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 — The EFL Professional's Written Forum

Table of Contents:

Book Reviews

June 2015 Foreword

by Wen-Chi Vivian Wu

First of all, I would like to welcome all the readers to the second issue of Asian EFL Journal in 2015. This issue includes six articles and two book reviews which touch upon vital topics in the field of teaching and learning English as a second/foreign language, addressing corrective feedback on writing, Grice's maxims of conversation, cross-cultural misunderstanding, stereotype threat, stigma consciousness, video dubbing tasks, computer-based testing, communicative language teaching, and test preparation. In addition to the importance of the various topics, this issue includes authors from diverse geographic regions, spanning Asia and North America, indicating the diversity of the journal. This demonstrates the vitality of the Asian EFL community, as well as the continued positive response to this journal.

Sayyah Al-Ahmad and Rasheed S. Al-Jarrah, in their study comparing the effectiveness between two feedback types (i.e., feedback only and direct feedback plus oral meta-linguistic explanation), investigated the impact of direct corrective feedback type on the linguistic accuracy of low-intermediate EFL learners' writing. With occurring errors of English past simple tense usages found in narrative story compositions based on sequential pictures, the researchers found that direct feedback along with oral meta-linguistic explanation was the most effective feedback type to enhance L2 writers' linguistic accuracy. Suggestions for future researchers who are interested in investigating related issues have also been provided by the authors.

Pino Cutrone examined cross-cultural misunderstandings between EFL speakers and native speakers of English, specifically focusing on the use of Grice's maxims of conversations for Japanese learners to bridge the sociolinguistic gap. The author stated that despite the universality of Grice's maxims of conversations, the four maxims were not culture-independent. Instead, fundamental issues underpinning cross-cultural misunderstandings between Japanese EFL speakers and native speakers of

English needed to be taken into account so that reasons for pragmalinguistic failure could be better identified. Pedagogical implications for language educators and cross-cultural communication trainers who are interested in investigating the related field in intercultural communication have also been provided by the author.

To avoid limitations from considering internalized culture or ascribed personality traits to be explanations for EFL learners' reticence, Ling-Hui Hsu probed deeply into the interplay of learner and teacher variables and identified three subtle psychological factors related to students' willingness to participate, including student stereotype threat, stigma consciousness, and teacher language attitude. The results revealed that low stereotype threat and positive teacher attitude contributed to significantly better oral communication performance. The author cautioned that student reticence should not be taken as a norm in Asian EFL setting; a carefully orchestrated communicative teaching curriculum should be implemented to have students psychologically ready for instructional engagement.

In their study investigating the effects of video dubbing tasks on improving English oral proficiency, Pin He and Sukhum Wasuntarasophit had Chinese EFL learners complete a video dubbing task in four weeks. The results from the questionnaire and interview indicated that the participants' oral proficiency (comprehensibility, fluency, and accentedness) was enhanced, and that the participants held positive attitudes towards the task, even with challenges encountered. Future improvements of the video dubbing task, pedagogical implications, and suggestions for future research are provided.

In an attempt to contribute to the insufficiency of computer-based testing (CBT) in EFL settings, Reza Dashtestani adopted a mixed-method study to examine EFL teachers' perspectives on the implementation of CBT in Iranian EFL contexts, with a questionnaire and an in-depth interview employed for data collection. The author found that while teachers held positive attitudes towards the implementation of CBT, the major challenge that CBT was not utilized and developed led Iranian teachers to ponder on strategies which might facilitate the incorporation of CBT into EFL programs. This article can serve as guidance for English instructors who are interested in integrating computerized assessment into EFL curriculum.

Sky Lantz-Wagner, in his study exploring how CLT impacted high stakes test preparation (i.e., the Test for English Majors Band 4, TEM4) in China, focused on CLT implemented during test preparation, students' perceptions of the course, and the effects of the course on TEM4. Diary entries proved to be a valuable tool in accomplishing the goal of preparing students for the TEM4 and reflective analysis yielded some assumptions and principles with which the course was designed and implemented. The participants engaged in the TEM4 test preparation course performed substantially better than those who took the test without structured preparation in the previous year. This article brings new insights for researchers and instructors who are interested in test preparation in an EFL setting.

We hope you find the articles in this June 2015 issue to be informative, inspiring, and enjoyable to read. We also hope that this issue will help provide new insights into the formulation of future research and innovations for EFL practitioners, so as to contribute to continuous improvements in English Language instruction around the world. We would like to express our greatest appreciation to three distinguished scholars (Dr. Sabrina Priego, Dr. Anne-Charlotte Perrigaud, and Dr. Michael Marek), for their constructive suggestions and insightful inputs on the proofreading to ensure overall quality of the journal. We sincerely acknowledge their investment of considerable time and effort into polishing the papers included in this issue. Special thanks also go to my research assistant, Mr. Ming-Yi Scott Chen Hsieh, for his devotion to careful formatting and overall layout of the issue. Finally, our sincere expressions of thanks are also extended to the contributors and reviewers of articles and book reviews who have made this issue possible. Their quality scholarly work and careful peer review is vital to the success of Asian EFL Journal.

The Impact of Direct Corrective Feedback Type on the Linguistic Accuracy of EFL Students' Writing

Sayyah Al-Ahmad and Rasheed S. Al-Jarrah

Yarmouk University, Jordan

Bio data

Sayyah Al-Ahmad is an Associate Professor of English, teaching English Language and Linguistics at the English Department and Language Center/Yarmouk University-Jordan. He graduated from Indiana University of Pennsylvania/USA, with PhD in Rhetoric & Linguistics in 2003. His research interests include Language Learning and Teaching, Language, Foreign Language Writing. E-mail: drsakash@hotmail.com

Rasheed Al-Jarrah is an Associate Professor of English, teaching English Language and Linguistics at the English Department and Language Center at Yarmouk University/Jordan. He graduated from Ball State University-Muncie/Indiana-USA, with PhD in Applied Linguistics in 2002. Areas of research include Second Language Writing, Optimal Solutions, Language Testing, Discourse Analysis. E-mail: rsaljarrah@yu.edu.jo

Abstract

Building on the existing literature which targets the efficiency of corrective feedback for helping L2 writers improve their writing, the present study aims to find out if there were any differential effects for the type of direct feedback on the linguistic accuracy of low intermediate EFL learners' writing. Concisely, by experimenting with two distinct classroom practices (namely direct feedback only versus a combination of direct feedback plus oral meta-linguistic explanation), the goal was to see which feedback option is more effective for our learners to retain the corrections on some linguistic structure in the course of time. Planning to target almost all occurring errors in the use of the English past simple tense in the writing tasks of three groups of EFL learners in our

Address correspondence:
Language Center & English Department,
Yarmouk University, Irbid, Jordan

feedback, the subjects were asked to produce three narrative story compositions based on sequential pictures they were provided with. Apart from the control group, it turned out that the participants who received direct feedback along with oral meta-linguistic explanation were more apt to sustain the same corrections on subsequent writing tasks.

Key words: Corrective feedback, direct and indirect feedback, focused and unfocused feedback

Introduction

Research on L2 written corrective feedback has been mainly concerned with two main topics. These are (1) what to correct, and (2) how to correct L2 students' writing. Despite the relatively huge amount of research that has accumulated so far, there is little consensus among researchers and teachers as to what type(s) of written corrective feedback should be adopted in the classroom. From the various options available to them, teachers' choice is hardly theoretically motivated. As each option has its own limitations, teachers' pedagogical decision is more constrained by the adverse realities of the learning environment than by clear-cut answers from carefully designed studies. In order to bridge the gap between the theory and the practice, one research inquiry should then be on which corrective feedback technique is most effective for correcting students' writing in a given context.

This study was intended to investigate the relative effectiveness of just two types of direct feedback on the accuracy performance of low-intermediate EFL learners in new pieces of writing in a particular L2 learning context. The two direct feedback options used were (1) providing some students with direct feedback only, and (2) providing others with a combination of direct feedback plus some oral meta-linguistic explanation.

The current research endeavour, however, makes at least two main departures from the bulk of research completed to date. First, the study was designed to target only one linguistic structure, namely the simple past tense. Probably with the exception of Sheen's and Bitchener's studies (namely Sheen, 2007a; Bitchener, 2008; Bitchener and Knoch, 2009; Bitchener and Knoch, 2010), most studies on feedback options (e.g. Truscott, 1996; Ferris, 2002; 2003; Chandler, 2003; Bitchener, Young, and Cameron, 2005; Cañado and Bedmar, 2006; Evans, Hartshom, and Tuioti, 2011) have adopted the unfocused approach to corrective feedback, i.e. providing a combination of several error categories on L2 students writing. One reason why we have chosen to adopt this focused approach to error correction is that experimenting with different error categories (e.g. past simple and conditional structure such as

9

I'd help you if I were in a position to help) could be a possible cause of a whole set of error types (see Al-Jarrah & Al-Ahmad, 2012). Another reason is that the focused approach could enhance, at least momentarily, learning because, according to Ellis (2008: 102), "the more intensive the attention, the more likely the correction is to lead to learning."

Second, although we adopted the focused approach, we tried to provide corrective feedback on all the functional uses of the targeted linguistic structure. This clearly distinguishes our study from Bitchener's (2008) and Bitchener's and Knoch's (2010) studies which focused on some 'treatable' functional uses of the linguistic structure under scrutiny.

On the whole, this study was then intended to address the following two intertwined research inquiries:

- Is linguistic accuracy improvement in the use of the English simple past tense determined by the type of corrective feedback provided?
- Is this accuracy improvement in the use of the English simple past tense extended and retained over time in new pieces of writing?

Literature Review

Despite the strong voices which advocate the view that corrective feedback is ineffective, and could possibly be harmful (Truscott, 1996; Polio, Fleck, and Leder, 1998; Truscott, 1999; Fazio, 2001; Truscott, 2007), a sizable portion of research has set itself the aim to invalidate this claim (e.g. Ashwell, 2000; Ferris, 1999; Ferris and Roberts, 2001; Chandler, 2003, 2004; Russell and Spada, 2006; Bitchener, 2008; Evans, et al, 2011). In order to do just this, their aim was therefore twofold. First, they needed to show that corrective feedback has a significant positive effect on learners' abilities to write accurately. For Truscott, the effect, if any, could be attributed in part (or possibly on whole) to some external forces such as research design, including the population's level of proficiency, procedure of data collection, longitudinal vs. cross-sectional research designs, etc. (Cf. Ferris, 2004; Guenétte, 2007). Extraneous variables such as classroom instruction, the strategy of avoidance that some students use, maturation of the population, etc. may also have had some effect on research findings, and could therefore have biased the findings of some studies towards slight gains of grammatical accuracy (for an illuminating discussion, see Truscott, 2007). Second, they needed to provide empirical evidence to prove that the gains obtained from the provision of corrective feedback can stand the test of time. This was badly needed because a number of studies (e.g. Ashwell, 2000; Semke, 1984; kepner, 1991;

Fathman and Whalley, 1990; Sheppard, 1992; Polio et al., 1998; Fazio, 2001; Ferris and Roberts, 2001; Chandler, 2003; and Truscott and Hsu, 2008) have shown that corrective feedback is only momentarily constructive.

However, one major area of dispute among researchers who experimented with corrective feedback is concerned with which error correction technique is the most effective (Cf. Lalande, 1982; Robb, Ross, and Shortreed, 1986; Ferris and Roberts, 2001; Sheen, 2007a; Sheen, Wright, & Moldawa, 2009; Bitchener, Young, & Cameron, 2005; Bitchener, 2008; Bitchener and Knoch, 2009; Bitchener and Knoch, 2010). To them, part of the problem then lies in how corrective feedback is administered (Cf. Cohen and Robbins, 1976; Hyland, 2003, Bitchener, et al., 2005). This is what we call the treatment which, to Guénette (2007: 13), is considered "the crux of the matter". For a number of researchers, the lack of positive effects could be attributed in part (or on whole) to "inconsistencies in research design", rather than to the corrective feedback provided to the learners (for an illuminating discussion see Ferris, 2004 and Guénette, 2007). Therefore, the feeling has always been that for error correction to be maximally effective, it has to be "the appropriate feedback, given at the right time and in the proper context," to use Guénette's (2007: 11) words.

However, despite all caution researchers take to keep the independent and intervening variables constant (so that the effect of error feedback is not attributed to these parameters), their findings showed conflicting evidence not only for the efficacy of error correction (whether feedback is/isn't effective), but also for the type of feedback. When experimenting with different types of indirect feedback, Ferris and Roberts (2001) and Robb, Ross, and Shortreed (1986), for example, found that they all have relatively equal impact. However, Bruton (2009) believed that "there is no guarantee that the students revisions will be correct" if only indirect feedback is provided. Lalande (1982), Frantzen (1995), Ferris, Chaney, Komura, Roberts, and McKee (2000), Ferris and Hedgcock (2005) argued that indirect feedback could be more effective than direct feedback. However, Chandler (2003) found completely the opposite. Findings of research on oral corrective feedback in second language learning contexts support this view (see Carroll and Swain, 1993; Ellis, 1998; Ellis, Lowen, and Erlam, 2006). Similar findings had been reported in Lalande (1982) and Robb et al. (1986). Still, Robb, et al. (1986), Semke (1984), and Ferris and Roberts (2001) found no significant differences between the two correction methods.

At greater levels of detail, it turned out that all correction methods (including coded and uncoded) were relatively equally effective in one of Chandler's (2003) experiments. Several studies investigated

the effect of meta-linguistic feedback in combination of other types of feedback on improving students' writing performance. Sheen (2007a) compared direct feedback alone versus direct feedback in combination with meta-linguistic explanation and found that the latter approach was more effective than the former because the meta-linguistic feedback could, according to her, advance "a deeper level of cognitive processing" (see Sheen 2007a: 260). Lyster (2004) claimed that meta-linguistic feedback was more effective than recasts. Sheen's (2007b) study showed that oral recasts and oral meta-linguistic feedback resulted in differential effects on acquiring English articles with oral meta-linguistics corrective feedback but not recasts facilitating learning. He added that meta-linguistic corrective feedback helped learners improve the accuracy of their use of English articles because when receiving a rule, and sometime for processing information, learners were able to engage in the cognitive comparison hypothesized to facilitate learning. Sheen (2010) argued that oral meta-linguistics corrective feedback was as effective as written meta-linguistic feedback in facilitating the acquisition of English articles. Li (2010) and Lyster and Saito (2010) reported higher effect size for explicit feedback type (e.g. explicit correction and meta-linguistic feedback). Ellis et al. (2006) found that meta-linguistic feedback (explicit feedback) was more effective than recasts (implicit feedback) in acquiring regular past tense. He also indicated that meta-linguistic feedback helped the development of implicit and explicit feedback of L2 learners.

One interesting finding of Bitchener et al. (2005) is that the type of feedback was not equally effective for specific structures. To illustrate, whereas those who received direct error correction and oral meta-linguistic explanation outperformed those who did not for the past simple tense and the definite article, they failed to do so for prepositions. Their explanation was like this: whereas the past tense and the definite article structures are rule-based, prepositional forms are more idiosyncratic. Al-Jarrah & Al-Ahmad (2012) questioned Bitchener et al.'s (2005) finding as it does not explain why some linguistics forms can be rule-based while others cannot. Added to this, Bitchener and Knoch (2009) found that there were no differential effects for the different types of direct feedback on the linguistic accuracy of low intermediate ESL learners' writing for two functional uses of the English article system (namely referential indefinite ''a" and referential definite ''the"). What this basically means is that, as Guénette (2007: 40) rightly points out, "the results of the many experimental studies on written corrective feedback carried out over the last 20 years have been so contradictory."

Methodology

Participants and Setting

This study was conducted at Yarmouk University, Jordan. Thirty-nine EFL undergraduate students participated in this study over a semester of 16 credit hours of writing instruction. The participants were all second-year English majors taking a required writing course in the English Department. Their English language proficiency can be rated as low-intermediate, or as Basic Users A2 (Waystage), with reference to Common European Framework (p. 24), according to which, they

- can understand sentences and frequently used expressions related to areas of most immediate relevance (e.g. very basic personal and family information, shopping, local geography, employment)
- can communicate in simple and routine tasks requiring a simple and direct exchange of information on familiar and routine matters
- can describe in simple terms aspects of his/her background, immediate environment and matters in areas of immediate need"

On the whole, the participants' foreign language proficiency skills were judged to be homogeneous in at least three ways. First, they all had the same native language background in that they were all native speakers of Arabic with an average of 10 years of English language instruction in a foreign environment. Second, most of them had approximately similar scores in the General Secondary Certificate Examination (high school graduation exam) which is an indispensible condition for admission to the program. Third, their writing capabilities were assessed in terms of their general English language proficiency level by the researchers themselves. Therefore, the participants were divided and arbitrarily assigned to three groups, two experimental and one control, with thirteen students in each group. The writing task was carried out in class that met three times a week, each of which lasted for 50 minutes.

Design

The linguistic accuracy in the use of the past tense form by the participants was measured by means of a pre-test, immediate post-test, and delayed post-test over one semester. The participants were divided into three treatment groups:

- Experimental (1) received direct full explicit written corrective feedback only above each targeted error
- Experimental (2) received direct full explicit written feedback as well as oral meta-linguistic explanation
- Control received no corrective feedback at all on the targeted feature, but to meet ethical requirements, the participants in this group were given general content feedback on the organization and quality of their work like "well-organized", "nice ideas", etc.

