaggerated”. Even so, process-centered approaches have generally been integrated with other types of approach in the L2 context.

Criticism of process-centered approaches has tended to focus on the following issues: the fact that the needs of L1 and L2 writers may be different; the lack of a socially-oriented perspective; and the lack of explicit instruction in specific aspects of writing, including discourse organization and linguistic selection.

Pennington and So (1993, p. 58) have insisted that writing processes are similar in the case of L1 and L2 (see also Zamel (1983); Cambourne (1988)). There are, however, researchers who have argued that this is not the case (see, for example, Arndt, 1987; Wolff, 2000). Whatever similarities there may be, one difference that must be acknowledged relates to “the constraints imposed by imperfect knowledge of the language code involved” in the case of novice L2 writers (Caudery, 1995, ¶41; Badger & White, 2000, p. 15).

Process-centered approaches to the teaching of writing have also been criticized for lack of a social perspective, something that “leaves students innocent of the valued ways of acting and being in society” and lacking “the cultural and linguistic resources necessary for them to engage critically with texts” (Hyland, 2003, pp. 18, 20 & 25). Knapp and Watkins (2005, pp. 8 & 14) have argued that in treating language learning largely as “an individualized phenomenon” and in under-valuing instruction in textual form, process-centered approaches result in students failing to develop an adequate understanding of contextualized language knowledge. In the absence of adequate lexical, grammatical and text construction skills (Hinkel, 2004, p. 7), and with “no recognizable discourse structure to speak of” (Atkinson & Ramanathan, 1995, p. 564), learners may be evaluated negatively in academic and employment contexts (Hinkel, 2004, p. 124).

Focusing on genres

Genre-centered approaches to the teaching of writing are generally considered within the context of what is often now referred to as ‘post-process’ pedagogies, pedagogies that have been described by Matsuda (2003, pp. 78-79) as involving “the rejection of the dominance of process at the expense of other aspects of writing and writing instruction”. In order to fully understand the motivation for the shift in emphasis that led to the introduction of the term ‘post-process’, it is important to bear in mind the overall educational climate in which process-centered approaches emerged and thrived and the ways in which that climate has changed over time.

Many of those who attended the 1966 Dartmouth Conference referred to above seem to have been ready to accept pockets of research that purported to demonstrate the specific instruction in language could be positively harmful (see, for example, Harris, 1962). Much of that research was based on the teaching of decontextualized traditional, Latin-based grammar. Nevertheless, there was a considerable period during which specific instruction in language was largely removed from the L1 curriculum. So far as L2 teaching is concerned, the broader context was one in which the demise of full-blown structuralism led to a move away from a

focus on sentence grammar and towards a focus on ‘communicative competencies’ and ‘communicative language teaching’. In its most extreme form, this led initially to a total rejection of instruction in language (see, for example, the discussion of this in Celce-Murcia, Dornyei, & Thurrell, 1997). However, this was strenuously resisted.

There was initially much less resistance to the disappearance of instruction in language in the L1 context. The situation has now changed, with a vociferous rejection in many countries of what is often referred to as the ‘whole language’ movement’ (Adams, 1991). Thus, for example, the New Zealand English Curriculum reintroduced the teaching of grammar into New Zealand classrooms in the early 1990s (MoE, 1994), and in England, the National Literacy Strategy (NLS) included a focus on systemic-functional grammar (DfEE, 1998). At the same time, increasing criticism of process-centered approaches to the teaching of writing in both L1 and L2 contexts often involved a rejection of the “inherent liberal individualism” associated with such approaches (Hyland, 2003, p. 17).

It is in the context outlined above that educationalists began to explore the possibility of incorporating insights from various different strands of research on genre into the teaching of writing. In some cases, particularly in North America, the emphasis has been on what are referred to below as ‘social genres’; in others, particularly in Australia and New Zealand, the emphasis has been on what are referred to below as ‘cognitive genres’.

Use of the terms ‘genre’ and ‘text-type’ vary widely. In this project, following Bruce (2003, pp. 4-5), the authors refer to constructs that are defined in terms of their overall social purpose (e.g. *novels, academic articles*) as ‘social genres’ and to constructs that are defined in terms of an overarching rhetorical function (e.g. *argument; explanation*) as ‘cognitive genres’. Although the latter are sometimes referred to as ‘elemental genres’ (Hyland, 2007, p. 153), the authors prefer the term ‘cognitive genres’ because it emphasizes the important role that cognitive processes (e.g. *associative, logico-deductive* and *temporal* processes) and the textual relationships that are linked to them (e.g. *amplification, means-purpose* and *temporal sequence*) play in differentiating them. A particular instance of a social genre (e.g. *a personal letter*) may exhibit a range of different cognitive genres although some social genres (e.g. *car manuals*) are characterized by a tendency to exhibit a single cognitive genre (Crombie & Johnson, 2004, p. 144). Research relating to both social genres (often associated with English for Specific Purposes (e.g. Swales (1990))) and cognitive genres (e.g. Bazermann, 1994) can play a role in the teaching of writing. As Bhatia (1998, pp. 26-27) observes:

We need the sophistication and subtleties of ESP but at the same time we need to master the power of generalizations across disciplinary boundaries. . . . [In] order to deal with the complexity of generic patterns so commonly intertwined in academic discourse across disciplines, one needs a system of linguistic analysis which is powerful enough to account for the intricacies of academic genres across disciplines.

Approaches to genre that are cognitive in orientation can be applied in a cross-disciplinary way. It is these with which the authors were primarily concerned in the design of this writing course.

Method

The participants

Although 59 students initially signed up for the course (which was not credit bearing), only 30 registered on the first day and only 28 (25 female and 3 male) completed it. The age range of the participants was from 17 to 31. Most of the participants had a score of between 180 and 240 (one student had not taken the test and a few had scored slightly higher) in the College Students' English Proficiency Test (CSEPT), an English language proficiency test involving listening, reading, and grammar use with a total possible score of 360. This represents a very wide proficiency range, being roughly equivalent to anywhere between 3.5 and 5.5 in the IELTS test (Crombie & Johnson, 2009b, p. 12). The participants came from different areas of a single institution: five-year junior college (8); two-year college (6); four-year college (11); four-year evening college (3). All of them, including those assigned to the online group, were expected to work in classrooms and/or computer rooms from 9.00am to 3.00pm over a period of 10 days. The total time devoted to the course itself was 50 hours (with the additional time being spent on administrative matters, pre- and post-tests, etc.).

The participants had no choice in terms of the learning mode to which they were assigned. In general terms, they were placed in groups in relation to their year of study and major subject area so that they would be with other students they were already likely to know and the class dynamic would be easier to establish in a relatively short period of time. Allocation to delivery mode groups also took account of the need to have students of varying proficiency in each group and to have as wide a variety of learning profiles as possible. There were 9 students in Group O (fully online), 10 in Group B (blended), and 9 in Group F (face-to-face). The learning styles profiles of each group are illustrated in *Table 1* below

Table1

Learning style profiles

| Groups | Learning Styles | | | | | |
|-------------------------|-----------------|----------|----------|------|------|----------|
| Online (O) | ENFP (3) | ENFJ (2) | ISTJ (2) | ISFJ | INFJ | |
| Blended (B) | ENFP (2) | ENFJ | ISFP (2) | ESFJ | INFJ | ESTJ (3) |
| Face-to-face (F) | ENFP | ENFJ | ISTJ | ESFJ | INFP | ESFP |
| | ENTP | ENTJ | ISTP | | | |

Those in Group B spent most of their time working online but had one hour of face-to-face tuition each day, during which time the emphasis was on providing opportunities for discussion of the online materials to which they had been introduced. In some cases, supplementary materials designed to clarify and/ or exemplify aspects of the online materials were also used.

The learning styles inventory

With the permission of Dr John Shindler of California State University, the 52-item Paragon Learning Style Inventory (PLSI) was translated into Chinese and used as part of this study. For further information about the PLSI, see Shindler & Yang, 2004. On the basis of responses to the questions included in the inventory, each participant was assigned a learning style profile made up of a combination of four characteristics/ tendencies: introversion or extroversion (I/E); sensation or intuition (S/N), thinking or feeling (T/F), and judgment or perception (J/P). The first two relate to inner-centred/ ideas-centred (I) versus people-centred (E) tendencies; the next two relate to the extent to which ideas are approached in a holistic (S) or detailed (N) way; the third pair relate to a tendency towards basing decisions on logic (T) or their likely impact on others (F); the final pair relate to the extent to which judgments tend to be based on facts (J) or on context-related interpretation (P). These profiles would be used at the conclusion of the course as a tool to measure any potential correlation between learning and cognitive style, reaction to delivery mode and issues of overall engagement and success.

The instructors

Three experienced language teachers, two of whom were also involved in language teacher training, were involved in the study. Two of them were located in Taiwan during the course: one taught Group F (the face-to-face mode group); the other monitored Group O (the online mode group) and monitored and taught Group B (the blended mode group). The third teacher, located in New Zealand for the duration of the study, graded and commented on all of the in-

class writing assignments done during the course which were received as email attachments in the evening and returned with feedback to the students by the following morning.

The learning platform and the face-to-face mode materials

The learning platform (E-learning) used in this study was locally developed and familiar to the participants. Nevertheless, they were all given an introduction to the environment of the learning at the beginning of the course to be sure that learning time was not impeded by potential problems with computers or difficulties with navigating the learning platform. A set of paper-based materials (relating directly to the online materials) was prepared for use by the blended mode and face-to-face mode teacher. The handouts and PowerPoint presentations which were part of the face-to-face and blended modes included model texts and summaries of the main teaching points.