As for the direct written feedback, learners were provided with the correct form for each past simple tense error they made in their pieces of writing by means of deletion (crossing out unnecessary morpheme), addition (inserting a missing morpheme) or replacement (crossing out the whole erroneous form and writing the correct form above it). However, although research has shown that unlike indirect feedback options which could lead to long term learning by getting the learners engaged in a problem-solving process of detecting and correcting the errors they make for themselves (e.g. Ferris, 2002), this feedback option may not contribute to long-term retention, but still saves the instructor time and effort to get the learners detect the erroneous forms for themselves. This is generally manifest in at least two ways: (1) many of our learners could not detect the erroneous forms they themselves have made; and (2) the majority of them could not provide the correct alternative form for the detected errors. This is probably because as low intermediate EFL learners, our subjects have a relatively limited linguistic repertoire to draw on.

As for the oral meta-linguistic explanation, the researcher instructor made brief oral comments on the simple past tense erroneous forms that the student writers had made. When needed, some additional illustrative examples were provided on the whiteboard and discussed with the class. The belief was that these comments might assist the learners to self-correct by evoking previous background information and relating it to the current writing task. Research has shown that although oral meta-linguistic explanations are more time and effort consuming on the part of the instructor, they may be more rewarding on the long run for the learner especially in new pieces of writing.

A point worthy of mention is that one reason why the writing researchers chose to experiment with these two corrective feedback options (namely direct feedback and a combination of direct feedback plus meta-linguistic explanation) is that they are the most common error feedback techniques in this learning environment. Practical constraints that teachers face (for example, large classes, heavy

workloads, and tight teaching schedules) make these feedback techniques most common in this learning environment (see Al-Jarrah and Al-Ahmad, 2013). Another driving force was that very few studies have compared the effects of these two corrective feedback types in promoting acquisition of specific grammatical structures such as the past simple tense.

Targeted linguistic error

Due to the positive results of recent written CF studies (Bitchener, 2008; Bitchener, and Knoch, 2008a, 2008b, 2009, 2010; Ellis, Sheen, Murakami, and Takashima, 2008; Sheen, 2007a; Sheen, Wright, & Moldawa, 2009) and oral CF studies (Carroll & Swain, 1993; Ellis et al, 2006; Muranoi, 2000) in targeting the use of a single specific linguistic form, it was decided by the researchers to choose the focused approach to providing corrective feedback on each piece of students' written assignments. The targeted structure was the use of simple past tense in a narrative task. This was targeted because (1) the researchers noticed that many student writers make errors in the use of the English simple past form; (2) the past simple has numerous forms (e.g. copula verbs, regular and irregular forms, negative past simple) which pause a challenge to learners with limited language proficiency; (3) the past simple tense is a relatively functionally complex structure especially when used in combination with other structures such as the conditionals (for example, *if I were you*).

It should be noted, however, that previous research has attempted to draw a distinguishing line between focused and unfocused corrective feedback. Whereas the more common unfocused classroom feedback practice target multiple errors made by the learners simultaneously, focused feedback targets very few functional uses of some specific structure(s) at a given time; and therefore dealing with the learners' occurring errors as if they were unrelated lists of grammatical features. Another problem was that the terms (focused and unfocused CF) were operationalized in a number of ways by different researchers. Whereas some selected one or two error types of some specific structure (e.g. the definite article for first mention and the definite article for subsequent mention), others experimented with a whole range of error types of seemingly unrelated structures (e.g. simple past tense, prepositions, articles). Not only this, but researchers were also fastidious in their choice of the functional uses for each specific target structure. For example, Sheen et al. (2009) chose to target articles, past tense, and prepositions. However, among the functional uses of the articles, they choose to provide corrective feedback on referential indefinite 'a' and referential definite 'the'; for the past tense they choose to

experiment with copula 'be', regular past tense 'ed', and irregular past tense; and for prepositions their corrections targeted only temporal and locative prepositions (namely at, in, on).

Unlike possibly all previous research, focused corrective feedback is operationalized in the current study to refer to *all* the functional uses of *one* specific structure (namely the past simple tense). The theoretical motivation behind this choice is that because the acquisition of language involves subtle processes that require more than just a collection of discrete items, there must be global approaches towards corrective feedback that establish a meaningful whole (roughly equivalent to, though not identical with, Sheppard's (1992) holistic comments). For the effect of the feedback, as Sheen et al. (2009) acknowledge, should not only be noticeable just by the improved accuracy rate on the targeted structure only, but on "a broader range of grammatical structures" (see Sheen et al., 2009: 259). On the more practical plane, it is probably impossible in a longitudinal study of a writing course to keep classroom practice focused on one or two error types of a single linguistic structure (Cf. Ellis, Sheen, Murakami, & Takashima, 2008; Sheen et al., 2009). In doing so, we are therefore echoing the concern of Ferris (2010: 188) who stressed that "a heavy emphasis on a few narrowly drawn structures in instruction and feedback would seem too limited a focus for a writing class".

Treatment

Experimental group (1) received direct error correction only above the targeted feature. Experimental group (2) received direct written corrective feedback plus oral meta-linguistic explanation. The control group did not receive formal instructions about the targeted structure but instead received some content related feedback notes. It was anticipated that errors might take place in at least three ways: (1) using the wrong past tense structure (i.e. past continuous or past perfect instead of past simple structures); (2) using the wrong simple past tense form (e.g. *costed* instead of *cost*); and/or (3) using the wrong tense form (i.e. present simple or present continuous instead of past simple). For clarification, writing samples are provided in Appendices (1, 2, 3)

In the oral meta-linguistic explanation, the participants in experimental group (2) were given a 30-minute mini-lesson in which the researcher explained the rules which govern the use of the English simple past tense. Additional examples were provided on the white board and discussed with students. This was followed by a short "controlled practice" exercise in which students were asked to fill in the blanks with the correct form of simple past (see Appendix 4). The students were given six minutes to complete the task. Again, the control group received no such feedback on the targeted feature.

The oral meta-linguistic explanation was only given in the treatment session that took place two weeks after the pre-test and on the same day as the immediate post-test. Direct error correction above the targeted feature was provided on all the writing tasks (pre-test, immediate post-test, and delayed post-test) for students in experimental groups (1) and (2). Group (3) did not receive any corrective feedback on the targeted error.

What is worth mentioning here is that although the researchers were involved in teaching, correcting, and commenting on all the groups' written tasks, they never volunteered to give explicit instruction on the targeted linguistic feature between the writing tasks. However, it was difficult to account for extraneous factors that might have taken place outside the classroom such as receiving additional input.

Instruments

The writing students were given three writing tasks on three different occasions. In these three testing occasions (pre-test, immediate post-test, and delayed post-test), the participants were asked to narrate what was happening in the pictures provided. Task one for the pre-test consisted of six sequential pictures (see Appendix 5). In task two for the intermediate post-test, the subjects had eight sequential pictures to narrate from (see Appendix 6). In task three for the delayed post-test, the writing students were provided with six sequential pictures (see Appendix 7). In each writing task, the students were asked to write a coherent story based on these pictures which the researchers believed would predispose the students to using the targeted structure. Besides, word prompts were provided under each picture to help students elicit as many phrases as possible within the specified writing time. Not only this, but they were also allowed to ask questions about vocabulary that they think they need to narrate about the pictures. Due to the nature of the writing tasks which required the telling of a story that took place in the past and due to the proficiency levels of the participants who were basically low-intermediate EFL learners, the strategy of avoidance was not expected to be used effectively by the participants. The students were given about forty minutes to complete their writing tasks.

Procedure

Three days before the pre-test, the participants were briefed on the study and were given the opportunity to ask questions about it. The data were collected by the researcher instructor in the following way:

1. On the first day of class, the pre-test was administered.

2. Ten days later, the texts with written corrective feedback were returned to students in experimental groups (1) and (2). Students were then given five minutes to look at the corrections. Experimental group (2) was given a thirty-minute oral meta-linguistic feedback lesson on the targeted structure.

3. Immediately after this mini lesson, students were asked to write a second writing task (immediate post-test). Students in experimental group (1) were given five minutes to look over the corrections that had been provided on their pre-test. Then they were asked to do the second writing task. The control group did the second writing task as soon as the pre-test writing assignments were returned to them.

4. The second writing task was returned to them five days after it had been written. Experimental group (1) received only direct error correction above the errors made by students in the group, but no oral meta-linguistic explanation was provided to this group. Experimental group (2) received direct corrective feedback as well as oral meta-linguistic explanation.

5. The delayed post-test was administered two months later. The students were not informed that there would be a delayed post-test so that they would not check on the corrections made in the pre-test and post-test. The third writing task was returned to the students five days later. The researcher did not provide any form of corrections or explanation during the interim period.

Scoring guidelines

In scoring the results of the written narrative tests for all three groups, the linguistic accuracy score was calculated as follows: The researcher instructors counted both the number of correct and incorrect uses of the targeted linguistic form (simple past tense). Then they counted the number of the correct uses of this grammatical form. Finally, the total number of the correct forms was divided by the total number of the correct and incorrect forms to obtain the accuracy score for each student in the form of a ratio of the correct to incorrect use of the targeted feature (see Bitchener et al., 2005).

Data Analysis

One of the researchers, a writing instructor for a about ten years made error identification and correction on the students' texts. A month later and after the initial scoring, the same researcher rescored the texts to examine the reliability of the writing test. The texts were also scored by another experienced writing instructor. The inter-rater reliability was 93%. Statistical Package for the Social

Sciences (SPSS) was implemented to account for descriptive and referential statistics for each of the three tests. ANCOVA and Bonferroni's tests for post-hoc comparisons were used to address the two research questions.

Results

This part displays the results of examining the effect of different types of corrective feedback on improving the accuracy scores of the simple past tense as used by the participants, and whether this effect continues to exist in new pieces of writing over time. The researcher instructor counted the simple past tense errors made by students in the three groups (direct feedback group only, direct combined with oral meta-linguistic feedback, and control group) before and after treatment as shown in Table 1 below.

Table 1

Number of errors made by students in the three groups

Group														
Direct Feedback Only					**Direct and Oral Meta-Linguistic Feedback**					**Control**				
id	pretest	Immediate Post	Delayed Post	**Total**	**id**	pretest	Immediate Post	Delayed Post	**Total**	**id**	pretest	Immediate Post	Delayed Post	**Total**
1	6	7	4	**17**	**1**	3	6	4	**13**	**1**	5	3	2	**10**
2	8	6	3	**17**	**2**	5	3	2	**10**	**2**	9	14	6	**29**
3	3	5	6	**14**	**3**	4	1	1	**6**	**3**	11	9	14	**34**
4	15	2	4	**21**	**4**	5	1	2	**8**	**4**	16	11	14	**41**
5	10	4	2	**16**	**5**	3	2	1	**6**	**5**	13	13	11	**37**
6	9	8	5	**22**	**6**	9	2	6	**17**	**6**	13	10	12	**35**
7	4	3	10	**17**	**7**	2	1	1	**4**	**7**	6	16	11	**33**
8	3	4	4	**11**	**8**	5	4	2	**11**	**8**	8	9	23	**40**
9	2	3	4	**9**	**9**	2	1	1	**4**	**9**	14	18	8	**40**
10	2	6	7	**15**	**10**	2	1	0	**3**	**10**	11	17	17	**45**
11	10	3	8	**21**	**11**	9	3	2	**14**	**11**	16	13	7	**36**
12	6	6	7	**19**	**12**	1	0	0	**1**	**12**	17	18	13	**48**
13	4	2	7	**13**	**13**	5	3	7	**15**	**13**	12	21	12	**45**

To answer the first research question, "*Is linguistic accuracy improvement in the use of English simple past tense determined by the type of corrective feedback provided?*", the means and standard deviations of the students' scores in the pre-test and immediate post-test were calculated as shown in Table 2.

Table 2

Group means and standard deviations of the accuracy of past simple tense (Immediate post-test)

Group	N	Pretest- (Covariate)		Immediate Posttest			
		Mean	Std. Dev.	Mean	Std. Dev.	Adj. Mean	Std. Error
Direct written feedback only	13	60.348	21.54	74.312	11.36	74.312	4.08
Direct written & oral meta-linguistic feedback	13	82.924	11.38	88.662	11.80	88.664	5.59
Control group	13	28.899	13.88	32.521	18.85	32.519	5.90

Table 2 shows the descriptive statistics for mean scores on the narrative writing tasks, which measured the accuracy of English simple past tense over two testing periods (pre-test, and immediate post-test) for experimental group (1) which received direct feedback only; experimental group (2) that received a combination of direct feedback and an oral-meta-linguistic explanation; and the control group which received no feedback. The analysis shows that there were differences between the students' mean scores on the pre-test and those on the immediate post-test in the three groups. In order to find out whether the differences in the groups' scores over time were statistically significant, ANCOVA test was used as illustrated in Table 3.

Table 3

ANCOVA test results (Immediate post-test)

Source	Sum of Squares	Degree of Freedom	Mean Square	F	Sig.	Partial η^2
Pretest (Covariate)	0.000	1	0.000	0.000	1.000	0.0%
group	7780.191	2	3890.095	**18.198**	0.000	51.0%
Error	7481.878	35	213.768			
Total	29600.191	38				

Accordingly, there were significant differences between the students' pre-test and immediate post-test scores (p < 0.05). To uncover which groups had these significant differences, Bonferroni's post-hoc pairwise comparisons were implemented as displayed in Table 4 below.

Table 4

Results of Bonferroni's test for post-hoc comparisons (Immediate post-test)

Group		Control group	Direct written feedback only	Direct written & oral meta-linguistic feedback
Bonferroni	Adj. Mean	32.519	74.312	88.664
Control group		32.519		
Direct written feedback only		74.312	**41.793**	
Direct written & oral meta-linguistic feedback		88.664	**56.145**	14.352

This analysis reveals that at the time of the immediate post-test, students in the two treatment groups (direct feedback only versus a combination of direct feedback and meta-linguistic explanation) outperformed those in the control group, but neither of these groups performed better than the other. In other words, the students' linguistic accuracy in the use of simple past tense improved significantly right after the treatment had been provided.

To address the second question, "*Is this accuracy improvement in the use of the English simple past tense extended and retained over time in new pieces of writing?*", the means and standard deviations of the students' scores on the immediate and delayed post-tests were calculated as shown in Table 5.

Table 5

Group means and standard deviations of the retention of accuracy over time

Group	N	Immediate (Covariate)		Posttest- Delayed Posttest			
		Mean	Std. Dev.	Mean	Std. Dev.	Adj. Mean	Std. Error
Direct written feedback only	13	74.312	11.36	68.874	14.98	65.483	4.41
Direct written feedback & oral meta-linguistic feedback	13	88.662	11.80	90.290	8.40	81.579	5.77
Control group	13	32.521	18.85	39.895	20.84	51.996	6.96

Accordingly, there were differences between the immediate post-test and delayed post-test scores in the three groups. While the group receiving the combination of direct feedback and oral meta-linguistic explanation showed an increase in terms of mean percentage at the time of delayed post-test, the direct written feedback group showed a decrease. In order to examine if the differences among the groups' scores over time were statistically significant, the ANCOVA test was performed as shown in Table 6.

Table 6

Results of ANCOVA Test

Source	Sum of Squares	Degree of Freedom	Mean Square	F	Sig.	Partial η^2
Immediate post-test (Covariate)	1028.164	1	1028.164	**4.660**	0.038	11.8%
group	1878.501	2	939.250	**4.257**	0.022	19.6%
Error	7722.148	35	220.633			
Total	25381.948	38				

Table 6 shows that these differences between the immediate post-test and delayed post-test scores were statistically significant (P < 0.05). Bonferroni's post-hoc comparisons were performed to find out which groups experienced these significant differences (see Table 7).

Table 7

Results of Bonferroni test for post-hoc comparisons

Group		Control group	Direct written feedback only	Direct written& oral meta-linguistic feedback
Bonferroni	Adj. Mean	51.996	65.483	81.579
Control group	51.996			
Direct written feedback only	65.483	13.487		
Direct written & oral meta-linguistic feedback	81.579	**29.583**	**16.096**	

Accordingly, in the delayed post-test, the combination of direct and oral meta-linguistic feedback group not only outperformed the control group but also the direct written feedback group. Another note worthy of mention here is that there were no significant differences between the direct group's delayed post-test scores and the control group's scores. In short, the linguistic accuracy in the use of the simple

past group (2) was extended and retained in the third piece of writing while that of the direct and control was not.

Discussion

The limited research on the efficacy of corrective feedback and its conflicting findings has stimulated further investigations and theory-building. To illustrate, some research that targeted the effect of error correction on improving L2 student writing reported that grammar correction did not have a positive effect on the development of L2 writing accuracy (Semke, 1984; Rob et al., 1986; Kepner, 1991; Sheppard, 1992; Polio et al., 1998; Truscot, 1996, 2007). Research which aimed to counteract this claim (e.g. Sheen, 2007a; Bitchener, 2008; Bitchener and Knoch, 2009, 2010) still produce conflicting evidence. Our current research is an attempt to aid this debate in one way or another.

Based on the findings of this study, it turned out to us that the researchers' inability to settle the dispute over which feedback option is most effective is operational for at least two reasons. On the one hand, they have not crystallized their definitions of terms. On the other, they have not sharpened the tools to put them to test. For example, according to Ellis, et al. (2008: 355), direct feedback and indirect feedback are defined in very general terms. To them, whereas direct feedback refers to all corrective feedback that "supplies learners with the correct target language form when they make an error", indirect feedback "refers to various strategies (e.g. simply indicating errors) to encourage learners to self-correct their errors". To Bitchener and Knoch (2010), the distinction could be made more precise by having the direct feedback refer to the explicit corrections provided above or near the linguistic error itself, and can be done by either "(1) underlining or circling an error and (2) recording in the margin the number of errors in a given line". Although the division between the two broad categories of direct feedback and indirect feedback seems relatively standard, the case is not crystal clear for meta-linguistic feedback. According to Ellis et al. (2008: 355), it "involves providing some kind of meta-linguistic clue as to the nature of the error that has been committed and the correction needed". According to Ferris, 2003, coding an error in the margin is regarded as indirect feedback, but to Bitchener and Knoch (2010: 209) it is direct feedback "because it supplies additional meta-linguistic information about the type of error from a linguistic perspective". To them, what this basically means is that by providing meta-linguistic feedback, the instructor helps the learners to understand the nature of the error, rather than just drawing the attention of the learners to the error and leaving it for them to resolve the problem for themselves. Because it is an explicit corrective feedback strategy, meta-

linguistic feedback is, according to Sheen (2010), regarded as direct feedback provided that the correct form is supplied along with the grammatical information. If only some grammatical information is provided with no provision of the correct form, meta-linguistic feedback is considered indirect feedback. However, a number of studies (e.g. Lyster, 2004; Ellis et al., 2006; Sheen, 2010) had drawn a distinguishing line between meta-linguistic feedback on the one hand and recasts on the other, depending on the observation that the former is more explicit than the latter.

What all this basically means to us is that the boundaries between the feedback strategies available to practitioners are not always clear cut. The directness (or the degree of explicitness) of the feedback options should be seen as a continuum rather than as independent strategies that fall into discrete larger units such as those detailed in Ellis's (2009) "*A typology of written corrective feedback types*". This is probably so because within the same category, some strategies can be more (or less) direct than others. For example, a strategy such as indicating plus locating the error is less indirect than locating the error only, though both options are often categorized under the umbrella term of indirect feedback. Likewise, the use of a written meta-linguistic code on the margin of the page is definitely less direct than using a written meta-linguistic code and locating it above or near the erroneous form. Running comparisons between (and drawing conclusions from) studies which use these strategies relatively loosely is, the argument goes, an analytical flaw that could misguide classroom instructors. According to (Ellis et al., 2008: 355), one good reason to believe why studies on different feedback options still produce mixed findings is due to the observation that "these studies operationalized the direct and indirect **CF** in very different ways". Therefore, for statistical significance tests to be valid, they should be run on findings which would reflect the differential effect of well-defined classroom practices, rather than wholesale grouping of technical terminology such as direct vis-à-vis indirect CF. Towards this end, our current study investigated the effect of basically two feedback options that can be split fairly distinctly as classroom practices, and provided tangible experimental evidence to practitioners that these distinct classroom practices do have differential effects on the accuracy performance of one targeted linguistic error in new pieces of writing.

The relative 'operationalization' of terms has extended to the division between the focused and unfocused approach to corrective feedback. Unlike possibly all previous studies which used the 'highly' focused approach to error corrections by targeting very few treatable errors of one specific structure (e.g. referential indefinite ''a" and referential definite ''the"), and unlike probably all previous studies which used the 'loosely' unfocused approach to error corrections by targeting a few functional uses of a

combination of seemingly unrelated structures (past simple tense, prepositions, and articles), the present study defined its approach in a more practically rigorous fashion: targeting all the functional uses (and therefore providing corrective feedback on all occurring errors) of just one specific structure (the past simple tense). Following Al-Jarrah & Al-Ahmad (2012), the study reported above adopts the approach which argues against segmenting error corrections on the targeted structure(s) on students' written work. By examining corrective feedback on few functional uses of some structure(s) at a given time, most previous investigations into the efficacy of error correction targeted errors as if they were unrelated lists of structures (for a discussion of this point see Ferris, 2010), a state of affairs which made us choose to totally reject Bitchener et al.'s (2005) line of demarcation between rule-based structures and idiosyncratic forms or Ferris's (2002) distinction between more treatable rule-based grammatical features (e.g., definite article and past-tense) and less treatable item-based features (e.g. many prepositions). For us, a theory of language learning will definitely have more explanatory power when it can pull things together (for a discussion on how to prioritize the treatment, see Al-Jarrah & Al-Ahmad, 2012).

Until some evidence emerges as to (1) how to split our feedback options, (2) how to experiment with a combination of them successfully, (3) which structures should be targeted, and (4) how many functional uses of each structure should be tested at one time, we believe the comparison should be made only between tangible classroom practices. For, irrespective of how the issue is debated at the theoretical plane, what is of utmost concern to writing teachers, the deliverers of feedback, is to see which classroom practice(s) from the various options available for correcting students' writing work(s) best with their own students.

In addition to these internal variables, there is another whole set of external variables that researchers may not succeed to deal equally with, and could remain beyond their control and vigilance (for more details, see Guénette, 2007). Our research endeavor was an attempt to make findings-based conclusions and generalizations that could be credited to the effect of the feedback practice itself rather than to some external detrimental factors such as the adverse reality of the context of the study, previous instruction in L2, amount of exposure to L2, Learners' goals, nature of the written task, and instructor's influence.

On the whole, conducting a field investigation on EFL learners who were highly homogenous in terms of the amount of previous and current exposure to English as a foreign language, their relative linguistic proficiency level in L2, their learning capabilities, their previous instruction in English as a

foreign language, their goals, motivation expectations, etc. in learning English should enable us to draw conclusions from group scores rather than individual performances. According to Sheen et al. (2009: 558), measuring the effects of CF in different ways "makes it very difficult to compare results and reach any conclusions". In Guénette's (2007: 3) words, "if we are looking to compare the efficacy of teacher feedback across studies, proficiency levels have to be carefully measured and reported." Otherwise, it could be difficult to determine whether improvements in performance are attributable to the feedback itself when in fact the findings come from groups that are not comparable in terms of their proficiency levels.

That being said, the present study provided an extra piece of evidence against Truscott (1996) who believed that correcting learners' errors in a written composition may not help improve the learners' grammatical accuracy in a new piece of writing although it may reduce the number of errors in a subsequent draft on the same topic. Not only this, but it also provided counter evidence against those (e.g. Ferris and Roberts, 2001; Robb et al., 1986) who believed that the treatment type (direct, coded, uncoded, etc.) is irrelevant. The findings not only showed the immediate effect of written corrective feedback after it had been provided in the immediate post-test on new piece of writing but also the extent to which the level of accuracy was retained over a two-month period (delayed post-test) in new writing tasks without receiving additional corrective feedback and writing instruction (Bitchener, 2008; Bitchener & Knoch, 2008a, 2008b, 2010; Sheen, 2007a; Sheen et al., 2009). However, the accuracy improvement over the two-month period was only retained by the group which received the combination of direct and oral meta-linguistic feedback, but not the group which received direct feedback only. In other words, the students who received a combination of direct and oral meta-linguistic feedback successfully retained their accuracy over a two-month period, whereas those who received direct feedback only could not retain the gains obtained in the immediate post-test.

It was also found that the treatment type had a significant effect on the accuracy scores of the participants when the past simple tense was used. This contradicts the claim of Kepner, 1991; Truscott, 1996, Truscott, 2007, and Sheppard, 1992 that providing corrective feedback on students' writing is ineffective and even harmful, but is congruent with studies that showed greater effectiveness for certain feedback strategies on improving students' performance (e.g. Bitchener et al., 2005; Ellis et al., 2008, Sheen, 2007a, 2010). This illustrates the positive long-term effect of providing L2 writers with meta-linguistic feedback as reported by Sheen (2007a) in her study of intermediate L2 writers. As pointed out by Bitchener and Knoch (2010), providing writers with clear and simple meta-linguistic

explanation specifically rules with examples is the best type of corrective feedback for long-term accuracy.

In addition, the findings of this study support the claim that the effectiveness of written feedback can only be measured when accuracy is traced on new pieces of writing rather than on the revision of a single writing task (Truscott, 1996, 1999, 2004; Ferris, 1999; Bitchener, 2009). They also show that targeting one linguistic form (focused feedback) is a key factor in improving students' writing accuracy (sheen, et al., 2008; Han, 2002). Since focused corrective feedback is clear and systematic, writers are likely to focus their attention on form. As Ellis (2002) indicated that targeting one aspect of language is important in that it helps to "(1) monitor language use and, thereby, to improve accuracy in output, (2) facilitate noticing of new forms and new form-function mapping and (3) make possible noticing the gap……" (p. 164). However, our study departed from probably all previous investigations in the way focused feedback is parameterized: targeting all the functional uses of one specific target structure at a time.

Conclusion

The aim of the study was to investigate the effect of certain types of corrective feedback on enhancing L2 writers' linguistic accuracy and whether this accuracy was maintained over a two-month period. The study showed that the combination of direct written corrective feedback and meta-linguistic explanation resulted in improving the accuracy in the use of simple past tense and this accuracy was retained two months later without providing additional writing instruction on the students' writing assignments. This study showed the effectiveness of corrective feedback not on revised drafts of the same composition but on new pieces of writing, lending support to a carry-over effect of error corrections. Furthermore, it showed that the group which received direct corrective feedback and meta-linguistic explanation on the targeted linguistic form outperformed the control group that received no corrective feedback at all, but not the direct feedback only group.

Due to the effect of meta-linguistic explanation on improving and retaining L2 writers' accuracy over the time reported, the researchers recommend that teachers continue to provide their students with mini-lessons on particular error categories. This stresses Sheen's stance in that meta-linguistic feedback results in differential effects on foreign language acquisition, probably because it advances "a deeper level of cognitive processing" (Sheen 2007a: 260). However, further future studies are needed to investigate the benefits of different single feedbacks or combinations of these feedbacks, which

feedback option is more rewarding on the long term, how many structures should be targeted, and which functional uses of each targeted structure should be corrected at one time.

Needless to say, replicating the present study for longer periods of time and other writing genres, with students of varying learning styles in other foreign language learning contexts will surely help settle the dispute among researchers who experimented with corrective feedback, namely which error correction technique is the most effective, a finding long waited for by writing teachers, the deliverers of feedback in the classroom.

References

Al-Jarrah, R. & Al-Ahmad, S.(2012). An optimality-theoretic account of corrective feedback in process writing. *Asian EFL Journal, 14*(3), 10-40.

Al-Jarrah, R. & Al-Ahmad, S.(2013). Writing instruction in Jordan: Past, present, and future trends. *System*, 41, 84-94.

Ashwell, T. (2000). Patterns of teacher response to student writing in a multi-draft composition classroom: Is content feedback followed by formal feedback the best method? *Journal of Second Language Writing, 9*(3), 227-257.

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

Bitchener, J. (2009). Measuring the effectiveness of written corrective feedback: A response to ''Overgeneralization from a narrow focus: A response to Bitchener (2008)''. *Journal of Second Language Writing, 18*(4), 276–279.

Bitchener, J., & Knoch, U. (2008a). The value of written corrective feedback for migrant and international students. *Language Teaching Research Journal, 12*(3), 409–431.

Bitchener, J., & Knoch, U. (2008b). The value of a focussed approach to written corrective feedback. *ELT Journal, 63*(3), 204–211.

Bitchener, J., & Knoch, U. (2009). The relative effectiveness of different types of direct written corrective feedback. *System, 37(*2), 322–329.

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

Bitchener, J., Young, S., & Cameron, D. (2005). The effects of different types of corrective feedback on ESL student writing. *Journal of Second Language Writing*, 14, 191-205.

Bruton, A. (2009). Improving accuracy is not the only reason for writing, and even if it were. *System*, 37, 600-613.

Cañado, P. & Bedmar, D. (2006). Data-driven learning and awareness-raising: An effective tandem to improve grammar in written composition? *International Journal of Emerging Technologies in Learning 1*(3), 1-11.

Carroll, S., & Swain, M. (1993). Explicit and implicit negative feedback: An empirical study of the learning of linguistic generalizations. *Studies in Second Language Acquisition*, 15, 357–386.

Chandler, J. (2003). The efficacy of various kinds of error feedback for improvement in the accuracy and fluency of L2 student writing. *Journal of Second Language Writing, 12*(3), 267–296.

Chandler, J. (2004). A response to Truscott. *Journal of Second Language Writing*, 13, 345-348.

Cohen, A., & Robins. M. (1976). Toward assessing interlanguage performance: The relationship between selected errors, learners' characters, and learners' explanations. *Language Learning*, 26, 45-66.

Council of Europe (2011). *Common European framework of reference for languages: Learning, teaching, assessment*. Cambridge: Cambridge University Press.

Ellis, R. (1998). Teaching and research: Options in grammar teaching. *TESOL Quarterly, 32*,39-60.

Ellis, R. (2002). Methodological options in grammar teaching materials. In E. Hinkel & S. Fotos (Eds.). *New perspectives on grammar teaching in second language classroom* (pp. 155-179). Lawrence Erlbaum Associates.

Ellis, R. (2008). A typology of written corrective feedback. *ELT Journal, 63*(2), 97-107.

Ellis, R. (2009). Task-based language teaching: Sorting out the misunderstandings. *International Journal of Applied Linguistics 19*, 221–246.

Ellis, R., Loewen, S., & Erlam, R. (2006). Implicit and explicit corrective feedback and the acquisition of L2 grammar. *Studies in Second Language Acquisition*, 28, 339–368.

Ellis, R., Sheen, Y., Murakami, M., & Takashima, H. (2008). The effects of focused and unfocused written corrective feedback in an English as a foreign language context. *System*, 36, 353–371.

Fathman, A., & Whalley, E. (1990). Teacher response to student writing: Focus on form versus content. In B. Kroll (Ed.), *Second language writing: Research insights for the classroom* (pp. 178–190). Cambridge: Cambridge University Press.

Fazio, L. (2001). The effect of corrections and commentaries on the journal writing accuracy of minority-and majority-language students, *Journal of Second Language Writing*, 10, 235-249.

Ferris, D. R. (1999). The case for grammar correction in L2 writing classes. A response to Truscott (1996). *Journal of Second Language Writing*, 8, 1–10.

Ferris, D. R. (2002). *Treatment of error in second language student writing*. Ann Arbor: University of Michigan Press.

Ferris, D. R. (2003). *Response to student writing: Implications for second language students*. Mahwah, NJ: Lawrence Erlbaum Associates.

Ferris, D. R. (2004). The grammar correction debate in L2 writing: Where are we, where do we go from here and what do we do in the meantime? *Journal of Second Language Writing*, 13, 49-62.

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

Ferris, D. R., & Hedgcock, J. S. (2005). *Teaching ESL composition: Purpose, process, and practice*. Mahwah. NJ: Lawrence Erlbaum Associates.

Ferris, D. R., Chaney, S. J., Komura, K., Roberts, B. J., & McKee, S. (2000). *Perspectives, problems, and practices in treating written error*. In Colloquium presented at International TESOL Convention. Vancouver, B. C., March 14-18.

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

Frantzen, D. (1995). The effects of grammar supplementation on written accuracy in an intermediate Spanish content course. *Modern Language Journal*, 79, 329-344.

Guenette, D. (2007). Is feedback pedagogically correct? Research design issues in studies of feedback on writing. *Journal of Second Language Writing*, 16, 40–53.

Han, Z. (2002). A study of the impact of recasts on tense consistency in L2 output. *TESOL Quarterly*, 36, 543-572.

Hyland, F. (2003). Focusing on form: Student engagement with teacher feedback. *System*, 31, 217-230.

Kepner, C. G. (1991). An experiment in the relationship of types of written feedback to the development of second language writing skills. *Modern Language Journal*, 75, 305-313.

Lalande, J. F. (1982). Reducing composition errors: An experiment. *Modern Language Journal*, 66, 140–149.

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

Lyster , R. (2004). Differential effects of prompts and recasts in form-focused instruction. *Studies in Second Language Acquisition*, *19*, 399 – 432.

Lyster, R., and Saito, K. (2010). Oral feedback in classroom SLA: A meta-analysis. *Studies in Second Language Acquisition*, 32, 265-302.

Muranoi, H. (2000). Focus on form through interaction enhancement: Integrating formal instruction into a communicative task in EFL classrooms. *Language Learning*, 50, 617–673.

Polio, C., Fleck, N., & Leder, N. (1998). "If only I had more time": ESL learners' changes in linguistic accuracy on essay revisions. *Journal of Second Language Writing*, 7, 43-68.

Robb, T., Ross, S., & Shortreed, I. (1986). Salience of feedback on error and its effect on EFL writing quality. *TESOL Quarterly*, 20, 83–93.

Russell, J., Spada, N. (2006). The Effectiveness of corrective feedback for the acquisition of l2 grammar: A meta-analysis of the research. In J. Norris & L. Ortega (Eds.), *Synthesizing research on language learning and teaching* (pp. 133-164). Benjamins, Amsterdam.