The writing pre-test and post-test

Lin (2006) advises that, so far as language is concerned, writing assessment should focus only on those language features that are highlighted during a course. However, for the purposes of this study, it was resolved that much could also be gained from being able to compare student scores that related specifically to what was taught, with student scores that related to more general criteria that are often applied in writing courses. Therefore, a two part criterion-referenced analysis/grading scheme was devised for use in relation to the pre- and post-test writing tasks. *Part A* related specifically to what was taught (see *Appendix 1 Table 1a*); *Part B* was more general in nature (see *Appendix 1 Table 1b*). The grading process was trialled and validated as part of a pilot study conducted in New Zealand prior to the Taiwanese study. In both the writing pre-test and post-test, participants were given 100 minutes to compose two texts (exhibiting two different cognitive genres) of approximately 250 words each. In order to ensure that there was no opportunity for plagiarism, participants were not provided with access to computers during the pre-test and post-test.

The texts produced were graded according to criteria that related directly to the objectives of the course:

- overall text structuring (up to 10 points);
- appropriate paragraphing and paragraph linkage within text segments (up to 10 points);
- use of discourse relations appropriate to the genre (up to 10 points);
- accurate and appropriate use of discourse relational signalling (up to 10 points); and

- accurate and appropriate use of other language features characteristic of the genre (up to 10 points).

The pre-test and post-test writing tasks were graded independently by two of the teachers involved in the study and spot checked by the third.

The course materials

The course materials were adapted from a book focusing on cognitive genres that was designed to teach academic writing to learners of English in Taiwan (Crombie & Johnson, 2009a). Although five main cognitive genres (*instruction*; *argument* (one-sided and two-sided), *description & classification*; *recount*; and *explanation*) are included in the book, time constraints meant that one of them (*explanation*) was not included in the course. The four included in the course were selected because, although differently labelled, they are consistent with those ‘elementary genres’ proposed by Quinn (1993, pp. 34-35) as a basis for instruction in English for Academic Purposes (EAP) and are related to those ‘text types’ identified by Biber (1989, pp. 29, 31 & 38) as being most common in academic prose (that is, *learned scientific exposition*, *learned exposition*, *involved persuasion* and *general narrative exposition*). The other one, *instruction*, was included because the writers’ experience in Taiwan indicated that it is a genre that is rarely taught in writing classes there.

A basic template for the overall rhetorical structuring of texts (*topic*, *focus*, *detail*, *conclusion*) was adapted in relation to each genre. Thus, for example, in the case of *instruction*, the *topic* section included the goal or goals (e.g. *how to use Skype*) the *focus* section included the materials and equipment needed, the *detail* section included warnings (optional) and the steps or stages involved, and the *conclusion* section included advice (optional) and/ or comment.

So far as the internal discourse structuring of texts is concerned, discourse relations and their realization and signaling played a central role. Thus, for example, in the section dealing with instruction texts, the relations in focus were *Reason-Result*, *Means-Purpose* and *Temporal Sequence*. In the case of *Reason-Result*, the realization focus was a combination of declarative (reason) and imperative (result):

Camera lenses are very delicate and easily damaged (REASON);

DO NOT clean your lens more often than is strictly necessary (RESULT).

This is a particular pragmatic variety of Reason-Result in which the result is an instruction.

In the case of *Means-Purpose*, the focus was on a combination of infinitive (purpose) and the imperative (means):

To keep your camera lens clean (PURPOSE),
always use your lens cover when you are not using your camera and always avoid touching the lens when you are taking photographs (MEANS).

It should be noted here that, as in the case of the result member of the *Reason-Result* relation, there is an imperative in the means member of the *Means-Purpose* relation here (functioning as an instruction) so that there is an overall focus on the use of imperatives in instructions.

In the case of *Temporal Sequence*, the realization focus was sentence initial conjuncts:

First, blow . . . Next, apply . . . Finally, dry . . .

By the time they had completed the course, the participants had been introduced to many discourse relations and to a range of different ways of encoding and signaling them. In addition, they had focused on other aspects of language in context. This included, for example, the use of the present simple tense to refer to general truths (*description and classification*), various types of conditional construction in the context of past time (e.g. *If she borrowed . . . , she would have . . .*), and present and/ or past continuous in initial position (*recount*).

Also included in the course were a range of model texts and writing tasks. In many cases, these were the same as those included in the source textbook. However, a number of changes were made in order to (a) make the course more directly relevant to the backgrounds, needs and interests of the participants, (b) take full advantage of the capabilities of the learning platform (which allowed for a more varied form of presentation than is possible in a printed textbook), and (c) ensure that the course could be completed in the time available.

The writing exercises included were of a variety of types. Thus, for example, participants might be provided with the title and introductory section of a text about the uses of personal computers and asked to write the detail (argument) and conclusion sections or they might be asked to make use of a text template and notes (not presented in any particular order) in responding to a question about the reasons why used computers should not be dumped along with other household items.

In genre-centered writing courses, joint construction of texts is often followed by individual text construction (see, for example, Derewianka, 1994). In this course, the joint

construction phase was replaced by a gradual unfolding of model texts (in sections), along with discussion of the principles guiding the construction and the content of each text section.

Findings

Pre- and post-test writing task scores

Table 2 outlines the scores of each group in the pre-test and post-test writing tasks.

Table 2

Writing pre-test and post-test scores by group

| Mode | Online | | Blended | | Face-to-face | |
|---|-------------|------|---------------|------|--------------|------|
| | Pre | Post | Pre | Post | Pre | Post |
| Writing test | | | | | | |
| Generic structure | 60.5% | 89% | 17.5% | 54% | 59% | 89% |
| TOTAL GAIN | +29% | | +36.5% | | +30% | |
| Discourse relations and discourse relational signalling | 51% | 72% | 75% | 84% | 65% | 83% |
| TOTAL GAIN | +21% | | +9% | | +18% | |
| Other aspects of language associated with the genre | 48% | 69% | 59% | 65% | 57% | 65% |
| TOTAL GAIN | +21% | | +6% | | +8% | |
| Increase in all 3 areas | 54% | 78% | 49% | 68% | 61% | 82% |
| TOTAL GAIN | +24% | | +19% | | +21% | |

In terms of overall performance (in relation to average improvement in post-test scores over pre-test scores), Group O (+24%) outperformed Group F (+21%) and Group B (+19%). This is also the case for both discourse relations and their signalling and other aspects of the language associated with the genre (where the rank order of the groups remains the same). However, in the case of generic structure, the increase is greatest in the case of Group B (+36.5%) followed by Group F (+30%) and Group O (+29%).

Relationship between learning style profiles and improvement in writing scores

The learning style profiles of the participants are outlined in *Table 1* above. Student learning style inventory profiles were examined in relation to student learning mode preferences (as indicated in the questionnaire responses), learning mode assignment and pre-test and post-test results. When the pre-test and post-test scores were examined in relation to the learning style profiles, it was found that the overall improvement was slightly greater in the case of those with I (introversion) in their profile (+21.6% as compared with +19%). Apart from this slight difference, there was no discernible relationship between mode of delivery and overall writing improvement.

Course questionnaire responses

Twenty-eight (28) fully or partially completed questionnaires were collected and the responses were analyzed. Overall, the students' responses were very positive. Data from the questionnaires are attached in *Appendix 2*.

Focus group discussion and email comments

All of the course participants sent comments on the course by email and ten of them accepted an invitation to participate in a focus group discussion (approximately 90 minutes) at the end of the course. Both the positive response to the course and the overall preference for face-to-face and blended modes were reinforced by participants in the focus group and in email comments. In addition, some of the participants noted that they believed that the course had led to an improvement in reading as well as writing. Those in Groups F and B also commented on the fact that they believed that there had also been an improvement in their listening and speaking skills. The following comments are included to provide some flavour of the responses. Those that are in italics have been translated from Chinese.

I love this course very much because the teaching materials are detailed and clear. I can learn it easily.

I am so glad I can join this course. Because of this course, I have learned how to write an article with logic.

I really learned a lot in this writing course. . . . [It] taught me different writing types step by step so I could understand the content easily. . . . Before this course, the teacher of writing lessons always gave me some description and asked me to create a paragraph.

Before I attended the course, I only knew writing an article should include introduction, body, and conclusion. . . . I have learned not only writing skill but also listening skill and speaking.

Being someone who lacks self-control and concentration, I would rather have been placed in a face-to-face group. I still learned a lot but I would have made more progress if I had been placed in face-to-face mode.

I learned a lot in the course because I had interaction with other students in person and also communicated with them online. I was able to ask questions and accomplish my daily work on time. The face-to-face session gave me a chance to practice my listening skills.

The course was very well organized and everything was done in sequence.

The different text templates are quite helpful to follow. The model texts are also useful in our own writing. They are also practical in helping us with structural principles for example.

In traditional classes I don't pay much attention but this intensive course was really unusual and I have gained something very valuable from it. If we have lessons like this in other courses, it will really help me to improve my writing ability. I feel that I have gained a lot from this kind of learning.

Discussion

Engagement

In questionnaire responses, email comments and focus group discussion, the participants expressed a high level of satisfaction with the course. All except two indicated that they 'liked the course' or 'liked it a lot', the exceptions being one member of the face-to-face mode group and one member of the online mode group (who selected 'so-so'). Participants in the blended mode group (70%) and the face-to-face mode group (67%) were more likely to select 'I liked it a lot' than were participants in the online mode group (33%). When asked whether they would like to do a course of a similar type in the future, 4 indicated that they would not, 3 (33%) in the online mode group and 1 (11%) in the face-to-face mode group. However, two of those in the online mode group who indicated that they would not like to do a similar course in the future added a note to the effect that they would if they could join a face-to-face mode group. Those in the blended mode group (80%) and the face-to-face mode group (78%) were more likely than those in the online mode group (56%) to select 'I would very much like to do such a course'.