Semke, H. (1984). The effects of the red pen. *Foreign Language Annals*, 17, 195–202.

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

Sheen, Y. (2007b). The effects of corrective feedback, language aptitude, and learner attitudes on the acquisition of English articles. In A. Mackey (Ed.) *conversational interaction in second language acquisition. A collection of empirical studies* (pp. 301-322). Oxford: Oxford University Press.

Sheen, Y. (2010). Differential effects of oral and written corrective feedback in the ESL classroom. *Studies in Second Language Acquisition*, 32, 203– 234. doi:10.1017/S0272263109990507

Sheen, Y., Wright, D., & Moldawa, A. (2009). Differential effects of focused and unfocused written correction on the accurate use of grammatical forms by adult ESL learners. *System,* 37, 556–569.

Sheppard, K. (1992). Two feedback types: do they make a difference? *Regional English Center Journal*, 23, 103-110.

Truscott, J. (1996). The case against grammar correction in L2 writing classes. *Language Learning*, 46, 327–369.

Truscott, J. (1999). The case for "the case for grammar correction in L2 writing classes: A response to Ferris. *Journal of Second Language Writing, 8*(2), 111-122.

Truscott, J. (2004). Dialogue: Evidence and conjecture on the effects of correction: A response to Chandler. *Journal of Second Language Writing, 13*(4), 337-343.

Truscott, J. (2007). The effect of error correction on learners' ability to write accurately. *Journal of Second Language Writing*, 16, 255-272.

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

Appendix (1): Pre-test writing sample

Joha and the cat.

One day, there was a man and his wife
lived in a small house. In the morning, Joha
went go to buy some meat. After that, his wife
cooked the meat, and invited her friends to eat
with her. in the evening, Joha came come to the
house and asked his wife about the meat. but he
it seemed that his wife didn't leave any meat for
him. So he asked his wife, "where is the meat?"
she lied and said the cat had ate eaten the meat,
but Joha didn't beleive that, so he followed
the cat and caught catch it d because he wanted
to know the cat weight. After that he found find
that the cat's weight
the weight didn't change. Joha beleived
that the cat didn't eat had eaten the meat.

33

The Tiddles

One day, there was a women called Deborah living in a small house with her cat called Tiddles, she loved her cat so much.

That day she missing [missed] her cat, and she offered reward who found the cat. In the evening Mr and Mrs Branksome enjoy^ed watched TV, but it was very old tv. Suddenly, something wrong happend in the picture. They went outside and took the torch to see what was the wrong, they see^saw a cat on the aerial. Then they phoned a fire brigade to discovered the wrong. after that, firefighter came and rescued the cat. next day Mr and Mrs Branksome returned the cat to her owner. Deborah gave the reward to Mr and Mrs Branksome. and they went to the Mall and bought a new television.

The burglar

When I was alittle girl, my Grand mother told me

astory about aburglar who's brok into ahouse in order to steal

it. The burglar began with the down stairs, so he collected the

in abig bag

most expensive things in that house, but after that he felt

hungry, so he went to the Kitchen to eat. The burglar opened

was

the fridge and took a piece of cheese, but he still hungry. The

burglar remembered that he saw acake in the fridge, so

he returned there and ate it. The burglar decided to

go up stairs to complete his work. During this time, The

burglar felt tired, so he went to the bed and slept.

The nighboures saw a light in the house, so they called

the police men. When the police men arrived they found

the burglar sleeping on the bed, so they cought hem and

but him in the prison. Although this story was very

funny, but I felt sorry on this folish on this burglar

who was very stupied.

Appendix (4): Controlled exercise

Complete the sentences using the past simple affirmative or negative.

1. The book was boring. I _____ it. (like)

2. I _____ party to celebrate my birthday. It was great. (have)

3. I _____ well because there was a lot of noise. (sleep)

4. She _____ a lot because she was very hungry. (eat)

5. The concert was too expensive, so they _____. (go)

6. We went to the best restaurant in London. The food _____ delicious. (be)

Appendix (5): The pre-test writing task

The pictures of this writing task told the story of a famous legendary mean character called Juha, who did not believe his wife that their cat ate the meat he brought for lunch. He weighed the cat and found out that his wife was not telling the truth.

Appendix (6): Immediate post-test writing task

The pictures of this writing task told the story of a woman who gave a couple a reward after they had called the fire brigade to rescue her missing cat which was stuck on their TV antenna.

Appendix (7): Delayed pot-test writing task

The pictures of this writing task were about a hungry burglar who was caught by the police after he had broken into a house, eaten the food in the refrigerator, and fallen asleep.

Examining Potential Sources of Miscommunication between Japan and the West: Using Grice to Bridge the Sociolinguistic Gap for Japanese EFL Learners

Pino Cutrone

Nagasaki University, Japan

Bio data

Pino Cutrone is currently an Associate Professor at Nagasaki University. He received his PhD in Applied Linguistics from the University of Reading, UK. His research interests include pragmatics, conversational analysis, intercultural communication, sociolinguistics relating to Japanese EFL learners, CALL and study abroad programs. He has published widely in the field of Applied Linguistics and Intercultural Pragmatics. E-mail: k-pino@nagasaki-u.ac.jp

Abstract

The main objective of this paper is to shed light on some issues that Japanese EFL speakers sometimes experience when communicating across cultures. These issues often stem from a lack of sociolinguistic competence in English. Revisiting the long-standing debate regarding the universality of Grice's theory of conversation, this paper argues that Grice's maxims of conversations do not apply universally and independently of culture. With the intention of informing EFL pedagogy in Japan, the writer demonstrates how Grice's theory of conversation can serve as a useful framework for intercultural analyses. In considering the thought processes and ideologies involved in interpreting each of Grice's four maxims across cultures, this paper highlights some of the fundamental issues underpinning cross-cultural misunderstandings between Japanese EFL speakers and native speakers of English (NESs). By identifying some of the specific reasons for pragmalinguistic failure, this article helps language educators, as well as cross-cultural communication trainers involved with Japanese people, deal with these issues. To this end, it is suggested that ELT professionals begin by incorporating targeted awareness-raising strategies in their contexts and then follow up by providing students/trainees with opportunities to develop

Address correspondence:
School of Global Humanities and Social Sciences, Nagasaki University, 1-14 Bunkyo-machi, Nagasaki City, Nagasaki Prefecture, Japan, 852-8521

better product-oriented conversational management techniques.

Keywords: EFL pedagogy, Grice, intercultural communication (IC), Japanese, miscommunication, pragmatic failure, sociolinguistics

Introduction

The general aim of this paper is to inform EFL pedagogy in Japan by highlighting potential sources of misunderstanding of Japanese EFL speakers in intercultural encounters. First, in specifying an area of weakness among Japanese EFL/ESL learners (JEFL/ESLs hereafter), this section begins by examining the concept commonly known as communicative competence. In SLA, communicative competence most often refers to Hymes's (1971) seminal article outlining the skills thought to define L2 ability. This concept was further developed by Canale and Swain (1980), whose definition of communicative competence has become canonical in the field of Applied Linguistics. Canale and Swain (1980) define communicative competence in terms of four components: grammatical competence (i.e., words and rules), sociolinguistic competence (i.e., appropriateness), discourse competence (i.e., cohesion and coherence) and strategic competence (i.e., appropriate use of communication strategies).

While the general failure of English language education in Japan is well known and continues to generate a lot of discussion (Lockley, Hirschel & Slobodniuk, 2012), most analysts agree that oral skills are what Japanese EFL learners have the greatest trouble with (Ellis, 1991; Farooq, 2005; Helgesen, 1987; Hughes, 1999; Okushi, 1990; Matsumoto, 1994; Yano, 2001; Reesor, 2002; Roger, 2008; Takanashi, 2004). For instance, Ellis (1991) and Okushi (1990) have noted that regular Japanese high school and/or university graduates are seriously incompetent in their English skills, particularly where sociolinguistic competence is concerned. Farooq (2005, p. 27) describes JEFLs as having "extreme difficulties in interacting with native speakers in real-life situations even at a survival level". The term "false beginner" is often used to describe JEFLs in current course books and/or teacher instructional manuals designed for university classes (Helgesen, Brown & Mandeville, 2007; Martin, 2003). According to Peaty (1987, p. 4), JEFL university students are "prototype false beginners", because they have a background in English based on their study of grammar and translation in junior and senior high school, but have very little, if any, communicative abilities.

It is not difficult to fathom how people from different cultures, who may have a high degree of grammatical proficiency in English, will, at times, still have trouble communicating in English. Knapp

and Knapp-Potthoff (1987, p. 8) shed light on the process underpinning this difficulty by describing "intercultural communication as taking place whenever participants introduce different knowledge into the interaction which is specific to their respective sociocultural group". In other words, interactants in intracultural encounters are thought to implicitly share the same ground rules of communication and meaning of signals (O'Keeffe, 2004), whereas interactants in intercultural encounters are likely to experience a degree of uncertainty and ambiguity concerning the meaning of signals and the ground rules by which communication will occur (Gudykunst & Nishida, 2001; Gudykunst, Nishida & Chua, 1986; Gudykunst, Yang & Nishida, 1985). In oral/aural exchanges, the meanings of utterances are negotiated jointly by speaker and listener; thus, it is always necessary for the receiver to draw inferences about the intentions of the sender (Scollon & Scollon, 1995). Despite the great interest in intercultural (mis)communication, a great majority of intercultural analyses seems to be anecdotal and lacking in a theoretical foundation. Thus, concerning the former, one of the aims here is to extend beyond anecdotal observations and stereotypical representations by providing empirical data as evidence of support or refutation. Further, concerning the latter, this paper has adopted the ideas first proposed in Nunn's (2003) article, in which he demonstrated the benefits of using Grice's (1975) cooperative principle as a theoretical basis for analyzing intercultural communication (hereafter IC). Specific to the writer's teaching context, this paper uses a Gricean framework for intercultural analyses in order to identify some potential sources of miscommunication experienced by JEFL learners.

Grice's Theory of Conversation

A General Summary

Much of the literature involved in developing politeness theory by scholars such as Brown and Levinson (1978, 1987), Lakoff (1973) and Leech (1983) stems from Grice's (1967, 1975, 1989) well-known theory of conversation. The assumption of Grice's theory rests on the notion that people are intrinsically cooperative in order to construct meaningful conversations. This assumption is known as the Cooperative Principle (CP). Examining the components that make up Grice's (1975) CP, and considering how members of different cultures may interpret these components differently, may shed some light on some of the misunderstandings in IC caused by different communication styles. As stated in Grice's (1975, p. 45) seminal work, Logic and Conversation, interactants tend to "make [their] conversational contribution such as is required, at the stage at which it occurs, by the accepted purpose or direction of the talk exchange in which [they] are engaged". Grice further suggests that there are a

number of conversational rules, or maxims, that regulate conversation by way of enforcing compliance with the cooperative principle. Taken from Grice's (1975, pp. 45-47) work, these maxims, and submaxims within, are divided into four categories (i.e., quantity, manner, quality, and relation) and presented as follows:

In the category of quantity, there are the following two maxims:

1. Make your contribution as informative as is required (for the current purposes of the exchange), and

2. Do not make your contribution more informative than is required.

The category of manner involves the super maxim *be perspicuous* and the following four maxims:

1. Avoid obscurity of expression,

2. Avoid ambiguity,

3. Be brief (avoid unnecessary prolixity), and

4. Be orderly.

The category of quality has one main maxim and two submaxims (a and b) as follows:

1. Make your contribution one that is true,

a. Do not say what you believe to be false, and

b. Do not say anything for which you lack evidence.

The category of relation has one maxim:

1. Make your contribution relevant and timely.

Before a discussion can ensue regarding how Gricean theory can inform intercultural analyses, it is necessary to address the ongoing debate regarding the universality of Grice's (1967, 1975, 1989) theory of conversation.

Differing Interpretations

Some researchers have questioned the feasibility that the maxims can apply universally and independently of culture, style and genre (Keenan, 1976), and others have focused their attacks on the universality of Grice's CP in the context of politeness (Churchill, 1978; Mura, 1983). Demonstrating a context in which Grice's CP is not adhered to, Keenan's (1976) study showed that the people he observed in Madagascar tended not to give information when required, which intentionally and systematically violates Grice's quantity maxim. According to Keenan (1976), Malagasy speakers tend to be reluctant to share information because of the risk of losing face by committing oneself to the truth of the information, as well as the fact that having information is a form of prestige in their culture.

Other researchers, however, have staunchly defended Grice's CP, on the grounds that many linguists continue to misunderstand what Grice was trying to do (Horn, 2004; Levinson, 1983, 2000; Nunn, 2003). As Nunn (2003) rightfully pointed out, Grice (1989, p. 26) himself makes no explicit claims of universality, using typically modest language to refer to a "first approximation of a general principle" and a "rough general principle" in describing his theory. Grice (1989, p. 26) is equally cautious in choosing his words so as not to overstate the case for cooperation in his theory as he suggests "each participant recognizes in them [talk exchanges], to some extent, a common purpose or set of purposes, or at least a mutually accepted direction". By advancing only for the existence of some general principle as this, it is apparent that Grice is adopting what he believes to be the appropriate degree of certainty for a conversational principle (Nunn, 2003).

In short, Grice's maxims can be seen to encompass the basic set of assumptions underlying verbal exchanges; however, this is not to imply that these maxims are regularly followed in every verbal exchange as critics have sometimes thought. Grice (1975) did not prescribe these maxims as laws governing conversation; rather, Grice (1975) fully expected people to flout, violate, infringe, and opt out of the maxims. In fact, the instances when the maxims are not followed were of particular interest to Grice (1975), as they are useful for analyzing and interpreting conversation, and often generate inferences beyond the semantic content of the sentences uttered, which Grice (1975) called conversational implicatures. Grice's (1975) maxims provide the foundation to Brown and Levinson's (1983) theory of politeness because, similar to Lakoff (1973) and Leech (1983), "the theorists understand that deliberately broken maxims can implicate more information than what is actually being said" (Lindblom, 2001, p. 1614). The widespread and longstanding application of Grice's theory of conversation in the research and in EFL course and resource books is evidence of its value (Nolasco & Arthur, 1987). For the purposes of this paper, using a Gricean framework to assess potential ideological differences of individuals across cultures seems to be an effective way to uncover potential sources of intercultural miscommunication and negative perceptions. In other words, while these maxims may very well be universal on some levels, their interpretation may be influenced by culture as well as other variables such as personality, context, age, gender, etc.

Understanding the Finer Points of Grice's Theory

To demonstrate how a Gricean framework can inform this intercultural analysis, it is necessary to take into account various concepts underpinning the main issues in this analysis. As the writer discussed

above, not all discourse encompasses Grice's (1975) CP as deceit, long-windedness, irrelevance, obscurity, taciturnity are all, for good or ill, part of natural communication. Some of the terms used to describe instances when Grice's CP is not followed include the actions known as violating a maxim, flouting a maxim, and creating a conversational implicature. Regarding the first, violating a maxim refers to when a speaker intentionally does not follow a maxim. This can occur in the form of a major violation or as a minor violation. A major violation would be evident when the speaker openly opts out from the operation of the maxim and the CP, such as when the speaker deliberately and secretly subverts the maxim and the CP, for some usually selfish end such as trying to deceive the listener (i.e., covertly violating the first maxim of quality), or when the speaker intentionally dominates the conversation, persistently violates the first maxim of quantity, and repudiation of the CP along with it. A minor violation of a maxim, on the other hand, would entail the speaker attempting to maintain the CP by coming out and telling the listener they are violating a maxim and why, as the following examples *I don't know if this is relevant, but…* and *this is just what I heard in passing, so I can't really vouch for the quality* demonstrate (Gartsman & Hughes, 2007). In these examples, the speaker is seen to have minor violations of the relation maxim and the quality maxim respectively.

In contrast to violating a maxim, the action of flouting a maxim refers to instances when the speaker is clearly and deliberately not following a maxim in order to imply something beyond what they have uttered. A common example of this would be what Bouton (1994) refers to as *Pope Questions* (Pope Q) to convey (rather sarcastically) the reply *of course*. That is, if one person asked another if he/she liked sushi and the other person responded with "*Is the pope catholic?*" (or another variant such as "*Is the sky blue?*" or "*Do fish swim?*"), the conversational implicature would be "*Yes, of course, I love sushi*". In this example, the Pope Q blatantly flouts the maxim of relevance, yet (assuming he/she were proficient in English) the person posing the original question "*Do you like sushi?*" would understand that the obvious answer to the question "*Is the pope catholic?*" (i.e., of course) becomes the answer to their original question "*Do you like sushi?*" In this way, the listener recognizes the speaker's intention and tries to draw the implied meaning, or conversational implicature, out of the utterance. For such flouting to be interpreted as such and the subsequent conversational implicature to be drawn, the speaker and the listener, apart from being cooperative, must share similar cultural and linguistic norms.

Intercultural encounters often involve interactants who do not share tacit knowledge that would enable them to achieve a higher level of understanding and communication, so it may be especially

difficult for listeners to interpret the implied meanings that the speaker intended. When these norms are not shared, the misunderstanding of utterances may arise if taken at face value. Utterances in conversations, and the inferences that are sometimes made through conversational implicature, can easily be misconstrued by a conversational participant as their interlocutor not being cooperative. Besides the difficulties associated with recognizing such inferences in intercultural exchanges, misunderstandings across cultures can occur in other more obvious ways. For instance, as Murray (2011) points out, culture or context may cause the instinctive suspension of Gricean maxims, such as not talking about something taboo in a particular culture, or not being brief in the context of preparing a legal document. Intercultural misunderstandings are examples of pragmatic failure, which Blum-Kulka and Olshtain (1986) define more specifically as a communication problem that occurs whenever two conversational participants fail to understand each other's intentions. While such miscommunication can even occur between interactants who share similar linguistic and cultural backgrounds, Blum-Kulka and Olshtain (1986) assert that it is much more likely to occur in conversations involving participants from different origins and languages. In the same way, models of intercultural communicative competence commonly assume people from different cultural backgrounds may have differing expectations about communication that serve as a framework for interpreting, responding to, and evaluating verbal and nonverbal communication (Spitzberg, 2000).

Using a Gricean framework to identify sources of miscommunication across cultures

Part 1: Potentially Different Interpretations of Grice's Maxim of Quantity

As introduced above, the category of quantity consists of the following two maxims: make your contribution as informative as is required (for the current purposes of the exchange), and do not make your contribution more informative than is required. In these maxims of quantity, the phrase "as is required" seems open to interpretation as too much and/or too little would seem to be relative concepts. In other words, what is enough for an introverted person may not be enough for an extroverted person. In the same way, an English speaking person may have different ideas from a Japanese person regarding what enough means. Nonetheless, it is not the writer's intention to propagate the shy Japanese stereotype here. Rather, recognizing that individual interpretations will vary, the writer seeks to examine possible culture-related ideologies that might also influence this phenomenon.

There is a wealth of literature describing JEFL speaker/learners' disposition towards taciturn behavior (Anderson, 1993; Ellis, 1991; Greer, 2000; Nozaki, 1993; Townsend & Danling, 1998);

however, much of it is anecdotal, cast in an essentialist light and has been challenged by some scholars. For instance, in his book <u>Japan's Modern Myth</u>, Miller (1982) refutes the idea that silence plays a distinctive role in the social life of Japanese. Miller (1982) contends that Japanese people are no more silent than other cultural groups, and this myth only serves to perpetuate the Japanese image of themselves as a mysterious, unique, undecipherable and hence profound culture. According to Anderson (1993), some Japanese people do talk, and sometimes they talk a lot, but the contexts in which they speak are culturally sanctioned and do not correspond to the cultural codes of the West. Thus, when thought of in this way, the notion of Japanese silence may simply be a form of Western ethnocentrism in some respects and may have more to do with speakers not being familiar with other people's differing social and cultural codes for speaking. Barnlund (1989, p. 143) describes how in the Western world, speech is often thought to be associated with the cultured, while "silence seems to be considered neutral at best, and at worst, as a symptom of social inadequateness or even emotional illness". Indeed, the French still use the word *sauvage* to mean not only *savage* but also *unsociable* or someone who does not have the skills and/or willingness to partake in the art of discourse (Yamada, 1997). Similarly, the following well-known quote by the German novelist Thomas Mann (cited in Yamada 1997, p. 17) exemplifies the degree to which some Westerners are thought to value speech: "Speech is civilization itself. The word, even the most contradictory, preserves contact – it is silence which isolates". This is consistent with the way in which the silence of Native Americans has been interpreted in the literature. While early depictions by scholars in the literature, as well as those held by the general population, attributed the Native Americans' failure to communicate as savagery, later studies conducted by anthropologists and linguists were more sympathetic as they portrayed the Native Americans as victims of the modern world (McDermott, 1987). Still, even in these later studies, their silence symbolized death, handicap, or the absence of civilization, while any ability to speak up signaled progress (Yamada, 1997). In the United States, a great deal of the research conducted in this area seems to treat silence as a symptom of pathology relating to shyness, which leads to communicative failure and a deteriorating relationship, rather than a sign of growth (Barnlund, 1989; Hendersen, Zimbardo & Carducci, 1999). The position adopted in this paper is that perceptions of silence (and shyness) will likely differ among individuals, as well as across cultures. Moreover, the views stated above, which describe the Western negative view of silence, seem too general, simplistic and convenient to be satisfied with. As Bruneau (1973) warned long ago, the role of silence in complex cultures is profound and needs to be studied in much greater depth to be truly understood.

Moreover, while agreeing with Andersen (1993, p. 102) above that the "the Japanese are silent" stereotype is far from the truth, there does appear to be some legitimacy to the idea that Japanese culture may value taciturnity over verbosity in some ways as demonstrated by the long list of famous Japanese proverbs to that effect, some of which include the following: *Chinmoku wa kin nari* (Silence is golden), *Kuchi wa wazawai no moto* (The mouth is the source of the calamity), *Kuchi ni mitsu ari, hara ni ken ari* (Honey in the mouth, dagger in the heart). In modern times, some of these proverbs have been rewritten with irreverent twists and used by TV personalities, such as *Oshaberi wa kuchi no onara* (talkative is a mouth's fart), and *Tori no nakaneba utaremaji* (if the bird had not sung, it would not have been shot). The first proverb above *Silence is golden* implies a general positive impression of silence, which is also evidenced in the Japanese ideographic Kanji symbol for the word *ma*, (間, pause or space). This symbol is drawn to represent the sun shining through the gates, illustrating how implied communication can shine through silence. The idea that implicit communication is desirable in Japanese is central to the cultural concept known as Haragei (literally belly art). For now, the purpose is to communicate that *ma*, or silence in Japanese conversation, is more than just a pause or empty space; rather, it is an important element that helps construct communication. The other four proverbs above encompass the widespread view that the Japanese may be somewhat skeptical of talk (Kenna & Lacy, 1994; Townsend & Danling, 1998). Relating this to the discussion of the Japanese cultural concept *wa* (i.e., striving for group harmony), these proverbs communicate the idea that words have the power to hurt people and, thus, potentially disrupt group harmony. Consequently, some scholars have explained Japanese silence as an effect of *wa* in which Japanese people avoid talking to limit the chances of hurting someone's feelings because it is safer to adopt a listener's role (Elwood, 2001; Matsumoto & Boyè Lafayette, 2000; Yamada, 1997). With these points in mind, it is important to note that despite the many contrastive studies that quote proverbs as support (Lebra, 1987; Nonaka, 1996; Scollon & Scollon, 2001), any conclusions derived from the meaning of a proverb should be treated with great caution as there often exist numerous proverbs displaying entirely opposing values (as shown by Rose, 1996 and Susser, 1998).

A Look at Some of the Studies in this Area

While a good portion of the descriptions portraying the Japanese as valuing taciturnity over verbosity seem to have been anecdotal, there are numerous empirical studies which seem to support the notion that some Japanese may be more comfortable with silence than citizens of some Western nations such

as the United States. For instance, in a contrastive study comparing the communication styles of Japanese and American businessmen, Yamada (1997) reported an average rate of silence of 5.15 seconds per minute in the Japanese meeting and only .74 seconds in the American meeting, and the longest pause in the Japanese meeting was 8.5 seconds and only 4.6 seconds in the American meeting. There seems to be a strong belief that the long pauses and brief utterances commonly found in Japanese may negatively transfer to the L2, as several studies involving the intercultural analyses of communication styles have shown; the JEFL 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). This may be contrary to what some cross-cultural interlocutors might hope to encounter 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, 2003; McCroskey, 1992; Ross 1994; Sato, 2008; Stubbe, 1998; Yashima, 2002). When fundamental behaviors are not shared and/or do not conform to one's expectations, there is a danger that those behaviors may be negatively perceived, lead to stereotyping, and in the worst case scenario, be misinterpreted as transgressions against one's value system (Armour, 2001, 2004; Chapman & Hartley, 2000).

How Do Differences in Quantity of Speech Affect Individuals' Perceptions across Cultures?

Some insights towards answering this question can be found in studies administered by Cutrone (2005) and Sato (2008). First, Cutrone (2005) examined listener responses and their effect on IC in eight dyadic casual conversations in English between Japanese and British participants. In follow-up interviews with the participants, Cutrone (2005) found evidence to suggest that listening behavior which was not shared between cultures may have contributed to negative perceptions across cultures. In relation to potential differences in communication styles concerning the maxim of quantity, the data revealed that the British participants spoke more than double the amount of the Japanese participants, with each group uttering 1985 and 887 words respectively. Although many of the British participants anticipated and accepted that they would be responsible for carrying the conversation (as they were more proficient in English), the following excerpts from the qualitative data seem to imply that it may have detracted from their enjoyment of the conversation:

> Victoria: Her reactions made me feel like she didn't want to speak, be put on the spot as she seemed content to just let me take it, (the primary speakership in the conversation) but I get tired after a while ya know.

> William: If I didn't ask him direct questions he probably would just continue nodding. I felt as though he didn't

really want to speak. Maybe he was nervous.

Elizabeth: Of course I'd love for her to have spoken more but I don't think it's in her nature to do so. She seems much more comfortable in a listener's role.

Charles: Well in a real life situation, like if I was in a bar or something, I doubt that I'd try so hard to keep the conversation going.   (Cutrone, 2005, p. 267)

Many of these sentiments were echoed in Sato's (2008) study, which investigated the oral communication problems and strategies of a group of 32 intermediate JEFL learners. In this study, Sato (2008) video-recorded face-to-face oral proficiency interviews between a learner and a native-English speaker (NES) interlocutor-assessor. From the interview data, Sato (2008) found that the JEFL learners tended to provide minimal responses and were not prone to elaboration. Although there was some cross-learner difference, the Japanese learners generally tended to provide short answers with solely factual information and did not show much, if any, awareness of the need for elaboration. From the subsequent verbal report sessions with the NES assessor, as well as the feedback provided from two additional NES co-raters, Sato (2008) reported that these under-elaborated or minimal responses gave the NESs the impression that these learners were uncooperative participants. As suggested in the research, such minimal responses could undermine interpersonal relationships as they may be perceived as an unwillingness to communicate and/or may even tire or bore the interlocutor (Andersen, 1994; Gumperz, 1995; Sato, 2008). Accordingly, the NES assessor in this study commented that brief answers, even those that may have been linguistically correct, would be graded down under test conditions as such responses would exhibit a lack of sociopragmatic awareness.

Possible Explanations of Ideological Differences across Cultures

In an attempt to explain some of the misunderstandings involving JEFL speakers, several IC researchers (such as Andersen, 1994, Ishii & Bruneau, 1994; Koreo, 1988) have taken somewhat of an essentialist approach drawing on the distinction between collectivist and individualistic cultures proposed by Hofstede (1991), high context and low context communication posited by Hall (1981), and differing religious traditions by Yamada (1997). The first to be discussed is the influential work of the anthropologist E. T. Hall (1981), who is considered one of the founders of IC study. In his book Beyond Culture, Hall (1981) presents a broad-based theory that describes a continuum ranging from what he called a High Context Culture to a Low Context Culture, terms he used to describe cultural differences between societies. In an archetypal sense, a high context culture refers to societies or groups

in which members have close connections over a long period of time, and many aspects of cultural behavior are not made explicit because most members base their behavior on years of interaction with each other. From a communication standpoint, talking is seen to be less valuable as members tend to rely heavily on the context for the interpretation of their messages, with the meaning being partly implied instead of put into words. On the other end of the spectrum, a typical low context culture refers to societies where people tend to have many connections but of shorter duration or for some specific reason. In these societies, cultural behavior and beliefs may need to be spelled out explicitly so that those coming into the cultural environment know how to behave appropriately. Communication is generally thought to be goal-oriented, and members tend to assign great value to talking and communicate mostly through verbal language rather than tacit understanding. According to Copeland and Griggs (1985) and Diez Prados (1998), Japanese society is among the higher context cultures, whereas American society, in contrast, is among the lower context cultures.

Another theory put forward to explain the Japanese use of silence involves their seeming collectivistic orientation (Gudykunst & Ting-Toomey, 1988; Triandis et al., 1988). Referring to the work of Hofstede (1991), some researchers argue that cultural differences in beliefs about talk are due to the individualism-collectivism dichotomy, as shown in the following excerpt:

> Individualists have a choice among many groups…to which they do belong, and usually belong to these groups because they volunteer. Collectivists…are born into a few groups and are more or less stuck with them. So, the collectivists do not have to go out of their way and exert themselves to be accepted. Hence, the individualists often speak more, try to control the situation verbally, and do not value silence. (Triandis, 1988, p. 61)

Triandis's (1988) assessment seems to be consistent with the concept of *wa* in Japanese culture whereby members of a group do not usually wish to stand out from their group, and, thus, will not frequently perform actions that will cause this to happen such as initiating talk and/or volunteering answers in class. In subsequent work, Triandis (1994) describes how norms are very powerful regulators of behavior in collectivist cultures as the threat of ostracism is an especially powerful source of fear. As an example, he describes the oft-heard plight of Japanese returnees. After spending some time abroad, they are often criticized when they return home for non-Japanese behaviors such as being too outspoken and uncooperative (Triandis, 1994). This is evidenced by the fact that the term used for individualism in Japanese, *kojinshugi* also has negative connotations which imply selfishness (Ito, 1989).

Furthermore, as Yamada (1997) contends, Judeo-Christian principles may have played a role in spreading the modern day importance given to speech vis-à-vis silence in the United States. The

following New Testament verse uttered by John (1:1) seems to point to this origin: "In the beginning was the Word, and the Word was with God, and the word was God". In this verse, John is equating words with God, and by doing that, he, thus, seems to be elevating speech to the highest position of power and eminence (Yamada, 1997). By the eighteenth century, religion competed with science for prestige, and this affected the way people thought of speech. Lexicographers began to see the deification of words as unacademic, and words, although believed to be backed by God, were then thought to be created by humankind in a science of language. As such, science implied exactness; thus, a greater emphasis was not only put on speech but on clear and precise speech. Accordingly, the desire to speak, and particularly the ability to express one's feelings clearly and explicitly, are thought to be a virtue in American society (McCarthy, 2003, McCroskey, 1992).

In Japan, however, a completely different picture developed from the two influential religions, Buddhism and Taosim as well as from the principles of Confucianism, together contributing the belief that silence is sacred. A central theme to understanding this is to examine how it relates to the concept of emptiness. In Buddhism, one of the main objectives is to realize the emptiness of words. This is evidenced by the fact that Buddhist practitioners, who attend the *okyoo* (i.e., a Japanese pronunciation and rendition of the sutras originally chanted in Sanskrit), do not seek to understand or interpret the meaning in *okyoo* because the point is to realize the emptiness of the chants. Conversely, participants of Judeo-Christian services may not always understand the sermon in a service they are attending, but it is generally desired to do so. Similarly, in Taoism, forgetting language and remaining speechless is advocated as the ideal way (tao) to emptiness. The Confucian code of conduct aims to guide group members through compatible relationships and smooth interaction. The function of talk, in turn, is to act as a social lubricant, and straight forward speech is generally discouraged as saying whatever you felt was viewed as socially inappropriate (i.e., tactless and blunt). Consequently, from this convergence of religious ideas developed a belief that explicit talk with definite meanings was often undesirable, and since talk always presents the opportunity of being overly explicit, Japanese began to approach talk as a communicative medium that warrants suspicion and caution (Yamada, 1997).

Another reason put forward to explain the silence sometimes found in conversations including JEFL speakers involves the concept of face. That is, the use of silence by JEFL/ESL speakers in conversations is sometimes the result of rule-conflict in English and Japanese conversations. For instance, when confronted with questions which they cannot answer (for any number of reasons), they may resort to silence as a face-saving measure. They do this because saying *I don't know X* does not

connote the same things to a Japanese person that it might to a NES. Noguchi (1987) provides us with a useful example when he describes a common scenario of a Japanese speaker (called Mr. Suzuki) who is proficient in English conversing with a visiting American business man (called Mr. Jones) who speaks no Japanese. During a conversation that includes other members, Mr. Jones asks Mr. Suzuki what his occupation is and Mr. Suzuki discovers that he cannot answer the question because he does not remember the words *quality control*. Instead of answering *I don't know X* or *I don't know how to say X,* a long silence ensues with everyone growing more uncomfortable with every passing moment until Mr. Jones changes the topic. Noguchi (1987) describes the possible interplay in the thought process of Mr. Suzuki:

> From his Japanese language experience, he knows the conversational rule that if an appropriate question is addressed to him, he must provide an appropriate answer in the next speaking turn. Yet, try as he might, he can not (sic) recollect the needed English words. At the same time, he realizes that he cannot admit this lapse of memory in front of the group, for he fears his Japanese friends and Mr. Jones may belittle his intelligence or, perhaps, even begin to think that he really does no work at all on his job. Thus, the face-protecting rule takes effect on Mr. Suzuki. (Noguchi, 1987, p. 22)

This example seems to indicate that, in Japanese, *I don't know X* often presupposes that the speaker lacks knowledge of X because of a lack of intelligence or lack of interest in X. In contrast, the expression *I don't know X* in English seems to carry the presupposition that the speaker lacks knowledge of X but does not presuppose that the speaker lacks the intelligence to acquire knowledge of X. Further, the expression is neutral with regards to the speaker's interest in X as a topic. The dynamics of rule-conflict in English and Japanese conversations leading to uncomfortable silences as shown by the example above are common in EFL teachers' everyday interactions with their Japanese learners (Blanche, 1987; Cutrone, 2005).