Perceptions of the usefulness of the course

When asked about the overall usefulness of the course, all of the participants selected ‘very useful’ or ‘useful’ (rather than ‘so-so’ or ‘not useful at all’). Those in the blended mode group (80%) and the face-to-face mode group (67%) were more likely to select ‘very useful’ than those in the online mode group (22%). A similar pattern was evident when participants were asked about the usefulness of specific aspects of the course. Participants selected ‘very useful’ or ‘useful’ in all cases with the exception of (a) two (one in the online group; the other in the blended group) who selected ‘not very useful’ when asked how useful the course was in providing information about the language of the model texts, (b) two (in the online group) who selected ‘not very useful’ in relation to the writing of description/ classification texts, and one (in the online group) who selected ‘not very useful’ in relation to the writing of recount texts. Overall, ‘very useful’ was most often selected by members of the blended mode group (an average of 74%) than by members of the face-to-face mode group (an average of 65%) and the online mode group (an average of 31%).

Improvement in writing

Although a preference for face-to-face mode or blended mode seemed generally to be related to a desire for more social contact, some of the participants appeared to believe that greater progress could be made in face-to-face or blended mode groups. However, this was not borne out by the pre-test and post-test results.

Comparison of the average group scores in the pre-test and post-test revealed that the online mode group (+24%) outperformed the face-to-face mode group (+21%) and the blended mode group (+19%) overall and in terms of discourse relations and discourse relational signalling and other aspects of language typically associated with different genres. However, so far as overall rhetorical structure is concerned, the increase was greatest in the case of the blended mode group (+36.5%) followed by the face-to-face mode group (+30%), with the online mode group (+29%) in third position. It may be that the major increase in this area in the case of the blended mode group (+36.5%) is attributable to the fact the average score of that group in the pre-test (17.5%) was much lower than that of the other two groups (59%; 60.5%) and so there was more room for improvement.

The relationship between learning styles and improvement in writing

No relationship could be found between the learning styles of participants (as determined on the basis of the PSLI) and their preferred learning mode. However, irrespective of their

learning mode preferences and the learning modes to which they were assigned, students with I (introvert) in their learning style profiles improved more overall in terms of their writing performance (pre-test compared to post-test scores) than those with E in their learning styles profile (+21.6% versus +19%). This may have been partly because the course focused on individual writing rather than collaborative writing and partly because the intensive nature of the course, together with the fact that participants often completed tasks at different times, meant that there was less student/ student interaction than some participants may have wished.

The writing task performance of the online mode group: The possible impact of learning styles and English language proficiency

It has also been noted that the writing of participants with I (introversion) in their learning style profiles improved more overall than that of the other participants. The online mode group included a higher proportion of students with I in their learning style profiles (44%) than did the blended mode group (20%) and the face-to-face mode group (33%). It may be partly for this reason that the writing performance of the online mode group improved more overall than did that of the other two groups. Another contributing factor may have been proficiency in English. The average scores (out of 360) in the CSEPT were 259.5 (Group O), 229.8 (Group F), and 217.9 (Group B). Thus, the group (online) whose members performed best in terms of improvement in writing had the highest overall proficiency as well as the highest percentage of participants with I in their learning style profiles.

Conclusion

This study centred on an academic writing course involving tertiary-level EFL students delivered in three modes (fully online; blended; face to face). The course, one that focused on cognitive genres (and, in particular, on overall rhetorical structure, internal discourse structure and typical language features), was popular with the participants and led to demonstrable improvement in their writing (as indicated by criterion-referenced grading of pre-test and post-test writing tasks). Although most of the participants expressed a preference for face-to-face or blended modes, participants who were assigned to the online mode group had *at least* an equal chance of benefiting from the course. What emerges very strongly is that genre-centered writing courses have much to offer in adult EFL contexts, for example explicit instruction of internal discourse structure and specific language features associated with specific genres irrespective of the mode of delivery and irrespective of differing learning styles in the student body. In the case of this particular course, the focus was on cognitive

genres, the intention being to help the participants to “master the power of generalizations across disciplinary boundaries” (Bhatia, 1998, p. 26).

References

- Adams, M. J. (1991). Why not phonics and whole language. In W. Ellis (ed.), *All language and the creation of literacy* (pp. 40-53). Baltimore, MA: The Orton Dyslexia Society.
- Applebee, A. N. (1986). Problems in process approaches: Toward a reconceptualization of process instruction. In A. Petrosky & D. Bartholomae (eds.), *The teaching of writing* (pp. 95-113). Chicago: University of Chicago Press.
- Arndt, V. (1987). Six writers in search of texts: A protocol-based study of L1 and L2 writing. *ELT Journal*, 41(4), 257-267.
- Atkinson, D. (2003). L2 writing in the post-process era: Introduction. *Journal of Second Language Writing*, 12(1), 3-15.
- Atkinson, D., & Ramanathan, V. (1995). Cultures of writing: An ethnographic comparison of L1 and L2 university writing/ language programs. *TESOL Quarterly*, 29(3), 539-568.
- Badger, R., & White, G. (2000). A process genre approach to teaching writing. *ELT Journal*, 54(2), 153-160.
- Bazermann, C. (1994). Systems of genres and the enactment of social intentions. In A. Freedman & P. Medway (eds.), *Genre and the new rhetoric* (pp. 79-101). London: Taylor and Francis.
- Bhatia, V. K. (1998). Generic conflicts in academic discourse. In I. Fortanet, S. Posteguillo, J. C. Palmer, & J. F. Coll (eds.), *Genre studies in English for academic purposes* (pp. 15-28). Castello de la Plana: Publicacions de al Universitat Jaume.
- Biber, D. (1989). A Typology of English text. *Linguistics*, 27, 3-43.
- Bizzell, P. (1986). Composing processes: An overview. In A. Petrosky & D. Bartholomae (eds.), *The teaching of writing* (pp. 49-70). Chicago: University of Chicago Press.
- Braddock, R., Lloyd-Jones, R. & Schoer, L. (1963). *Research in written composition*. Urbana, IL: National Council of Teachers of English.
- Bruce, I. (2003). *Cognitive genre prototype modelling and its implications for the teaching of academic writing to learners of English as a second language*. PhD thesis. New Zealand: University of Waikato.
- Caudery, T. (1995). What the “process approach” means to practicing teachers of second language writing skills. *TESL-EJ*, 1(4), 1-11. Retrieved Feb 25, 2014, from <http://www.tesl-ej.org/wordpress/past-issues/volume1/ej04/ej04a3/>
- Cambourne, B. (1988). *The whole story: Natural learning and the acquisition of literacy in the classroom*. Auckland, NZ: Ashton Scholastic.

- Celce-Murcia, M., Dornyei, Z. & thurrell, S. (1997). Direct approaches in L2 instruction: A turning point in communicative language teaching? *TESOL Quarterly*, 31(1), 141-152.
- Crombie, W. & Johnson, D. (2004). Genre and text-type: putting discourses together. *Languages, Literary Studies and International Studies: An international Journal*, 1, 143-176.
- Crombie, W. & Johnson, D. (2009a). *Writing texts in English: A guide for intermediate and advanced learners*. Taichung, Taiwan: Ting Lung.
- Crombie, W. & Johnson, D. (2009b). *The Wenzao Ursuline College of Languages 36 credit English program: Issues and recommendations*. Kaohsiung, Taiwan: Wenzao Ursuline College of Languages.
- Cronbach, L.J. & Snow R.E. (1977). *Aptitudes and instructional methods*. New York: Irvington.
- Derewianka, B. (1994). *Exploring how texts work*. Victoria, (Australia): Australian Print Group.
- DfEE (1998). *The National Literacy Strategy Framework*. London: Department of Education and Employment.
- Dixon, J. (1967). *Growth through English: A report based on the Dartmouth Seminar 1966*. London: National Association of Teachers of English.
- Elbow, P. (1973). *Writing without teachers*. New York: Oxford University Press.
- Emig, J. (1971). *The composing processes of twelfth graders*. Urbana, IL: National Council of Teachers of English.
- Emig, J. (1983). *The web of meaning: Essays on writing, teaching, learning, and thinking*. Portsmouth, NH: Boynton/Cook.
- Entwistle, N.J. (1988). *Styles of learning and teaching*. London: David Fulton.
- Faigley, L. (1986). Competing theories of process: A critique and proposal. *College English*, 48(6), 527-542.
- Flower, L. & Hayes, J. (1981). A cognitive process theory of writing. *College Composition and Communication*, 32(4), 365-387.
- Fogarty, D. (1959). *Roots for a new rhetoric*. New York: Teachers College, Columbia University.
- Gao, Jiajing. (2007). Teaching writing in Chinese universities: Finding an eclectic approach. *Asian EFL Journal*, 20, Professional teaching article, Article 2. Retrieved at

<http://asian-efl-journal.com/teaching-articles/2007/05/05/teaching-writing-in-chinese-universities-finding-an-eclectic-approach/#thethe-tabs-1-1> 25 February 2014]

- Gagne, R. & Briggs, L.J. (1979). *Principles of instructional design*. New York: Holt, Reinhart and Winston.
- Gould, J. D. (1980). Experiments on composing letters: Some facts, some myths, and some observations. In L. W. Gregg & E. R. Steinberg (eds.), *Cognitive processes in writing* (pp. 97-128). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Gough, J. (1996). *Developing learning materials*. London: Institute of Personnel and Development.
- Harris, R. J. (1962). *An experimental inquiry into the functions and value of formal grammar in the teaching of English, with special reference to the teaching of correct written English to children aged twelve to fourteen*. PhD thesis, London: University of London.
- Hayes, J. & Allison, C.W. (1996). The implications of learning styles for training and development: A discussion of the matching hypothesis. *British Journal of Management* 7 (1): 63-73.
- Hayes, J. R. & Flower, L. S. (1980). Identifying the organization of writing processes. In L. Gregg & E. R. Steinberg (eds.) *Cognitive processes in writing* (pp. 3-30). Hillsdale, NJ: Lawrence Erlbaum.
- Hinkel, E. (2004). *Teaching academic ESL writing: Practical techniques in vocabulary and grammar*. New York: Lawrence Erlbaum Associates.
- Hyland, K. (2002). *Teaching and researching writing*. Essex: Pearson Education.
- Hyland, K. (2003). Genre-centered pedagogies: A social response to process. *Journal of Second Language Writing*, 12(1), 17-29.
- Johns, A. M. (1990). L1 composition theories: Implications for developing theories of L2 composition. In B. Kroll (ed.), *Second language writing: Research insights for the classroom* (pp. 24-36). Cambridge: Cambridge University Press.
- Knapp, P., & Watkins, M. (2005). *Genre, text, grammar*. Sydney: University of New South Wales.
- Lin, B. (2006). Vygotskian principles in a genre-based approach to teaching writing. *NUCB Journal of Language, Culture and Communication*, 8(1), 69-84. Retrieved from <http://www.nucba.ac.jp/cic/pdf/njlcc081/05LIN.pdf>
- Matsuda, P. K. (2003). Process and post-process: A discursive history. *Journal of Second Language Writing*, 12(1), 65-83.