While the explanations above offer some interesting insights into how and why different cultures may perceive talk the way they do, it is somewhat difficult to generalize this to the Japanese communication style as it does not account for the tremendous situational variability in Japanese society as discussed in several sections above. In addition to one's personality and willingness to communicate, the contextual variables of the conversation, which involve the status and familiarity of one's interlocutor(s), where they are speaking, and what purpose the communication serves, are among the many variables that will also influence the amount an individual speaks. For example, regarding individual differences, one person may feel more comfortable talking a lot at home but not at work, whereas another person may be quiet at home and talkative at work. Relating to the group dynamics and the status of the interlocutor(s), Gudykunst and Nishida (1994) describe how when there is an older

person with higher status present in a situation, it is often up to them to initiate speech, and should they choose not to speak, then silence would be the appropriate behavior for the others present.

Part 2: Potentially Different Interpretations of Grice's Maxim of Manner

The fundamental nature of the maxim of manner is clearly encapsulated in Grice's (1975, p. 46) super maxim "be perspicuous", which he then divides into the following four submaxims: avoid obscurity of expression, avoid ambiguity, be brief (avoid unnecessary prolixity), and be orderly. As the previous section alludes to, the degree to which any of the aforementioned submaxims are followed and/or preferred is open to interpretation, and individual preferences regarding the degree of ambiguity vis-à-vis clarity in communication will vary. In this section, the author discusses how potential culture-specific ideologies might also influence this phenomenon. There exists a great deal of literature claiming that Japanese speakers are ambiguous communicators, who tend to avoid direct, plain statements in favor of more suggestive, indirect comments in their L1 and in English (Hill, 1990; Kenna & Lacy, 1994; Loveday, 1982; Matsumoto & Boyè Lafayette, 2000). Further, as Haugh (2003) points out, this common view is held by both Japanese and non-Japanese linguists in works ranging from IC handbooks (Kitao & Kitao, 1989; McClure, 2000; Yamada, 1997), to academic papers (Akasu & Asao, 1993; Clancy, 1986; Doi, 1996; Gudykunst & Nishida, 1993; Nakai, 1999; Nittono, 1999), to dissertations (Books, 1995; Day, 1996; Iwata, 1999; Sato, 2008). Some of the oft-cited examples of Japanese indirectness and vagueness, according to Haugh (2003), involve phenomena such as the common omission of elements of Japanese utterances that would be made explicit in English (Akasu & Asao, 1993; Donahue, 1998), the common use of indexicals in Japanese such as *are* and *sore* (in English, that) in place of the topic of a subject (Akasu & Asao, 1993), the frequent use of hedges as *toumo kedo* … (I think that …) to convey hesitancy and uncertainty (Okabe, 1993; Sasasagawa, 1996), the tendency to use understatements rather than overstatements such as *tabun* (maybe/probably) rather than *zettai* (definitely), and the use of opaque formulaic utterances such as *chotto yoji ga arimasu* in reply to an invitation to go out together (in English, this may be translated as "Sorry I can't, I have some business to take care of").

A Look at Some of the Studies in this Area

Various researchers have attempted to attribute vagueness and indirectness to traditional Japanese values such as the importance of preserving harmony (*wa*) and group orientation (*shuudan shugi*)

(Nakane, 1970; Morita & Ishihara, 1989), or to the highly contextualized nature of Japanese communication (Arima, 1989; Ikegami, 1989) and seeming preference for non-verbal communication (Haga, 1998). Although the perception of Japanese as indirect and vague speakers seems to be widespread, there are a number of problems with this view. First, much of this description comes from anecdotal accounts and has not received convincing empirical support. Regarding some of the studies that have been conducted in this area, the results appear to have been influenced by the specific speech act under investigation and how the data is collected (i.e., naturally occurring data vis-à-vis discourse completion tests). While various studies have found Japanese to be more indirect and vague performing speech acts such as requesting and complimenting (Barnlund & Araki, 1985; Takahasahi, 1987), an equal number of studies have produced contrary results in reporting that Japanese can be more direct in some requesting, complaining, and conflict situations than NESs (Rose, 1992; Sato & Okamoto, 1999; Spees, 1994). In light of the evidence to date, the assumption that Japanese and JEFL speakers are more vague and indirect than NESs appears to be questionable. Thus, the perceptions of Japanese communication shared by many people, which include Japanese, non-Japanese, and academics in both groups, seem to be based on factors other than objective, empirical evidence (Haugh, 2003). Whether these views are driven by a lack of understanding of Japanese L1 and/or misunderstanding of Japanese people is not certain; however, it is evident that these perceptions do exist and thus may influence intercultural encounters between NESs and Japanese. Hence, the next step in this analysis is to describe some of the studies on Japanese and non-Japanese perceptions of Japanese communication and consider how these perceptions might influence IC.

Relating degrees of ambiguity to culture-specific views on politeness, Nisugi (1974) surveyed 250 Japanese native speakers regarding the terms they would choose to describe the Japanese language. 76 percent of the participants responded that *teineina* (polite) was an appropriate adjective for Japanese, and 73 percent responded that *amaina* (vague and indirect) was also an appropriate term. Nisugi (1974) also surveyed 20 NESs and found that 90 percent of them believed Japanese language to be vague and indirect, and 79 percent of them also considered Japanese to be a polite language. Although Nisugi's (1974) sample of non-Japanese was small, her findings were supported by studies conducted by Haugh (1998), Iwata (1999) and Sasagawa (1996), who also found that the majority of their NES respondents perceived Japanese communication to be vague and indirect. Iwata's (1999) study, which more specifically involved the perceptions of business communication across cultures between Japanese and North Americans, revealed that the North American respondents in her study generally perceived the

Japanese participants to be more indirect and nonverbal in their communicative patterns. Additionally, the findings demonstrated that the Japanese participants also consider themselves to be more indirect and vague than the North Americans; however, this view was not held as strongly as it was by the North American participants. Sasagawa (1996) surveyed 89 foreign students in Japan and found that 64 percent of them believed that Japanese often do not clearly express what they want to say. A smaller group of 55 foreign students were also asked if they thought there were many vague and indirect expressions in Japanese, and 76 percent of the respondents agreed that there were.

How Do Cross-Cultural Perceptions Concerning Manner of Speech Affect JEFLs in IC?

From the studies presented above, it appears that, although the empirical data does not necessarily support it, both Japanese and non-Japanese respondents believed Japanese communication to be generally ambiguous, vague and indirect. This then begs the question as to what extent these perceptions might affect JEFL speakers' intercultural encounters. While behaviors different from what one would expect in their own culture might be viewed negatively across cultures, it is also possible for negative perceptions to stem from preconceived notions of the other culture. Recognizing the inextricable link between expectations and perceptions in conversations, Guest (2002, p. 159) describes the perils of perpetuating stereotypes involving "the direct-talking American, who appears boorish and unsophisticated to his or her Asian hosts, while that same American is perplexed by the vague, indirect forms of speech used by the Asian interlocutor, and thus ascribes a certain 'sneakiness' or 'inscrutability' to his or her counterparts". While there exists a great deal of unsubstantiated literature making similar claims, the author was not able to locate any strong empirical evidence supporting the notion that perceptions of Japanese ambiguity negatively affects IC.

Nishiyama (1995) administered a public opinion survey in the US on American perceptions of Japanese people. The results of Nishiyama's survey (1995), which, at first glance, appear to indirectly support the above mentioned stereotype, in fact, only produce more questions and misunderstandings. One of the questions asked the respondent to choose the animal that seemed to best characterize Japanese people. The largest percentage of American respondents selected a fox. When this was reported in the Japanese media, a great many Japanese people were shocked and upset as a fox in Japan is associated with the image in Japanese folklore of an eerie, distrustful, and phantom animal. This image seems to be in sharp contrast with the American idea of a fox as a small, quick, and clever animal that possesses the astuteness to outrun and outmaneuver the hounds and hunters chasing it.

In another study, Graham (1990) examined Japanese-American business encounters and whether behavioral differences affected perceptions across cultures. The 12 Japanese participants surveyed in the study admitted that they found the Americans' openness and directness uncomfortable to some degree. Similarly, in a survey of 1346 Japanese people, Chung (1999) found that 79 percent of respondents believed that saying things plainly or directly (*hakkiri iu koto*) is impolite (*bushitsuke*). Interestingly, contrary to the explanation of ambiguity and indirectness occurring as a by-product of the Japanese desire to avoid conflict and promote harmony (*wa*), Chung (1999) found that 60 percent of the respondents felt that the use of expressions to avoid making clear judgments reflects a recent movement in Japanese society towards a passive approach to life, where one desires a peaceful and uneventful life (*kotonakare shugi*). In a somewhat different study, Yoshida et al. (2003) examined the perceptions that 486 Japanese students had of their classmates who had returned after having spent a prolonged amount of time living in a native English speaking country (i.e., returnees). Consistent with the results of similar studies such as Minoura (1988), one of the findings of Yoshida et al. (2003) was that the Japanese respondents perceived their returnee classmates as being too direct and individualistic. Interestingly, these were among the many instances in which the non-returnees held similar perceptions to the ones the returnees held of themselves.

Possible Explanations from a Japanese Perspective

While it is not possible to arrive at any clear-cut conclusions, some Japanese cultural concepts can help shed light on this area. For instance, the idea that implicit communication is desirable in Japanese society is central to the cultural concept known as *haragei*. *Haragei* is comprised of the kanji symbols *gei* (芸, acting) and *hara* (腹, guts), which combine to mean in literal terms - acting on guts alone. This represents the idea of communication without the use of direct words. According to McCreary (1986, p. 45), "the many formalities, conventions, and common standards developed in a society that gives priority to harmonious relations makes it easy to understand what is in the mind of the Japanese people". In other words, tacit understanding or *haragei* is made possible by the vertical relationships, the need for harmony, and the homogeneity found in Japanese society. Within a Vygotskyan (1962) perspective, which places emphasis on culture and society shaping cognitive development and, thus, language use, *haragei* would seem to be a form of other-regulating behavior as knowledge of the other determines the strategy to be employed. Two commonly used referents of *haragei* which help shed light on what this term means are *ishin denshin* (intuitive sense) and *sasshi* (surmise or guess). The

phrase *ishin-denshin* is translated literally as *what the mind thinks, the heart transmits*, and refers to the oft-essentialist descriptions of Japanese appearing to be using mental telepathy when they converse. Similarly, *sasshi* refers to the highly valued skill in Japanese society of being able to implicitly deduce the meaning of subtle messages.

Part 3: Potentially Different Interpretations of Grice's Maxim of Quality

The maxim of quality contains the supermaxim: make your contribution one that is true, which encompasses the following two submaxims: do not say what you believe to be false, and do not say anything for which you lack evidence. While violations of the quality maxim can result in using contradictions, irony, metaphors, and rhetorical questions, they may also be construed as exaggerations, deception and dishonesty. Similar to the discussions involving the maxims of quantity and manner above, the degree to which these maxims are followed are open to interpretation and are likely to differ according to each individual's personality as well as the specific contextual factors involved in each conversation. This section will consider whether culture-specific ideologies also influence this phenomenon.

A Look at the Research in this area, and Some Initial Explanation of Communication Styles

While there have not been many empirical investigations conducted in this area concerning Japanese behavior, one study that did investigate it was by Imai (1981). Imai assessed how Japanese businessmen respond to requests that they cannot or will not fulfill. As Nishiyama (1995) points out, there are many ways of saying *no* in Japanese without actually using the word and conveying the negative connotation that seems to go with it in Japanese society. Imai (1981) reported that a common strategy among the Japanese businessmen he surveyed was to use a number of the alternatives to the explicit word *no*, including answers which sound fairly similar to those deemed deceptive by Information Manipulation Theory (IMT), which, in brief, views deception as arising from covert violations of one or more of Grice's (1989) four maxims. One example from Imai's (1981) study occurred when some of the Japanese participants said *yes* and followed with long explanations, which may be equated to violating the maxims of quality and quantity. Other responses included using vague or ambiguous replies (non-observance of the manner maxim), avoiding the question, and changing the subject (non-observance of the relevance manner). In this way, IMT suggests that deceptive messages function deceptively because they violate the principles that govern conversational exchanges

(McCornack, 1992).

In another study, Nishiyama (1994) discusses deception in a cultural framework from an organizational perspective. Nishiyama (1994) examined the tactics and behaviors of Japanese negotiators and found a number of strategies and behaviors that her Japanese participants considered everyday business practice in Japan, yet may be interpreted as deceptive by American business people. Such commonly misunderstood messages may stem from the distinction between Japanese cultural concepts known as *tatemae* and *honne*. *Tatemae* (i.e., the public self) refers to the principle by which one is bound to the group vis-à-vis one's ranking in the vertical order of society, whereas *honne* (i.e., the private self), on the other hand, refers to one's true or inner wishes and desires. One example of tatemae in action would be if a Japanese person outwardly expressed agreement and support to a statement made by an older person (or a person of seeming higher status such as a superior at work) that they, in fact, did not agree with at all.

A study conducted by Robinson (1992), may help shed some further light in this area. In this study on 12 JESL learners' refusals in English, one of the methodological problems that arose was that several of the JESL respondents had a particularly difficult time issuing refusals and tended to accept requests rather than refuse them. Robinson (1992) attributed this to the nature of Japanese society, which he contends raises children, and especially girls, to say *yes*, or at least not say *no*. In an earlier study of the development of communication styles in children, Johnson and Johnson (1975) report that American children are socialized to speak the truth, to be honest. Miller (1994, p. 37) supports this claim by characterizing Americans as "forthright, direct, and clear". In a study comparing the American and Japanese communication styles, Okabe (1983, p. 36) concluded that "Americans' tendency to use explicit words is the most noteworthy characteristic of their communication style".

Further Explanations of Ideological Differences, and Potential Effect on IC
Although many claims given above have been cast in an essentialist light and only seem to include scant empirical support, they continue to influence perceptions. For instance, Kenna and Lacy (1994) summarize the American concept of truth as an absolute entity that is not dependent on circumstances. In other words, a fact is either true or false, and what is true for one person is likely true for everyone. In Japanese society, conversely, Kenna and Lacy (1994) contend that truth is relative and largely dependent on the situation and the parties involved. Further, the idea of communicating truth would seem to be given a much higher priority in American society than it might in Japanese society. As has

been documented by several sources (Hill, 1990; Loveday, 1982; Matsumoto & Boyè Lafayette, 2000), maintaining harmony and protecting face are thought to be much more important virtues than truthfulness, clarity and directness in Japanese culture. This last claim was supported by Cutrone's (2005, pp. 265-266) study, in which some of the JEFL interviewees admitted that they sometimes avoid giving their opinions and/or conveying the truth in an effort to preserve harmony and ensure smooth communication as in the following excerpts:

> Masami: I didn't have such a case here, but usually I wouldn't tell someone if I disagree (with) their opinion because I don't want to lose nice atmosphere. This is Japanese culture. Do you know omoiyari?

> Masahiro: I couldn't show I didn't understand because it's the Japanese mind. If I show, he loses his face, and I too lose my face.

The non-observance of the maxim of quality here is clearly not a violation in which the participants are purposefully lying and deceiving for selfish gains. Rather, these non-observances fit into the category of suspending a maxim, which, according to Murray (2011), occurs when a person does not observe a maxim due to various cultural or contextual factors. In the case of Cutrone's (2005) study, the JEFL respondents, who feigned understanding and agreement in the intercultural conversations, attributed this type of behavior to a cultural norm which stresses being polite, keeping conversations harmonious and avoiding confrontations. Nonetheless, while the JEFLs' intentions may have been good, feigning understanding and agreement appear to be in contrast to what some of the British participants desired of their JEFL interlocutors as shown in the following utterances:

> Victoria: Honestly I think she's just agreeing with everything I say no matter how she feels which is too bad because I'd like to hear her opinion; it would do heaps to stimulate the conversation.

> Charles: I don't know why they (Japanese EFL speakers) just can't give their opinions. You (referring to all people) can disagree without hurting people.

> Berenice: In Japan, shopkeepers on the street always nod and act like they understand but they really don't, and then we try to buy something and it's like did we miss something.

> (Cutrone, 2005, p. 268)

The responses above suggest that one strategy that some Japanese may use to mitigate potential face threatening acts such as disagreeing is to hide their true feelings and, in some cases, simply convey the sentiments they believe their interlocutor desires. The final excerpts given in each of the cultural sets above by Masahiro and Berenice respectively touch upon another issue that has been presented in the literature regarding ELT in the Japanese context. Blanche (1987) and Cutrone (2005) are among the

many researchers to demonstrate instances when intercultural miscommunication occurred due to differing ways of showing understanding. It appears that there may be rule-conflict processes at work that might hinder a JEFL speaker from showing that they do not understand. The example given above by Noguchi (1987) in which Mr. Suzuki had a particularly difficult time showing that he did not understand is a case in point.

Other attempts to explain the Japanese concept of truth relate to some of the cultural concepts mentioned above. For instance, situational variability, and particularly the dichotomy between the private self and the public self, can play an important role in a Japanese person's life. According to Doi (1986), this notion of the Japanese shifting self may be at the crux of some intercultural misunderstandings with Americans. Doi (1986) suggests that in American society it is thought to be important for these two selves to remain consistent; when the private self deviates from the public self, an individual might be considered to be a hypocrite. In Japanese society, however, being polite and preserving harmony is given a much higher priority, and an individual's actual feelings pertaining to an action are thought to be less important (Doi, 1986; Triandis, 1989). Lapinsky and Levine (2000) summarize these sentiments by stating that in collectivist cultures, there is not as strong an emphasis on maintaining consistency between what one feels and what one says, whereas in individualist cultures consistency between thoughts and actions is believed to be extremely important. This is not to conclude, however, that individualists always maintain consistency between thoughts and actions, but that there seems to be a greater emphasis placed on consistency by those from individualistic cultures.

Part 4: Understanding the Maxim of Relevance

The maxim of relevance is stated by Grice (1975, p. 46) as "make your contribution relevant and timely". Adherence to this maxim prevents random, incoherent conversations lacking continuity. The precise role of this maxim and its significance in relation to the other maxims continues to be a source of great discussion. In terms of adhering to the CP, the maxim of relevance may be the most difficult and far-reaching maxim to infringe (i.e., it can only be done through unintentional failure to provide relative information due to inferior language ability, cognitive disorder, physical problems and mistakes/slips). If any of the other maxims are not observed in these ways, it can be argued that communication of some kind can still go on. While the example of the Pope Question (*Is the pope catholic?*) above has shown how the maxim of relevance can be flouted to create a conversational implicature, it is not likely that the CP would be upheld at the point where the maxim of relevance is

infringed. For instance, a reply of "*Obama will be re-elected*" to the question "*Do you like sushi?*", would not make any sense and would, thus, severely impede the progress of the conversation.

It has also been pointed out that the other three maxims can, to varying degrees, be viewed in terms of the maxim of relevance (Talib, 2009). For instance, an undetected lie is thought to violate the maxim of quality, as the speaker is deliberately uttering a falsehood, but in a sense, it may also violate the maxim of relevance, as the same utterance to the listener is an utterance which is propositionally true, and not one that is false. Similarly, as Talib (2009) points out, in a hypothetical situation whereby five units of information are needed by the listener, but four or six units of information are uttered by the speaker, then, either one relevant unit of information is not given, or one irrelevant unit of information is given, which indicates that both the maxims of quantity and relation may have been violated. Sperber and Wilson (1986) further expanded upon Grice's (1975) principle of relevance in developing what is known as Relevance Theory. This principle, according to Sperber and Wilson, states that the utterance given has to be relevant for it to be understood, and presumes that the receiver will have available the contextual information necessary to derive the meaning of the utterance with minimum effort.

Examining Grice's Maxim of Relevance across Cultures

While the Gricean (1975) approach recognizes the importance of the maxim of relevance, it does not subsume all the other maxims under it. Similar to the other three maxims discussed above, interpretations as to what is deemed relevant and timely may differ according to each individual's personality, interests, age, gender, culture and subculture(s), etc. For a JEFL speaker, it may sometimes be a matter of simply not being proficient enough in L2 English to adequately connect what appear to be disparate thoughts and ideas. So, for instance, if a person were to abruptly change the topic in a conversation and did not utter something like "By the way, …" or "Changing the topic completely, …" before starting to discuss the new topic, their sudden topic shift might be perceived as awkward or confusing by the listener(s). While the link between English proficiency and the ability to produce coherent and connected speech seems quite straight forward, the role of Japanese culture, and whether JEFLs may feel somewhat less pressure to remain on topic in a conversation than other cultures, is worthy of further exploration.

Possible Explanations of Ideological Differences, and Call for more Research in this Area

Unfortunately, research into whether culture-specific ideologies affect this phenomenon has been scant. Much of the earlier discussion in this area centered on cross-cultural literacy styles and can be traced back to the controversial and oft-cited paper of Kaplan's (1966) graphic depiction of various modes of discourse structure according to exhibited patterns of textual development. According to his theory, linearity is at least a prima facie requirement of Anglo rhetorical patterning, whereas circuity is thought to characterize an East Asian rhetorical pattern. Regarding the latter, this circuity rhetorical mode seems to involve texts in which the writer avoids a direct delineation of thesis (i.e., statement of topic) in the opening sections of text, and it is left to the reader to extrapolate a position from seemingly unrelated facts or situations in the text. Discourse development seems to follow a pattern of turning and turning in a widening gyre, with the loops revolving around the topic and viewing it from a variety of positions, but never addressing it directly (Brown, 1998). In a similar way, Mulvey (1997) describes Japanese preferred rhetorical strategies as identified by Hinds (1980, 1983, 1984), Mulvey (1992), Ricento (1987) Takemata (1976), and Yutani (1977), among others. The overriding element found in three commonly used rhetorical strategies in Japanese texts involved seeming irrelevance, i.e., texts which seemed to contain a series of seemingly disconnected and semi-connected topics. While such an approach may indeed possess a coherent method of organization in their own way in Japan, students using such strategies in EFL/ESL classes with Western teachers risk having their efforts mistaken for poor organization (Hinds, 1983; Ricento, 1987). While this seems to lend support to the notion that the maxim of relevance may be interpreted differently across cultures, it is clear that more research is necessary to complete the picture, particularly involving how this phenomenon might affect spoken discourse.

In attempting to explain the circuitous rhetorical pattern described above and to draw insights into how the notion of relevance is thought to be interpreted in the Asiatic tradition, Leki (1991) asserts that rhetoric in the Asiatic tradition seems to have a historical purpose of announcing truth rather than proving it, whereas rhetoric in the Western tradition, conversely, often seems to be designed to convince people towards a certain position. Consequently, in the Asiatic tradition, the speaker/writer arranges the propositions of the announcement in such a way that references to a communal, traditional wisdom encouraging harmonious agreement, while in the Western tradition, much more prominence is placed on the speaker/writer's ability to reason and marshal evidence in order to persuade the reader/listener towards a certain position. Thus, consistent with the descriptions above, Brown (1998)

summarizes the Asian mode of text development as deferential, anecdotal, and circuitous, one which seeks to address an issue by describing the surrounding terrain. Fliegel (1987) concurs and goes on to point out three defining characteristics of this rhetoric: an emphasis on group collectivity, the elicitation of consent, and the avoidance of direct conflict.

Conclusion

From Theory to Practice

In conclusion, this paper has argued that culture-specific interpretations of Grice's maxims do indeed exist. With misunderstanding as the focus of this theoretical analysis, the writer found it useful to examine intercultural conversations using a Gricean framework. That is, intercultural misunderstandings and negative perceptions were often associated with differing interpretations of Grice's (1975) maxims. These findings would seem to question the notion that Grice's (1975) maxims can apply universally and independently of culture. In this way, Grice's (1975) maxims would seem to better serve linguists engaged in cross-cultural research as a tool for analysis rather than a set of norms expected in conversations. If researchers were seeking prescriptive laws to govern intercultural conversations, a good place to start would be with Clyne's (1994, p. 194) culturally sensitive modifications and additions to three of Grice's maxims, as follows (N.B. No revisions were suggested for the maxim of relevance):

Quantity: A single maxim – 'Make your contribution as informative as is required for the purpose of the discourse, within the bounds of the discourse parameters of the given culture.'

Quality: Supermaxim – 'Try to make your contribution one for which you can take responsibility within your own cultural norms.' Maxims (1) 'Do not say what you believe to be in opposition to your cultural norms of truth, harmony, charity, and/or respect.' (2) Do not say that for which you lack adequate evidence.'

Manner: The supermaxim can be retained in its original form – 'Be perspicuous.' Maxims (1) 'Do not make it any more difficult to understand than may be dictated by questions of face and authority.' (2) Avoid ambiguity unless it is in the interests of politeness or of maintaining a dignity driven cultural core value, such as harmony, charity or respect.' (3) 'Make your contribution the appropriate length

required by the nature and purpose of the exchange and the discourse parameters of your culture.' (4) 'Structure your discourse according to the requirements of your culture.' (5) 'In your contribution, take into account anything you know or can predict about the interlocutor's communication expectations.'

Undoubtedly, Clyne's (1994) revisions to Grice's (1975) maxims provide several useful pieces of advice to anyone embarking on international communication; however, from the perspective of Japanese EFL learners, using the original Gricean framework to analyze intercultural encounters seems to be a good way to shed light on the potential origins of pragmalinguistic failure. That is, by understanding why the pragmalinguistic failure is occurring, both teachers and students alike will be in a much better position to prevent and/or deal with the awkward moments that result in intercultural misunderstandings.

Before any practical suggestions can be given however, it is important that EFL teachers understand and respect how potentially sensitive cross-cultural issues can be in the EFL classroom. Accordingly, the writer cautions EFL teachers to never push their learners to communicate in ways that make them feel uncomfortable. It is important to keep in mind that students forge their identities and belief systems through their culture; thus, any attempts by teachers to alter their communicative style or behaviors may be met with resistance and disengagement from the class. Unquestionably, the degree to which learners choose to conform to new communication practices is entirely up to each individual. Still, many JEFLs are more than willing to adopt cultural practices different than their own in their efforts to learn English, and those who are reluctant at first often seem to find their way over time as they become more acclimatised to the norms of the international community. Since many learners may not have previously considered how differing communication styles can impact IC, raising their awareness of these issues would be a good first step in conversation classes. Ultimately however, to have successful conversational exchanges, learners will have to go beyond the raised-consciousness phase and on to developing more product-oriented conversational management techniques. To these ends, the writer would like to suggest the following two phases of instruction in this area, as advocated by Ellis (1991).

(1) The first phase of instruction would involve consciousness-raising activities. Such activities are designed to raise awareness of how conversational behaviors might be perceived (and sometimes misconstrued) across cultures. For instance, concerning the cross-cultural analysis of Grice's maxim of quantity provided above, there is ample evidence of Japanese people speaking much less than their

interlocutors in intercultural encounters in English, and this tendency to adopt more of a listening role (and failure to take on more of the primary speakership responsibility) has seemingly affected intercultural encounters in a negative way (i.e., miscommunication, stereotyping and negative perceptions across cultures). One of the ways which EFL teachers in Japan can raise their learners' awareness in this area is by having students watch video clips and/or listen to audio excerpts which were specifically chosen (or constructed) to highlight particular features of conversation and, subsequently, engage in deconstruction/discussion activities to become more aware of how different communicative styles might affect IC (N.B. The conversations used in this phase of instruction could easily be modified to highlight issues concerning any of the four maxims). To provide an example that highlights the effects of not speaking enough in a conversation (i.e., concerning the maxim of quantity), a teacher could administer the following two steps:

A. First, the teacher would present two conversations to their students (via audio or video): in the first conversation, one of the conversational participants would be seen carrying the conversation (in terms of taking on most of the primary speakership responsibilities), while the other participant is seen adopting more of a listening role (and relying mostly on providing short listener responses as a way to stay involved in the conversation); in the second conversation, the primary speakership of the conversation would be shown to be much more balanced between the conversational participants.

B. Second, after each conversation, the teacher would engage students in a deconstruction of what they thought happened in the conversation and how it might be perceived by different people, including themselves. Depending on the teacher's context and preferences, this could involve having the students sit in a circle (or in small groups) and discuss the following questions: *What did you think of the conversation? Did you notice any differences in the behaviors of the conversational participants? If so, do you have any guesses to explain why these differences occurred? How do you think the conversational participants perceived these differences? How do you think people in your culture would perceive such behaviors? Etc.*

Ideally, the students would bring up the issue of "how much each person spoke" on their own after watching/listening to the first conversation, which could elicit more specific questions and in-depth discussion of this topic; however, if students do not bring it up on their own, the teacher should be prepared to prompt students with various questions such as: *Did you notice if one conversational participant spoke more than the other? If so, how do you think this affected the conversation? Did you think the participant who spoke less was interested in what the other one was saying? Did you think the*

participant who spoke less was bored by what the other participant was saying? How do you think the person who spoke more felt about the person who spoke less? How do you think the person who spoke less felt about the person who spoke more? Etc.

At some point, teachers should steer students attention away from the recorded conversations and have them reflect more on their own general conversational behavior (in both L1 Japanese and L2 English) by posing the following questions: *Do you behave like any of the participants you observed at times? How do you think others perceive your behavior? How do you think your conversational behavior would be perceived across cultures? Does your communication style change when you speak English? If so, why? How would you change your conversational behavior if you could? Do you behave differently when you are speaking to your teacher, boss, or family? Do you think you act this way because of your personality or the personality of the other speakers? How much influence does the situation have on your listening behavior? Etc.* At the end of this step, it is hoped that the students will have developed a greater appreciation of how certain communicative behaviors might be perceived across cultures. In many cases, this serves to provide JEFLs with the impetus to want to speak more in their intercultural encounters in English.

(2) In the second phase of instruction, students focus more on developing strategies for application. In other words, continuing with the example above, students move from understanding why they should speak more (in the first phase) to actually learning how to do so (in the second phase). To this end, a teacher could administer the following two steps:

A. First, the teacher would have the students engage in exercises that allow them to carefully analyze the target language and develop strategies to speak more. In this initial skill-building stage, most of the analytical activities would involve written texts (with matching, multiple choice and cloze exercises) and, thus, do not involve the real time-pressure and spontaneity of authentic conversations. Instruction on how to speak more could focus on such conversational management areas as new topic initiation, expansion techniques, agreeing and disagreeing, supporting one's opinions, asking return and follow-up questions, the ability to initiate repair when there is a potential breakdown, etc. For instance, the teaching of expansion techniques could involve presenting students with a sample conversation (or the transcription of the first conversation above in which one participant spoke much more than the other), such as the one that follows:

John: I really love European art. (silence) I went to the Louvre Museum in Paris last year.

Kenji: Cool.

John: Yeah, I have family in Italy, so I spent some time in Rome as well.

Kenji: I see.

John: I wish there were more art museums in this area.

Kenji: Me too. (silence)

John: Hmm, ok then, well, I guess I should get going; I'll talk to you later.

Kenji: See you later.

Examining this conversation with the class, the teacher shows the class how Kenji might have been able to expand on some of his initial utterances by including a fact, opinion, or question. For instance, rather than just saying "Cool" to John's initial statement "I really love European art", Kenji could have responded, "Cool, there's an exhibit on Spanish art this weekend in Tokyo" (fact), "Me too, I especially love the work of Leonardo da Vinci" (opinion)/ Oh really, I'm more of an Asian art lover myself" (opinion), or "Cool, who is your favorite artist?" (question). Once students see some examples such as these, the teacher can ask them to rewrite the rest of the conversation by adding a fact, opinion and/or a question to their initial response, as follows:

John: I really love European art.

Kenji: *Me too; in fact, I'm planning to go to Italy next year.*

John:

Kenji:

John:

Kenji:

To prepare students for the next step, teachers should have students complete similar types of language-building exercises and subsequently provide extensive feedback on their work. Further, to give students the best chance to succeed, teachers would be well-advised to begin with easier tasks and gradually progress to more difficult ones as students gain competence and confidence. Hence, each new sub-skill should be introduced separately, and new dimensions and complexities should be added only when students show they are ready to take the next step.

B. Second, in an effort to apply what they have learned in the first step, the teacher provides students with practice opportunities and feedback. Unlike the first step (which afforded students the time to think things through), this step aims to simulate a real conversation and the constraints that go with it (i.e., spontaneity, real-time pressure, all dimensions of conversation mixed together, etc.). That

is, the students would participate in role-plays or conversations (with or without prompts, depending on how much assistance students need) and focus on applying the new conversational techniques they learned in the previous step. The teacher and/or other students should observe the conversations and offer constructive feedback, and the students should change partners/groups and repeat the practice-feedback cycle as many times as possible. Ideally, each of these conversations would be recorded (video is preferred but not necessary) and played back for reference as the teacher and/or peer provides the conversational participants with feedback.

The Importance of Avoiding Cultural Stereotypes

Finally, as the writer alluded to above, the concept of culture, as well as cultural differences, can pose great dilemmas in the classroom, as well as to researchers comparing the communication styles of different groups (see Guest 2006 for an in-depth discussion of this issue). Regarding the latter, a common categorisation of culture is based on nationalities, as evidenced by Hofstede's (1991) oft-quoted study in which he compared the cultural values of citizens from over 50 different countries across four dimensions: Power Distance, Collectivism vis-à-vis Individualism, Femininity vis-à-vis Masculinity and Uncertainty Avoidance. While Hofstede's (1991) study is indeed interesting in that it points out various national differences, it is problematic to view cultures purely from the perspective of nationalities as there is sure to be a wide range of subcultures within any nation. In Japan, this may involve the management subculture in the occupational dimension, the large corporation subculture in the firm-size dimension, the male subculture in the gender dimension, the Osaka subculture in the regional dimension, to name a few. Furthermore, it is important to recognize that individual differences within any group also exist. Certainly, a great deal of individual variation has occurred in the conversational exchanges referred to in this paper, and output has been influenced to varying degrees by the specific contexts of each conversation, the personality and demeanour of the participants, and the chemistry between the participants in the dyadic conversations, as well as seemingly peripheral variables such as the amount of sleep the participants had the night before and the mood of the participants at the time of the conversations. Thus, while this paper seeks to investigate how Japanese culture could potentially influence JEFL speakers' performance, it is important to proceed with extreme caution in arriving at any conclusions and/or generalisations where culture is concerned, so as to not fall into the culturist trap of reducing individuals to less than they are (Holliday, Hyde & Kullman, 2004). In other words, it would be imprudent, and a vast overgeneralization, to suggest that all Japanese

people (or any group for that matter) adhere to any one set of cultural principles or values; however, it is not a great leap to surmise that cultural influences do indeed affect the behavior of many Japanese people to varying degrees when speaking L2 English. Thus, in taking a cautious approach, it is the writer's modest hope that this paper will help raise awareness in this area and, thus, inform language pedagogy in Japanese EFL classes.