- Miller, S. (1991). *Textual carnivals: The politics of composition*. Carbondale, IL: Southern Illinois University Press.
- Ministry of Education (New Zealand). (1994). *English in the New Zealand curriculum*. Wellington, New Zealand: Learning Media.
- North, S. (1987). *The making of knowledge in composition*. London: Heinemann.
- Nystrand, M. (2006). The social and historical context for writing research. In C. A. MacArthur, S. Graham, & J. Fitzgerald (eds.), *Handbook of writing research* (pp. 11-27). New York, London: The Guilford Press.
- Nystrand, M., Greene, S., & Wiemelt, J. (1993). Where did composition studies come from? An intellectual history. *Written Communication, 10*(3), 267-333.
- Odell, L. (1980). Teaching writing by teaching the process of discovery: An interdisciplinary enterprise. In L. W. Gregg & E. R Steinberg (eds.), *Cognitive processes in writing* (pp. 139-154). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Pennington, M. C., & So, S. (1993). Comparing writing process and product across two languages: A study of 6 Singaporean university student writers. *Journal of Second Language Writing, 2*(1), 41-63.
- Pullman, G. (1999). Stepping yet again into the same current. In T. Kent (ed.), *Post-process theory: Beyond the writing-process paradigm* (pp. 16-29). Carbondale: Southern Illinois University Press.
- Quinn, J. (1993). A taxonomy of text types for use in curriculum design. *EA Journal, 11*(2), 33-46.
- Raimes, A. (1983). *Techniques in teaching writing*. Oxford: Oxford University Press.
- Raimes, A. (1985). What unskilled ESL students do as they write: A classroom study of composing. *TESOL Quarterly, 19*(2), 229-258.
- Rowntree, D. (1982). *Educational Technology in Curriculum Development*. London: Paul Chapman Publishing.
- Reigeluth, C.M. Instructional design. What is it and why is it? In C.M. Regeluth, ed., *Instructional design theories and models: an overview of their current status*, pp3-96. Hillsdale, NJ:LEA
- Riding, R.J. & Cheema, I. (1991). Cognitive styles- an overview and integration. *Educational Psychology 11*(3 & 4) 193-215.
- Sadler, E. & Riding, R. (1999). Cognitive styles and instructional preference. *Instructional Science 27*, 355-371.
- Schmeck, R.R. (1988). *Styles and strategies of learning*. London: Plenum.

- Shindler, J. & Yang, H. (2004, October 24). Paragon learning style inventory: Combinations. [Retrieved from <http://www.oswego.edu/plsi/combo1.htm> 25 February 2014]
- Susser, B. (1993). Networks and project work: Alternative pedagogies for writing with computers. *Computers and Composition*, 10(3), 63-89.
- Swales, J. (1990). *Genre analysis: English in academic and research settings*. Cambridge: Cambridge University Press.
- Tangpermpoon, T. (2008). Integrated approaches to improve students writing skills for English major students. *ABAC Journal*, 28(2), 1-9.
- Tobin, L. (1994). Introduction: How the writing process was born – and other conversion narratives. In L. Tobin & T. Newkirk (eds.), *Taking stock: The writing process movement in the '90s* (pp. 1-14). Portsmouth, NH: Boynton/Cook Heinemann.
- Tobin, L. (2001). Process pedagogy, In G. Tate, A. Rupiper, & K. Schick (eds.), *A guide to composition pedagogies* (pp. 1-18). New York: Oxford University Press.
- Wolff, D. (2000). Second language writing: A few remarks on psycholinguistic and instructional issues. *Learning and Instruction*, 10(1), 107-112.
- Young, R. E. (1978). Paradigms and problems: Needed research in rhetorical invention. In C. R. Cooper & L. Odell (eds.), *Research on composing: Points of departure* (pp. 29-48). Urbana, IL: National Council of Teachers of English.
- Zamel, V. (1983). The composing processes of advanced ESL students: six case-studies. *TESOL Quarterly*, 17(2), 165-187.

Appendix 1

Table 1a: Criterion-referenced analysis/grading scheme (Part A: specific to what was taught)

| Part A | Features | Points | | Possible score | | Actual score | |
|--------------------|--|---|--|----------------|----|--------------|--|
| | | | | 10 | 10 | | |
| | Generic structure (steps, stages etc. – includes appropriate paragraphing within text segments) | Up to 10 points for overall structuring; | Up to 10 points for appropriate paragraphing and paragraph linkage within text segment | 10 | 10 | | |
| | Semantic relations and their signaling (e.g. occurrence of Temporal Sequence in recount texts; Grounds-Conclusion in argument texts) | Up to 10 points for semantic relational occurrences | Up to 10 points for accurate and appropriate semantic relational signaling | 10 | 10 | | |
| | Language characteristic of the genre used accurately and appropriately (e.g. imperative constructions in instruction texts) | Up to 10 points for selection of language that is appropriate to the genre (e.g. imperative constructions in instruction texts) | | 10 | | | |
| TOTAL SCORE | | | | 50 | | | |

Table 1.b: Criterion-referenced analysis/grading scheme (Part B: general)

| Part B | Features | Points | Possible score | Actual score |
|--------------------|--------------------------------|--|-----------------------|---------------------|
| | Overall impression | Up to 10 points | 10 | |
| | Length | Up to 10 points (remove 1 point for every 10 words short of 250) | 10 | |
| | Ideas & ideas development | Up to 10 points | 10 | |
| | Grammatical accuracy | Up to 10 points (delete 1 point for each grammatical error (maximum of 2 point deduction for same grammatical error occurring more than once)) | 10 | |
| | Appropriate lexical selections | Up to 10 points (delete 1 point for each inappropriate lexical selection (maximum of 2 point deduction for same lexical selection error occurring more than once)) | 10 | |
| | Punctuation | Up to 5 points (delete 1 point for each punctuation error) | 5 | |
| | Spelling | Up to 5 points (delete 1 point for each spelling error (maximum of 2 point deduction for same spelling selection error occurring more than once)) | 5 | |
| TOTAL SCORE | | | 60 | |

Appendix 2

Table 3: *How much did you enjoy the course?*

| Groups | I liked it a lot | I liked it | So-so | I did not like it at all |
|--------------|------------------|------------|---------|--------------------------|
| | N (%) | N (%) | N (%) | N (%) |
| Online | 3 (33%) | 5 (56%) | 1 (11%) | 0 (0%) |
| Blended | 7 (70%) | 3 (30%) | 0 (0%) | 0 (0%) |
| Face-to-face | 6 (67%) | 2 (22%) | 1 (11%) | 0 (0%) |

Table 4: *How useful was the course in helping you to write texts?*

| Groups | Very useful | Useful | Not very useful | Not useful at all |
|--------------|-------------|---------|-----------------|-------------------|
| | N (%) | N (%) | N (%) | N (%) |
| Online | 2 (22%) | 7 (78%) | 0 (0%) | 0 (0%) |
| Blended | 8 (80%) | 2 (20%) | 0 (0%) | 0 (0%) |
| Face-to-face | 6 (67%) | 3 (33%) | 0 (0%) | 0 (0%) |

Table 5: *How useful was the course in providing you with information about the language of the model texts?*

| Groups | Very useful | Useful | Not very useful | Not useful at all |
|--------------|-------------|---------|-----------------|-------------------|
| | N (%) | N (%) | N (%) | N (%) |
| Online | 3 (33%) | 5 (56%) | 1 (11%) | 0 (0%) |
| Blended | 7 (70%) | 2 (20%) | 1 (10%) | 0 (0%) |
| Face-to-face | 5 (56%) | 4 (44%) | 0 (0%) | 0 (0%) |

Table 6: *How useful was the course in helping you to understand more about language (generally)?*

| Groups | Very useful | Useful | Not very useful | Not useful at all |
|--------------|-------------|---------|-----------------|-------------------|
| | N (%) | N (%) | N (%) | N (%) |
| Online | 3 (33%) | 6 (67%) | 0 (0%) | 0 (0%) |
| Blended | 8 (80%) | 2 (20%) | 0 (0%) | 0 (0%) |
| Face-to-face | 5 (56%) | 4 (44%) | 0 (0%) | 0 (0%) |

Table 7: *How useful was the course in teaching you to use language accurately?*

| Groups | Very useful | Useful | Not very useful | Not useful at all |
|--------------|-------------|---------|-----------------|-------------------|
| | N (%) | N (%) | N (%) | N (%) |
| Online | 3 (33%) | 6 (67%) | 0 (0%) | 0 (0%) |
| Blended | 8 (80%) | 2 (20%) | 0 (0%) | 0 (0%) |
| Face-to-face | 7 (78%) | 2 (22%) | 0 (0%) | 0 (0%) |

Table 8: *How useful was the course in helping you to write texts in each of the different genres? (face-to-face group)*

| Genres | Very useful | Useful | Not very useful | Not useful at all |
|----------------------------------|-------------|---------|-----------------|-------------------|
| | N (%) | N (%) | N (%) | N (%) |
| Instructions | 6 (67%) | 3 (33%) | 0 (0%) | 0 (0%) |
| Arguments | 5 (56%) | 4 (44%) | 0 (0%) | 0 (0%) |
| Descriptions/ Classifications | 6 (67%) | 3 (33%) | 0 (0%) | 0 (0%) |
| Recounts | 5 (56%) | 4 (44%) | 0 (0%) | 0 (0%) |

Table 9: How useful was the course in helping you to write texts in each of the different genres? (online group)

| Genres | Very useful | Useful | Not very useful | Not useful at all |
|----------------------------------|-------------|---------|-----------------|-------------------|
| | N (%) | N (%) | N (%) | N (%) |
| Instructions | 3 (33%) | 6 (67%) | 0 (0%) | 0 (0%) |
| Arguments | 3 (33%) | 6 (67%) | 0 (0%) | 0 (0%) |
| Descriptions/ Classifications | 2 (22%) | 5 (56%) | 2 (22%) | 0 (0%) |
| Recounts | 2 (22%) | 6 (67%) | 1 (11%) | 0 (0%) |

Table 10: How useful was the face-to-face section in helping you to write texts in each of the different genres? (blended group)

| Genres | Very useful | Useful | Not very useful | Not useful at all |
|----------------------------------|-------------|---------|-----------------|-------------------|
| | N (%) | N (%) | N (%) | N (%) |
| Instructions | 7 (70%) | 3 (30%) | 0 (0%) | 0 (0%) |
| Arguments | 7 (70%) | 3 (30%) | 0 (0%) | 0 (0%) |
| Descriptions/ Classifications | 6 (60%) | 4 (40%) | 0 (0%) | 0 (0%) |
| Recounts | 7 (70%) | 3 (30%) | 0 (0%) | 0 (0%) |

Table 11: How useful was the **online** section in helping you to write texts in each of the different genres? (blended group)

| Genres | Very useful | Useful | Not very useful | Not useful at all |
|----------------------------------|-------------|---------|-----------------|-------------------|
| | N (%) | N (%) | N (%) | N (%) |
| Instructions | 8 (80%) | 2 (20%) | 0 (0%) | 0 (0%) |
| Arguments | 7 (70%) | 3 (30%) | 0 (0%) | 0 (0%) |
| Descriptions/ Classifications | 7 (70%) | 3 (30%) | 0 (0%) | 0 (0%) |
| Recounts | 7 (70%) | 3 (30%) | 0 (0%) | 0 (0%) |

Table 12: How useful were the model texts that were included in the materials?

| Groups | Very useful | Useful | Not very useful | Not useful at all |
|--------------|-------------|---------|-----------------|-------------------|
| | N (%) | N (%) | N (%) | N (%) |
| Online | 4 (44%) | 5 (56%) | 0 (0%) | 0 (0%) |
| Blended | 8 (80%) | 2 (20%) | 0 (0%) | 0 (0%) |
| Face-to-face | 7 (78%) | 2 (22%) | 0 (0%) | 0 (0%) |

Table 13: What is your preferred learning mode?

| Face to face only | Blended only | Face to face or Blended | Online or Blended | Online or Face-to-face |
|-------------------|--------------|-------------------------|-------------------|------------------------|
| 4 (14%) | 3 (11%) | 17 (61%) | 3 (11%) | 1 (4%) |

Table 14: *Would you like to do another writing course of a similar type?*³

| Groups | Yes, I would very much like to do such a course. | Yes, that would be okay. | No. |
|--------------|--|--------------------------|---------|
| | N (%) | N (%) | N (%) |
| Online | 5 (56%) | 1 (11%) | 3 (33%) |
| Blended | 8 (80%) | 2 (20%) | 0 (0%) |
| Face-to-face | 7 (78%) | 1 (11%) | 1 (11%) |

³ Note that two of the online group members who selected 'No' indicated that they would like to do a course of a similar type in the future if they were not assigned to online mode. Also, the face-to-face group member who selected 'No' indicated that she would like to do a similar course in the future if she were not assigned to the face-to-face mode.

Multiple Perspectives on the Self in SLA

S. Mercer & M. Williams (eds). 2014. *Multilingual Matters*: Bristol, ix + 189 pages

Reviewed by Brian Wadman, Mahidol University, Thailand

Mercer and Williams provide the reader with a monolithic account on how the acquisition of a second language is affected by notions of the self. The book was written for experienced, novice and aspiring language teachers. The vastness of the knowledge required to fully understand the book is daunting, however, the authors do provide clear conceptualizations of major ideas related to the self and simplify psychological terminology which has only recently been applied to SLA. The expansive number of references and thought provoking ideas only whets the appetite for more reading. Scholars should add this book to their collection to be up to date with some of the most recent academic developments in the field.

Chapter 1 provides the readers with the authors' impetus for compiling the specific topics in the book. Williams and Mercer explain why it is important for language educators to have a good background knowledge of the self because unlike other subjects, which are additive, language teachers must "support learners in constructing and negotiating their identities through the use of language." (p.2).

Chapter 2 attempts to clarify the web of theories and definitions which often get confused and jumbled together by those with little background in psychology. The work of Zoltan Dornyei (among others) is highlighted, substantiating the role of motivation in the language learning process. As well, those who are unfamiliar with educational psychology can update their knowledge on the more recent research in this particular area.

Chapter 3 explores the role that confidence and anxiety plays in SLA. The potential benefits of Computer Mediated Communication are featured given its increasing role in the development

Address correspondence:
999 Phuttamonthon 4 Road,
Salaya, Nakhon Pathom
73170 Thailand

of L2 fluency. Included in this chapter, readers will find a framework for language contact, a visual paradigm that relates the varying levels of the involvement of self with the degree of

richness in their contact experience. This model is extremely useful when considering student development towards full proficiency in an L2.

Chapter 4 requires thoughtful consideration as it discusses the oft-banded about terms of self-esteem, self-efficacy and self-concept. These notions are of the highest order in SLA education, and the author provides great depth of knowledge as well as offering practical implications in the classroom. The particular model, which expresses the necessary conditions to develop positive self-images is explained, and the author validates the essentiality of notions of self in SLA.

The topic of chapter 5, Post-structuralism in SLA, appears to be inconsistent with the other chapters of the self. The research to support post-structural notions of identity is mostly qualitative and the author mentions her own research a number of times. The main topic of identity formation and its dynamic process is relevant, but mostly, the chapter just strews from such hot-button issues as feminism, racism and elitism adapted to SLA in a cookie cutter style. The work would be better found in a collection of essays on critical pedagogy.

Chapter 6 begins with a thorough examination of the different views of identity. It clarifies a number of concepts not well defined in previous chapters. There are a number of intriguing ideas to be found here as notional ideas of self are expressed and the varying perspectives on cultural relativity are considered. The second half of the chapter is a bit brief as it attempts to provide analysis on the identity of bilinguals. Sapir-Whorfian ideas are touched upon, but there are no clear conclusions and only one research study is investigated (in one context), which does little to add to the current global perspectives of dual identities in bilingualism.

Poets have been saying for a long time that there are multiple parts of the self. In Chapter 7 public and private views of the self are considered and the consequences of these divergent notions are applied to SLA. Interestingly, the research supports the idea of a true authentic self being fundamental to language learning. The chapter covers internalization, possible-selves and self-discrepancy theory.

At the heart of chapter 8 is the power of the imagination and the role of agency in SLA. The research is convincing and can be paraphrased in Henry Ford's maxim "whether you think you can or you can't – you're right." The chapter concludes with an overview of Dornyei's L2 motivational self-system and pedagogical applications of research on the imagination.

Chapter 9 focuses on the internal and external processes of the mind which shape the self. Readers will find both a more nuanced perspective of motivation, a fundamental element in the development of the self. In particular, this chapter clarifies for the reader the main ideas of how experience, social factors and our future visualizations of our selves effect development.

A burgeoning field of SLA research, neurophilosophy, is discussed and described in chapter 10. French philosopher Rene Descartes is brought into the material and much to this reader's surprise, "I think, therefore I am" is dismissed as a modern perspective of empirical truth. However, a potential new truth is revealed in its place. Described as being found in that small space between the pre-reflective sub-conscious and consciousness, new methodological techniques, including MRI brain scanning could offer scientific evidence of self-existence. Language is put at the forefront due to its organizing and high order cognitive functions

In chapter 11, Sarah Mercer calls for a more complex perspective of the self. The case is convincing as this perspective, while limited in scope due to the humility of what we actually know about the brain, does pin down some factual truths. The author suggests researchers should be chasing an integrative, dynamic and emergent self, and interestingly claims a lack of confidence, which may result from a low SCC (self-concept certainty), is a positive aspect in 2nd language learning due to the fact it is more dynamic and open to change. The author highlights her own research conducted in Austria, and admits to the difficulties in researching the self.

The final chapter reflects on the knowledge found in the research of the self and describes the different aspects which make up the wide area of knowledge. Importantly for language educators, a section outlining the pedagogical implications puts the research into practical terms.

Multiple Perspectives of the Self in SLA is a must-read for the language educator who wants a modern, comprehensive view of this extremely important concept of the self for SLA . The book offers a sound foundational knowledge of the self and offers a path which could lead the teacher to become more proficient in his/her pedagogical skill.

The submission has not been previously published or is not being considered for publication elsewhere.

Brian Wadman is a lecturer/visiting professor at Mahidol University's Faculty of Liberal Arts. He graduated from the University of Massachusetts at Boston in Applied Linguistics in 2009.

His research interests include critical pedagogy, virtual learning environments, and English as a Lingua Franca. He can be reached at bwad1976@gmail.com.

Language, Education and Nation-building- Assimilation and Shift in Southeast Asia

P. Sercombe & R. Tupas, Palgrave Macmillan, NY. 2014. Pp. XIV, 264

Reviewed by Raheb Zohrehfard, University of British Columbia, Vancouver, Canada

In an attempt to trace the connections between language, education and nation-building in Southeast Asia, the writers focus on how states and governments have employed language policies as tools to exert power on minorities and assimilate them into the mainstream cultural and linguistic politics. The common theme throughout the whole book shows how language use forms and demonstrates linguistic and cultural maintenance and adaptation. This book can potentially address policy makers, curriculum designers and school administrators to strive hard towards establishing an egalitarian pedagogical system in multicultural contexts. The early history of Southeast Asia and colonialism in the region is very well explicated with further information about the population, education system and ethnographic data (pp. 1-5).

In chapter one, *Sercombe* delineates how the ideology of one nation and establishment of *Bahasa Melayu* as the sole official language of the country creates an assimilationist incorporation of all citizens into a single community which marginalized the minority groups who speak neither English nor *Bahasa Melayu*.

Frewer in chapter two clarifies the fact that Chinese, Vietnamese and Muslim communities have been subject to assimilationist projects, which led to suppression of their identities in favor of tendencies to khmerize Cambodia. The development work, though problematizing issues relative to linguistic homogeneity, inequality and justice, has been complicit in creating an image of a solely khmerized Cambodia.

The focus of chapter three is on Timor Leste where *Curaming* and *Kalidjernih* describe the

Address correspondence:

2329 West Mall, Vancouver,
BC Canada V6T 1Z4

landscape of language policies as initially sentimental followed by a more workable trajectory. Globalization has given rise to

English and other local languages to vitalize the language and identity politics bringing about a push for multilingual education which has made the young, multilingual society of Timor Leste a site of struggle for smaller languages given its economic and political transience.

In chapter 4, *Musgrave* brings to the fore the polyglossic situation as the result of different varieties of ‘Bahasa Indonesia’ which is complicated by a linguistic divide between western and eastern parts of the country. The government’s policy of encouraging internal migration (p. 92) resulted in massive population movements, which contributed to language shift.

In chapter 5, *Cincotta-segi* tracks the tension between multiethnicity and multilingualism. The institutionalized support for the Lao language as the official language and medium of instruction leaves little or no chance for multilingual education. However, the Lao culture is based on a shaky ground since teachers and pupils relied on hybrid communicative practices through use of their first languages, which takes issue with the official assimilationist discourse of Lao-ization.

Chapter 6 by *Maya Khemlani David* and *James McLellan* addresses a dichotomy between ‘the colonial past’ and ‘the post-colonial present’ in Malaysia (Tupas, 2004). The country’s colonial past and the current trends towards globalization as well as the constitutional safeguards accorded to the ‘Bumiputera’ (a term used to refer to indigenous groups in Sabah and Sarawak, meaning ‘sons of the soil’), contributed to the language shift in four minority groups.

In chapter 7, *Khin* and *Sercombe* describe paradoxes in language policy-making and practice in Myanmar. They indicate how the state exerts power on minority groups through monolingual and monocultural policies and how the country’s dominant political circumstances ensnarl language policy to be more democratic.

Chapter 8 is where *Tupas* and *Lorente* discuss the combined official use of Filipino and English in the Philippines, which shaped bilingual education indicating positive effects on student learning. The current paradigm of World Englishes, vigorously promoted by Braj Kachru (1997b) has left its trace in the Philippines where indigenized varieties of English represent nationalist voices.

In Singapore, discussed in chapter 9 by *Wee*, language policy has brought about clines of linguistic minoritization. One intriguing tension, which is worth mentioning is the discussion of English as a Lingua Franca (ELF) and World Englishes (WE) in the context of Singapore (Pakir, 2009). Although both frameworks support a pluricentric view of the form and use of English,

they differ in that ELF implies borderless hybrid uses of English by nonnative speakers, whereas English in Singapore, though similarly used as a lingua franca, is forming nativized uniqueness denoting national linguistic identity.

Kosonen and Person, in chapter 10, unpack the monoculturalist and assimilationist policy-making in Thailand through the lens of Thai ethnolinguistic identities. The language policy in Thailand, due to its linguistic diversity (p.208), supports the use of non-dominant languages (NDLs) in education, which has encouraged a pluralistic position. However, what takes central stage is that plurilingualism calls for the development of cultural and linguistic competence, which prompts the perpetuation of the cultural and linguistic imperialism.

In chapter 11, *Phan Le Ha*, *Vu Hai Ha* and *Bao Dat* focus on how the convoluted nexus between nationalism and globalization has caused a dual embrace of English and Vietnamese, which affects ethnic minorities. The nationalist language policy has systematically and insidiously made other ethnic languages inaccessible, unequal and unqualified.

Through an anthropological perspective, Giordano concludes in her epilogue that language politics is deeply entwined with complex processes of colonization, nationalism and globalization, both protracting the hegemony of Europe and the United States as references for the subaltern societies and re-appropriating and adapting the Western conceptions of nation state, under the guise of being ethnically identified, politically autonomous and socioeconomically empowered.

This book is a high-quality, worthwhile read that provides (1) a useful theoretical understanding of linguistic and cultural assimilation and marginalization in the nation-building agendas of SE Asian countries, (2) a pragmatic and programmatic aspect of neoliberalism as ‘ a merge of the state and the market in a form of corporate governance with a focus on linguistic imperialism in perpetuating neoliberal agenda’ (Flores, 2013), and (3) an insight into how the denial of multilingualism for marginalized populations produces and justifies social violence (McNamara, 2011). Nevertheless, the book could have expanded the discussion on the limitations of current approaches by emphasizing how negotiating the multiple linguistic and cultural differences in societies can be essential to the pragmatics of the working, civic, and private lives of citizens and students. Such a negotiation would provide students with access to the language of power, work, and community, fostering the critical engagement to design their social futures.

References

- Flores, N. (2013). The unexamined relationship between neoliberalism and plurilingualism: A cautionary tale. *TESOL Quarterly*, 47, 500-520.
- Kachru, Braj B. (1997b) World Englishes and English using communities. *Annual Review of Applied Linguistics*, 17, 66-87.
- Kubota, R. (2014). The Multi/Plural Turn, Postcolonial Theory, and Neoliberal Multiculturalism: Complicities and Implications for Applied Linguistics. *Applied Linguistics*, amu045.
- McNamara, T. (2011). Multilingualism in Education: A poststructuralist critique. *The Modern Language Journal*, 95, 430-41.
- Pakir, A. (2009). English as a lingua franca: analyzing research frameworks in international English, world Englishes, and ELF. *World Englishes*, 28(2), 224-235.
- Tupas, T. R. F. (2004). The politics of Philippine English: Neocolonialism, global politics, and the problem of postcolonialism. *World Englishes*, 23(1), 47-58.

The submission of this book review has not been previously published or is not being considered for publication elsewhere.

Raheb is now a Master's student in TESL in the Language and Literacy Education Department at the University of British Columbia. His focus is on identity issues of those who immigrate to English speaking countries. He tries to help non-native English speaking students to negotiate and construct their identity and legitimacy in their new community.

English Literacy among Malaysian First Graders:

Boys' Achievement Relative to Girls

Ratnawati Mohd-Asraf¹
International Islamic University Malaysia

Hazlina Abdullah²
Islamic Science University Malaysia

Ratnawati Mohd Asraf is a professor at the Faculty of Education, International Islamic University Malaysia, and Dean of the Research Management Centre, IIUM. She has published in the areas of reading, literacy education, and second and foreign language learning and teaching. Email: ratnawati@iium.edu.my

Hazlina Abdullah obtained her M.Ed TESL from Universiti Pendidikan Sultan Idris (UPSI), Malaysia, and her B.Ed TESOL degree from the University of Warwick, England. She has vast teaching experience at the secondary and tertiary levels. Hazlina is currently a lecturer at the Islamic Science University of Malaysia (USIM). Email: hazlina@usim.edu.my

Abstract

In the past few decades, data from numerous countries such as the United Kingdom, Canada, Australia, the Philippines, Thailand, and Malaysia show that many boys are not performing well on a range of educational achievement measures such as literacy, and that the gap between boys' and girls' performance has increased over time. This paper discusses the results of a study that sought to determine the English literacy performance of Malaysian first graders in English. Using one of the states in the country as a case study, English literacy performance data on approximately 85,000 students from grade 1 were obtained. Teachers' views were also sought through focus group interviews and journal entries. The results show that girls outperformed boys on all English literacy constructs.

Keywords: boys' underperformance, literacy, English literacy, English literacy performance, reading.

Affiliation address:

¹ Kuliyyah of Education, International Islamic University Malaysia, P.O. Box 10, 50728 Kuala Lumpur, Malaysia. Email: ratnawati@iium.edu.my. Phone: +60122169693

² Faculty of Major Language Studies, Islamic Science University of Malaysia (USIM), Bandar Baru Nilai, 71800 Nilai, Negeri Sembilan, Malaysia. Email: hazlina@usim.edu.my

Introduction

In the past twenty years, the underachievement of boys has been an issue of central concern among researchers and academics throughout the world (Booth, Elliott-Johns, & Bruce, 2009; Noble, Brown, & Murphy, 2001; Warrington & Younger, 2006). Statistics from countries such as the United Kingdom, the United States, Australia, Canada, the Philippines, Thailand, and Malaysia reveal that girls are doing better in school settings, surpassing boys (Cappon, 2011; Lindsay & Muijs, 2005; Weaver-Hightower, 2008). In Australia for example, the evidence regarding boys' underachievement was drawn by the Commonwealth of Australia (2002) through several key indicators such as: a) *Early Literacy Achievement Measures*, where, in 2000, it was found that 3.4% fewer Year 3 boys and 4.4% fewer Year 5 boys achieved the national benchmarks than the girls in Year 3 and Year 5 respectively; b) *School Retention*, where the retention rate of boys was 11 percentage points lower than that of the girls in 2001; c) *Results* in most subjects at Years 10 and 12, in which girls are achieving a higher average and the gap between boys' and girls' total marks has widened markedly; d) *Admission to Higher Education*, where about 56% of university commencements are females; and e) *Other Indicators*, which include suspensions and expulsions, where many more boys than girls are involved (Commonwealth of Australia, 2002 p. xvi).

Studies have also shown that girls have outperformed boys in literacy (Carr & Pauwels, 2006; Cuttance & Thompson, 2008). In the United Kingdom, for example, the national achievement patterns show that not only has the gender gap in entry and performance at the age of 16 closed, but that new gender gaps have opened up, and that girls are now outperforming boys. From 1996 to 2004, girls consistently did better than boys in O-level examinations in England (<http://www.geog.cam.ac.uk/research/projects/>). In Canada, boys "...trail in reading and writing, ... [and many] of them land in the bottom quarter of standardized tests" (from *Globe and Mail*, October 15, 2010, cited in Cappon, 2011, p. 1) while in the United States, they were found to be about one and one-half years behind girls in reading and writing (Donahue, Finnegan, Lutkus, Allen, & Campbell, 2001; Snyder & Hoffman, 2001). The same pattern is also seen in the East Asia and Pacific countries, including the Philippines, Thailand, and Malaysia (UNESCO, 2009). Indeed, according to the Programme for International Student Assessment (PISA), in every country, girls have outshone boys in literacy (OECD 2010).

The worldwide trend in the academic performance of boys as well as their performance in literacy is a cause for concern because numerous studies have suggested that low academic achievement is highly related to weak literacy skills (Annamalai & Muniandy, 2013; TIMSS & PIRLS International Study Center, 2011). This is because literacy is “the prerequisite to academic achievement” (Snow & Biancarosa, 2003); and those that lack these basic skills are more likely to struggle not only in school, but also throughout life (Clark & Burke, 2012; National Literacy Trust, 2012). The concern for the importance of fostering early literacy prompted us to conduct a study on the reading performance of Malaysian first graders.

This paper presents the results of this study, in which we attempted to investigate the literacy performance of Malaysian boys and girls in their first year of formal schooling. Specifically, it sought to determine whether there are discernible patterns in the boys’ and girls’ literacy scores in their second language—the English language—so that efforts could be taken to help improve their literacy and academic performance.

Literacy and Its Assessment

Most countries in the world view literacy as being a significant issue for the development of their nations (National Literacy Trust, 2012). To succeed academically, it is vital for students to master literacy skills, and the literature has shown that the mastery of literacy skills at an early age is crucial (Connolly, 2004; Jones, 2003; Lindsay & Muijs, 2005; Rathvon, 2004; Warrington & Younger, 2006). Hence, it is important to have a coordinated system to monitor students’ literacy development.

A number of international assessments have been developed to assess students’ literacy. Among them are the Program for International Student Assessment (PISA), which measures the performance of 15-year-olds in reading literacy, mathematics, and science (OECD, 2010); the Progress in International Reading Literacy Study (PIRLS), to determine students’ reading comprehension covering reading for literary experience and reading to acquire and use information (TIMSS & PIRLS International Study Center, 2011); and the Literacy Assessment and Monitoring Program (LAMP), which measures the reading and reading component skills as well as the numeracy skills of youth and adults, (UNESCO Institute for Statistics, 2009).

In Malaysia, students’ literacy is assessed by the Literacy and Numeracy Screening (LINUS) program, which is conducted to identify students with difficulties in reading, writing,

and basic arithmetic. Within the first three years of their primary school education, all Malaysian children are to be equipped with a sound base in basic literacy and numeracy skills. The LINUS program, first started in 2010, assesses students from Grades 1 to 3 on their performance in literacy and numeracy. Two screenings are carried out; one in May and the other in September or October (Ministry of Education Malaysia, 2013). However, this study only focused on the assessment of literacy.

In the LINUS program, literacy in English is defined as being able to: (a) communicate appropriately with peers and adults in English; (b) read and comprehend simple texts and stories in English, and (c) write a range of texts from a variety of media in English. Twelve constructs are measured, as shown in Table 1:

Table 1: Twelve Constructs Measured in the Literacy Screening

| Construct | Description |
|------------------|--|
| 1 | Able to identify and distinguish shapes of the letters of the alphabet. |
| 2 | Able to associate sounds with the letters of the alphabet. |
| 3 | Able to blend phonemes into recognizable words. |
| 4 | Able to segment words into phonemes. |
| 5 | Able to understand and use the language at the word level. |
| 6 | Able to participate in daily conversations using appropriate phrases. |
| 7 | Able to understand and use the language at phrase level in linear texts. |
| 8 | Able to understand and use the language at phrase level in non-linear texts. |
| 9 | Able to read and understand sentences with guidance. |
| 10 | Able to understand and use the language at sentence level in non-linear texts. |
| 11 | Able to understand and use the language at paragraph level in linear texts. |
| 12 | Able to construct sentences with guidance. |

(Source: Ministry of Education Malaysia, 2013, p. 2-3)

The literacy assessment instruments were developed by a panel of national experts under the supervision of the Malaysian Examinations Syndicate, after which they were verified in terms of their validity, reliability, feasibility, and quality by another panel of national experts (Examinations Syndicate, Ministry of Education Malaysia, 2012; Ministry of Education Malaysia 2013). The instruments underwent several processes based on a standardized Work Procedure Manual, starting with the development and evaluation of the items in line with the

item specifications, and which contained the aspects of measurement, context, and the level of difficulty of the items (Examinations Syndicate, Ministry of Education Malaysia, 2012).

Methods

The study was guided by the following research question: *How do primary school boys in the state of Selangor perform in English literacy relative to girls?* The state of Selangor was chosen as a case study because of the diversity of creeds, cultures, gender, races, and education levels of its population, which, by and large, reflects that of Malaysia as a whole.

To obtain the most accurate results regarding the literacy performance of boys and girls in this state, we collected the *entire population* of English language literacy scores from *all* the students in Grade 1, from *all* the government primary schools in the state of Selangor. The data were obtained from the Selangor State Department of Education through its nationwide LINUS program. A focus group interview with English language teachers from 6 selected schools in the state was also conducted to get an insight into how they felt about the literacy performance of boys and girls. The interview was chosen as a means of data collection as it would provide the opportunity to observe a large amount of interaction on the topic being discussed, and at the same time allowed for a large amount of concentrated data that were directly related to the focus and interest of the study (Morgan, 1997). Since the interviewees were all primary school teachers teaching Grade 1, it allowed them to share their personal views in a similar and comparable context. This was indeed an advantage, as focus group interviews would generate “the best information” in a situation where the interviewees are alike, and can positively work together (Cresswell 2008, p. 226). Furthermore, according to Krueger & Casey (2009) and Onwuegbuzie et al. (2009), focus groups can establish a non-threatening setting for many participants, which permits them to discuss their views, opinions, and thoughts without restraint. Thick and rich data were also obtained through the teachers’ journal entries, where they recorded their thoughts and reflections regarding boys’ and girls’ literacy performance. We included journal entries as another means of collecting data because it has been shown, in numerous studies, that the teachers’ reflections can encourage the creation of meanings for new ideas and insights, rooted in their practices (Jarvis, 1992; Moon, 2003; Ortlipp, 2008). It was also aimed at triangulating the data obtained from the State Department of Education and the interviews.

Results

Data on boys' and girls' literacy scores for the state of Selangor, which approximately amounted to around 85,000 pupils (Please refer to Table 2), were obtained from the whole population of Grade1 students in two phases, Screening 1 (March to April) and Screening 2 (September to October), and comprised the following categories. Those who: i) had not mastered constructs 1-2, ii) had not mastered constructs 3-12, and iii) had mastered constructs 1-12 (Ministry of Education Malaysia, 2013). Students who have not mastered the first two basic constructs are considered very weak and are taught following the LINUS modules prepared by the Ministry, while those who have mastered all the twelve constructs are placed in the mainstream classes, and taught following the standard syllabus and textbooks. Those who have not mastered constructs 3-12 will be taught using either the LINUS module or the standard module, based on its suitability as judged by their teachers (Ministry of Education Malaysia, 2013). It is also necessary to point out that although literacy in general includes reading and writing, in the LINUS programme, the writing section only involves simple and basic writing, given that it is meant for Grades 1 to 3 (e.g. writing syllables, words, and simple sentences).

How do primary school boys in the state of Selangor perform English literacy relative to girls?

In the first screening, as was the case for the language of instruction, Bahasa Malaysia, (Mohd Asraf & Hazlina, in press), more girls were found to have mastered Constructs 1-12 (62.55%) compared to boys (51.75%); a substantial difference of about 11%. There were also more boys who had not mastered Constructs 1-2 (10.91%) and Constructs 3-12 (37.34%) than girls (7.23% and 30.22% respectively).

Table 2
Grade 1, Screenings 1 and 2 English Language Literacy Performance 2013

| | Screening 1 N = 85,104 (Boys = 43,869; Girls = 41,235) | | | Screening 2 N = 84,852 (Boys = 43,675; Girls = 41,177) | | |
|--------------------------------------|--|--------------------|--------|--|-------------------|--------|
| | Boys | Girls | Total | Boys | Girls | Total |
| Have Not Mastered (C1-C2) | 4,784 (10.91%) | 2,983 (7.23%) | 7,767 | 4,802 (10.99%) | 2,571 (6.24%) | 7,373 |
| Have Not Mastered (C3-12) | 16,379 (37.34%) | 12,459 (30.22%) | 28,838 | 11,865 (27.17%) | 8,499 (20.64%) | 20,364 |

| | | | | | | |
|----------------------------------|--------------------|--------------------|--------|--------------------|--------------------|--------|
| Have Mastered (C1-12) | 22,706 (51.75%) | 25,793 (62.55%) | 48,499 | 27,008 (61.84%) | 30,107 (73.12%) | 57,115 |
| Total | 43,869 (100%) | 41,235 (100%) | 85,104 | 43,675 (100%) | 41,177 (100%) | 84,852 |

In the second screening, it was also found that the boys were lagging behind the girls, although the percentage was lower. About 38% of the boys had not mastered the English language constructs (10.99%: Constructs 1-2, and 27.17%: Constructs 3-12) compared to 26.88% of girls (6.24%: Constructs 1-2, and 20.64%: Constructs 3-12). The results of the second screening also showed a sizeable increase (of about 10%) in the number of both boys and girls who had mastered all the literacy constructs, suggesting quite a substantial growth in their English literacy performance after a term in school. The results substantiate the assertions of Bonomo (2010), Gurian, Henley & Trueman (2001), Gurian & Stevens (2004), and King & Gurian (2006) that girls are better at languages, as they have “more cortical areas devoted to verbal functioning, they are better at sensory memory...tonality, and the complexity of reading and writing” – all of which are skills highly emphasized and valued in school (Bonomo, 2010, p. 258). However, the teachers, as determined from the focus group interview, seem to disagree with the fact that girls are necessarily better at languages than boys. Instead, they believed that girls are more hardworking and more concerned than boys about getting many As in examinations, which make them appear to perform better. Boys, on the other hand, are not really concerned about the number of As they get, but are able to “think more critically and more creatively”, which the teachers said made them more “street-smart”, unlike the girls, who are more “book-smart”. This, according to them, is because of the country's education system, which is very much exam-oriented. They argued that if the assessment were done differently, (e.g. via projects and presentations), boys could have performed better than girls. Additionally, some of the teachers pointed out that girls like reading more, and have more positive attitudes toward reading. Boys, on the other hand, dislike reading—they prefer playing with friends to reading. The teachers' views are consistent with the results of several earlier studies and assertions of authors (e.g. Canadian Council on Learning, 2009; Carr & Pauwels, 2006; Commonwealth of Australia, 2002; Connolly, 2004; Cuttance & Thompson, 2008). However, the teachers, nevertheless, admitted that examinations are also important as a yardstick to gauge students' understanding and capabilities.

Conclusion

The data drawn from nearly 85,000 students from Grade 1 show that the English literacy performance of primary school boys in the state of Selangor, Malaysia, is lower than that of the girls. This is consistent with other findings that have found the same pattern at different levels of education in many other countries around the globe.

The results of this study add to the body of literature on boys' underachievement in literacy, and have manifold implications for educational practice. Firstly, it calls for the development of a holistic approach toward the teaching of literacy across the entire primary school curriculum, where instruction should relate to students in terms of the content as well as methods that would foster the development of their interest in reading and writing in the second language. Relevant courses should also be introduced to help teachers provide differentiated instruction for boys and girls, and they should be made aware of the fact that boys tend to lag behind girls in literacy—especially at the early stages of schooling. Placing due emphasis on boys' and girls' differential performance in literacy at the earlier stages of primary school will also assist teachers, textbook writers, and curriculum developers to provide lessons and materials that would be suitable to their interests and needs. With this, it is hoped that we will be able to nurture a generation of *successful* readers, where no child would be left behind.

References

- Annamalai, S., & Muniandy, B. (2013). Reading habit and attitude among Malaysian polytechnic students. *International Online Journal of Educational Sciences*, 5(1), 32–41.
- Bonomo, V. (2010). Gender matters in elementary education: Research-based strategies to meet the distinctive learning need of boys and girls. *Educational Horizons*, 88(4), 257–264.
- Booth, D., Elliott-Johns, S., & Bruce, F. (2009). *Boys' literacy attainment: Research and related practice*. Retrieved from http://www.edu.gov.on.ca/eng/research/boys_literacy.pdf
- Canadian Council on Learning. (2009). *Why boys don't like to read: Gender differences in reading achievement*. Retrieved from http://www.ccl-cca.ca/pdfs/LessonsInLearning/02_18_09-E.pdf
- Cappon, P. (2011). *Exploring the "boy crisis" in Education*. Retrieved from <http://www.ccl-cca.ca/pdfs/OtherReports/Gendereport20110113.pdf>

- Carr, J., & Pauwels, A. (2006). *Boys and foreign language learning. Real boys don't do languages*. Basingstoke: Palgrave Macmillan.
- Clark, C., & Burke, D. (2012). *Boys' reading commission 2012: A review of existing research conducted to underpin the commission*.
- Commonwealth of Australia. (2002). *Boys: Getting it right*. Canberra: House of Representatives Standing Committee on Education and Training.
- Connolly, P. (2004). *Boys and schooling in the early years*. London: Routledge Falmer.
- Cresswell, J. W. (2008). *Educational research: Planning, conducting and evaluating quantitative and qualitative research* (Third Ed.). New Jersey: Pearson Education.
- Cuttance, P., & Thompson, J. (2008). Literature review of boys education report. Retrieved from http://www.victoria.ac.nz/education/pdf/whakapiki/boys_ed_lit_rev_nz.pdf
- Donahue, P. L., Finnegan, R. J., Lutkus, A. D., Allen, N. L., & Campbell, J. R. (2001). *The nation's report card: Fourth grade reading 2000*. Jessup. Retrieved from <http://nces.ed.gov/nationsreportcard/pdf/main2000/2001499.pdf>
- Examinations Syndicate, Ministry of Education Malaysia. (2012). *Laporan tahunan lembaga peperiksaan 2012*. Retrieved from <http://www.moe.gov.my/lp/files/penerbitan/BukuLaporanTahunanLP2012.pdf>
- Gurian, M., Henley, P., & Trueman, T. (2001). *Boys and girls learn differently! A guide for teachers and parents*. San Francisco: Jossey-Bass/John Wiley.
- Gurian, M., & Stevens, K. (2004). With boys and girls in mind. *Educational Leadership*, 62(3), 21–26.
- Jarvis, J. (1992). Using diaries for teacher reflection on in-service courses. *ELT Journal*, 46(April), 133–143.
- Jones, J. (2003). *Early literacy assessment systems: Essential elements*. Princeton, NJ. Retrieved from <http://www.ets.org/Media/Research/pdf/PICEARLYLIT.pdf>
- King, K., & Gurian, M. (2006). Teaching to the minds of boys. *Teaching to student strengths*, 64(1), 56–61.
- Krueger, R. A., & Casey, M. A. (2009). *Focus groups: A practical guide for applied research* (4th ed.). Thousand Oaks: Sage Publications.
- Lindsay, G., & Muijs, D. (2005). *Challenging underachievement of boys in Years 8 and 10 : A study for the Oldham Leading Edge Partnership*. Retrieved from

<http://www2.warwick.ac.uk/fac/soc/cedar/projects/completed05/oldhamboys/oldhamboys.pdf>

Ministry of Education Malaysia. (2013). *LINUS 2.0 (Literasi Bahasa Inggeris) manual am pentadbiran instrumen literasi 2013 (Saringan 2)*.

Moon, J. (2003). *Learning journals and logs, reflective diaries*. Retrieved from <https://www.deakin.edu.au/itl/assets/resources/pd/tl-modules/teaching-approach/group-assignments/learning-journals.pdf>

Morgan, D. L. (1997). *Focus groups as qualitative research* (2nd ed.). Thousand Oaks: Sage Publications.

National Literacy Trust. (2012). *Boys' reading commission*. Retrieved from http://www.literacytrust.org.uk/assets/0001/4056/Boys_Commission_Report.pdf

Noble, C., Brown, J., & Murphy, J. (2001). *How to raise boys' achievement*. London: David Fulton Publishers.

OECD. (2010). *PISA 2009 at a glance*. OECD Publishing. doi:10.1787/9789264095298-en

Onwuegbuzie, A. J., Dickinson, W. B., Leech, N. L., & Zoran, A. G. (2009). A qualitative framework for collecting and analyzing data in focus group research. *International Journal of Qualitative Methods*, 8(3), 1–21.

Ortlipp, M. (2008). Keeping and using reflective journals in the qualitative research process. *The qualitative report*, 13(4), 695–705.

Rathvon, N. (2004). *Early reading assessment: A practitioner's handbook*. New York: The Guildford Press.

Snow, C. E., & Biancarosa, G. (2003). *Adolescent literacy and the achievement gap: What do we know and where do we go from here?* Retrieved from [https://umdrive.memphis.edu/mransdll/public/Dr. Ransdell's old courses/RDNG 7544/Chapter 2 Adolescent Literacy and the Achievement Gap.pdf](https://umdrive.memphis.edu/mransdll/public/Dr._Ransdell's_old_courses/RDNG_7544/Chapter_2_Adolescent_Literacy_and_the_Achievement_Gap.pdf)

Snyder, T. D., & Hoffman, C. M. (2001). *Digest of education statistics 2000*. Retrieved from <http://nces.ed.gov/pubs2001/2001034.pdf>

TIMSS & PIRLS International Study Center. (2011). *PIRLS 2011 international results in reading: Executive summary* (Vol. 91). doi:10.1097/01.tp.0000399132.51747.71

UNESCO. (2009). *Promoting gender equality in education*. Retrieved from <http://unesdoc.unesco.org/images/0018/001864/186495E.pdf>

UNESCO Institute for Statistics. (2009). *The next generation of literacy statistics: Implementing the literacy assessment and monitoring programme (LAMP)*. Montreal, Quebec. Retrieved from <http://www.uis.unesco.org/Library/Documents/Tech1-eng.pdf>

Warrington, M., & Younger, M. (2006). *Raising boys' achievement in primary schools towards a holistic approach*. Maidenhead: Open University Press.

Weaver-Hightower, M. B. (2008). *The politics of policy in boys' education: Getting boys "right."* New York: Palgrave Macmillan.