References

Akasu, K. & Asao, K. (1993). Sociolinguistic factors influencing communication in Japan and the United States. In W. Gudykunst (Ed.), *In Communication in Japan and the United States* (pp. 88-121). Albany: New York State University Press.

Andersen, P. (1994). A model of intercultural communication competence. In L. Samovar & R. Porter (Eds.), *Intercultural Communication* (pp. 7-24). Belmont, CA: Wadsworth Publishing.

Anderson, F. (1993). The enigma of the college classroom: Nails that don't stick up. In P. Wadden (Ed.), *A Handbook for Teaching English at Japanese Colleges and Universities* (pp. 101-110). New York: Oxford University Press.

Arima, M. (1989). Japanese culture versus schizophrenic interpretation. *Text, 9*(1), 351-365.

Armour, W. (2001). This guy is Japanese stuck in a white man's body: A discussion of meaning making, identity slippage, and cross-cultural adaptation. *Journal of Multilingual and Multicultural Development, 22*(1), 1-18.

Armour, W. (2004). Becoming a Japanese language learner, user, and teacher: Revelations from life history research. *Journal of Language, Identity, and Education, 3*(1), 101-125.

Barnlund, D. (1974). The public self and the private self in Japan and the United States. In Condon, G. & Sato, M. (Eds.), *Intercultural Encounters with Japan* (pp. 27-96). Tokyo: Simul Press.

Barnlund, D. (1989). *Communicative Styles of Japanese and Americans.* Belmont, CA: Wadsworth Publishing Company.

Barnlund, D. & Araki, S. (1985). Intercultural encounters: The management of compliments by Japanese and Americans. *Journal of Cross-Cultural Psychology, 16*(1), 9-26.

Blanche, P. (1987). The case for a pedagogy of pragmatics in foreign or second language teaching. *RELC Journal, 18*(1), 46-71.

Blum-Kulka, S. & Olshtain, S. C. (1986). Too many words: length of utterance and pragmatic failure. *Studies in Second Language Acquisition. 8*(1), 165-179.

Books, M. (1995). *Communication Styles of Japanese and Americans: Six Key Styles Explored.* Cincinnati, Ohio: The Union Institute.

Bouton, L. (1994). "Conversational implicature in a second language: Learned slowly when not deliberately taught." *Journal of Pragmatics, 22*(2): 157-167.

Brown, D. D. (1998) Academic protocol and targeted rhetoric. *Literacy Across Cultures. 2*(1). 1-16.

Brown, P. & Levinson, S. C. (1978). Universals in language usage: Politeness phenomena. In Goody, E. N. (Ed.), *Question and Politeness* (pp. 56-311). Cambridge: Cambridge University Press.

Brown, P. & Levinson, S. (1987). *Politeness: Some Universals in Language Usage.* Cambridge: Cambridge University Press.

Bruneau, T. (1973). Communicative Silences: Forms and Functions. *Journal of Communication, 23*(1), 17-46.

Canale, M. & Swain, M. (1980). Theoretical bases of communicative approaches to second language teaching and testing, *Applied Linguistics*, *1*(1), 1-47.

Chapman, D. & Hartley, B. (2000). Close encounters of the unhomely kind: Negotiating identity and Japan literacy. *Japanese Studies*, *20*(1), 269-279.

Chung, S. (1999). Nihongo no gengo hassoo ni tsuite [On conceptions about Japanese language]. Heisei 11 Nendo Nihongo Kyooiku Gakkai Shun Taikai Yokooshuu. *Proceedings from the 1999 Spring Conference of the Association of Japanese Language Education* (pp. 107-112). Tokyo, Japan: AJLE.

Churchill, L. (1978). *Questioning Strategies in Socio-linguistics.* Rowley, MA: Newbury House.

Clancy, P. M. (1986). The acquisition of communicative style in Japanese. In B. Schieffelin & E. Ochs (Eds.), *In Language Socialization Across Cultures* (pp. 213-250). Cambridge: Cambridge University Press.

Clyne, M. (1994). *Inter-cultural Communication at Work: Cultural Values in Discourse.* Cambridge: Cambridge University Press.

Copeland, L. & Griggs, L. (2001) *Going International.* New York: Random House.

Cutrone, P. (2005). A case study examining backchannels in conversations between Japanese-British dyads. *Multilingua - Journal of Cross-Cultural and Interlanguage Communication*, *24*(1), 237-274.

Cutrone, Pino. (2011). *Toward the learning of listener responses in the Japanese EFL context*. Unpublished doctoral thesis, University of Reading, Reading, UK.

Day, M. (1996). *Aimai no ronri: The logic of ambiguity and indirectness in Japanese.* Berkeley, CA: University of California.

Diez Prados, M. (1998). Say enough, but no more pragmatics to the aid of language teaching. *Encuentro. Revista de Investigación e Innovación en la clase de idiomas*, *10*(1), 53-63.

Doi, T. (1986). *The Anatomy of Conformity: The Individual Versus Society.* Tokyo: Kodansha.

Doi, T. (1996). The Japanese psyche: myth and reality. In C. Strozier & M. Flynn (Eds.), *Trauma and Self* (pp. 197-203). London: Rowman and Littlefield.

Donahue, R. (1998) *Japanese Culture and Communication.* Lanham, Maryland: University of America.

Ellis, R. (1991). Communicative competence and the Japanese learner. *JALT Journal*, *13*(1), 103-127.

Elwood, K. (2001) *Getting Along with the Japanese.* Tokyo: Ask.

Farooq, M. U. (2005). A model for motivating Japanese EFL learners through real-life questioning strategies. *The Journal of Nagoya Gakkugei University*, *1*(1), 27-42.

Fliegel, D. (1987). Immigrant professionals must speak American. *The Boston Globe,* 16 June, p. 2.

Gartsman, A. & Hughes, L. (2007). "Introduction to linguistics study guide". Honors Junior/Senior Projects, Paper 30, Northeastern University. Boston. Retrieved March 30, 2014 from http://hdl.handle.net/2047/d10006552

Graham, J. (1990). An exploratory study of the process of marketing negotiations using a cross-cultural perspective. In R. Scarcella (Ed.), *Developing Communicative Competence in a Second Language* (pp. 239-270). Rowley: Newbury House.

Greer, D. (2000). The eyes of Hito: A Japanese cultural monitor of behavior in the communicative classroom. *JALT Journal*, *22*(1), 183-195.

Grice, H. P. (1967). *Logic in conversation*. Paper presented at the William James Lectures, Harvard University, Cambridge, MA.

Grice, H. P. (1975). Logic and conversation. In P. Cole & J. Morgan (Eds.), *Syntax and Semantics 3: Speech Acts* (pp. 41-58). New York: Academic Press.

Grice, H. P. (1989). *Studies in the Way of Words*. Cambridge: Harvard University Press.

Gumperz, J. (1995). Mutual inferencing in conversation. In: I. Marková, I., C. F. Graumann & K. Foppa (Eds.), *Mutualities in Dialogue* (pp. 101-123). Cambridge: Cambridge University Press.

Gudykunst, W. & Nishida, T. (1993). Interpersonal and intergroup communication in Japan and the United States. In W. Gudykunst (Ed.), *Communication in Japan and the United States* (pp. 149-214). Albany, NY: State of New York Press.

Gudykunst, W. & Nishida, T. (1994). *Bridging Japanese/North American Differences*. Thousand Oaks, CA: Sage.

Gudykunst, W. & Nishida, T. (2001). Anxiety, uncertainty, and perceived effectiveness of communication across relationships and cultures. *International Journal of Intercultural Relations*, *25*(1), 55-71.

Gudykunst, W. B., Nishida, T. & Chua, T. (1986). Uncertainty reduction in Japanese-North American dyads. *Communication Research Reports*, *3*(1), 39-46.

Gudykunst, W. & Ting-Toomey, S. (1988). *Culture and Interpersonal Communication and Conversation*. Newbury Park: Sage.

Gudykunst, W. B., Yang, S. & Nishida, T. (1985). A cross-cultural test of uncertainty reduction theory: Comparisons of acquaintances, friends, and dating relationships in Japan, Korea, and the United States. *Human Communication Research*, *11*(3), 407-455.

Guest, M. (2002). A critical 'checkbook' for culture teaching and learning. *ELT Journal*, *56*(2), 154-161.

Guest, M. (2006). Culture research in foreign language teaching: Dichotomizing, stereotyping and exoticizing cultural realities? *Zeitschrift für Interkulturellen* Fremdsprachenunterrich, *11*(3), 1-19.

Haga, Y. (1998). *Nihongo no Shakaishinrin*. Tokyo: Ningen no Kagakusha.

Hall, E. T. (1981). *Beyond Culture*. Garden City: Anchor Books.

Haugh, M. (1998). Native-speaker beliefs about Nihonjinron and Miller's "law of inverse returns". *Journal of the Association of Teachers of Japanese*, *32*(2), 64-96.

Haugh, M. (2003). Japanese and non-Japanese perceptions of Japanese communication. *New Zealand Journal of Asian Studies*, *5*(1), 156-177.

Heinz, B. (2003). Backchannel responses as strategic responses in bilingual speakers' conversations. *Journal of Pragmatics*, *35*(1), 1113-1142.

Helgesen, M. (1987). False beginners: Activating language for accuracy and fluency. *The Language Teacher*, *11*(14), 23-29.

Helgesen, M., Brown, S. & Mandeville, T. (2007). *English Firsthand*. Tokyo: Pearson Longman.

Hendersen, L. M., Zimbardo, P. G. & Carducci, B. J. (1999). Shyness. In Hendersen, L. M. & P. G. Zimbardo (Eds.), *Encyclopedia of Mental Health.* San Diego, CA: Academic Press.

Hill, T. (1990). Sociolinguistic aspects of communicative competence and the Japanese learner. *Dokkyo University Studies in English*, *36*(1), 69-104.

Hinds, J. (1983). Contrastive rhetoric: Japanese and English. *Text*, *3*(2), 183-195.

Hofstede, G. (1991). *Cultures and Organizations*. Berkshire: McGraw-Hill Book Company Europe.

Holliday, A., Hyde, M. & Kullman, J. (2004). *Inter-cultural Communication*. New York: Routledge.

Horn, L. (2004). Implicature. In L. Horn & G. Ward (Eds.), *The Handbook of Pragmatics* (pp. 3-28). Oxford: Blackwell.

Hughes, H. J. (1999). Cultivating the walled garden: English in Japan. *English Studies*, *80*(6), 556-568.

Hymes, D. H. (1971). *On Communicative Competence*. Philadelphia: University of Pennsylvania Press.

Ikegami, Y. (1989). Introduction: Special Issue on Discourse Analysis in Japan. *Text*, *9*(3), 263-273.

Imai, M. (1981). *16 Ways to Avoid Saying No: An Invitation to Experience Japanese Management from Inside*. Tokyo: The Nihon Keizai Shimbun.

Ishi, S. & Bruneau, T. (1994). Silence and silences in cross-cultural perspective: Japan and the United States. In L. Samovar & R. Porter (Eds.), *Intercultural Communication* (pp. 246-251). Belmont, CA: Wadsworth Publishing.

Ito, Y. (1989). Socio-cultural background of Japanese interpersonal communication style. *Civilizations*, *39*(1), 101-136.

Iwata, Y. (1999). *Toward bridging the communication gap: Cross-cultural variation in Japanese-American business communication*. Unpublished doctoral dissertation, University of Mississippi, Mississippi, USA.

Johnson, C. & Johnson, F. (1975). Interaction rules and ethnicity. *Social Forces*, *54*(1), 452-466.

Kaplan, R. B. (1966). Cultural thought pattern in inter-cultural education. *Language Learning*, *16*(1/2), 1-20.

Keenan, E. O. (1976). On the universality of conversational implicatures. *Language in Society*, *5*(1), 67-80.

Kenna, P. & Lacy, S. (1994). *Business Japan: A Practical Guide to Understanding Japanese Business Culture*. Lincolnwood: Passport Books.

Kitao, K. & Kitao, S. K. (1989). *Intercultural Communication Between Japan and the US*. Tokyo: Eichosha Shinsha.

Knapp, K. & Knapp-Potthoff, A. (1987). Instead of an introduction: Conceptual issues in analyzing intercultural communication. In K. Knapp, E. Werner & A. Knapp-Potthoff (Eds.), *Intercultural Communication. Studies in Anthropological Linguistics 1* (pp. 1-13). Berlin: Mouton de Gruyter.

Koreo, K. (1988). On Language habits of the Japanese. *English Today*, *5*(1), 19-25.

Lakoff, R. (1973). The logic of politeness; or, minding your p's and q's. In T. C. Corum, C. Smith-Stark & A. Weiser (Eds.), *Papers from the Ninth Regional Meeting of the Chicago Linguistic Society, April 13-15* (pp. 292-305). Chicago: Chicago Linguistic Society.

Lakoff, R. (1989). The limits of politeness: Therapeutic and courtroom discourse. *Multilingua - Journal of Cross-Cultural and Interlanguage Communication*, *51*(1), 51-73.

Lapinsky, M. K. & Levine, T. R. (1989). Culture and information theory: the effects of self construal and locus of benefit on information manipulation. *Communication Studies*, *8*(1), 101-129.

Lebra, T. (1976). *Japanese Patterns of Behavior.* Honolulu: The University of Hawaii Press.

Leech, G. (1983). *Principles of Pragmatics*. London: Longman.

Leki, I. (1991). Twenty-five years of contrastive rhetoric: text analysis and writing pedagogies. *TESOL Quarterly*, *25*(1), 123-143.

Levinson, S. (1983). *Pragmatics*. Cambridge: Cambridge University Press.

Levinson, S. (2000). *Presumptive Meanings: The Theory of Generalized Conversational Implicature*. Cambridge: MIT Press.

Lindblom, K. (2001). Cooperating with Grice: a cross-disciplinary metaperspective on uses of Grice's cooperative principle. *Journal of Pragmatics*, *33*(10), 1601-1623.

LoCastro, V. (1999). A Sociocultural functional analysis of fragmentation in Japanese. *Multiligua: Journal of Cross-Cultural and Interlanguage Communication*, *18*(4), 369-389.

Lockley, T., Hirschel, R. & Slobodniuk, A. (2012). Assessing the Action Plan: Reform in Japanese High School EFL. *Electronic Journal of Foreign Language Teaching, 9*(2), 152-169.

Loveday, L. (1982). *The Sociolinguistics of Learning and Using a Non-native Language*. Oxford: Permagon.

Matsumoto, K. (1994). English instruction problems in Japanese schools and higher education. *Journal of Asian Pacific Communication*, *5*(4), 209-214.

Matsumoto, M. & Boyè Lafayette, D. (2000). *Japanese Nuance in Plain English.* Tokyo: Kodansha.

McCarthy, M. (2003). Talking back: "Small" interactional response tokens in everyday conversation. *Research on Language and Social Interaction*, *36*(1), 33-63.

McClure, W. (2000). *Using Japanese. A Guide to Contemporary Usage*. Cambridge: Cambridge University Press.

McCornack, S. A. (1992). Information manipulation theory. *Communication Monographs*, *59*(1), 1-16.

McCreary, D. R. (1988). *Japanese-US Business Negotiations: A Cross-Cultural Study*. New York: Praeger Publishers.

McCroskey, J. C. (1992). Reliability and validity of willingness to communicate scale. *Communication Quarterly*, *40*(1), 16-25.

McDermott, R. P. (1987). Inarticulateness. In D. Tannen (Ed.), *Linguistics in Context*. Norwood, NJ: Ablex.

Miller, L. (1994). Japanese and American indirectness. *Journal of Asian Pacific Communication*, *5*(1), 37–55.

Miller, R. (1982). *Japan's Modern Myth: Language and Beyond*. New York: Weatherhill.

Minoura, Y. (1988). The psychological reorganization of overseas experience after returning to Japan: A symbolic interactionist approach to returnees. *Shakai Shinrigaku Kenkyuu*, *3*(2), 3-11.

Morita, A. & Ishihara, S. (1989). *"No" To Ieru Nihon [The Japan That Can Say "No"]*. Tokyo: Kobunsha.

Mulvey, B. (1992). *Japanese and English rhetorical strategies.* Unpublished Masters' Thesis, California State University, San Bernardino, USA.

Mulvey, B. (1997). Japanese and English rhetorical strategies. *Literacy Across Cultures*, *1*(2). Retrieved August 20, 2009 from http://www2.aasa.ac.jp/~dcdycus/LAC97/rhetoric.htm

Mura, S. S. (1983). Licensing violations: Legitimate violations of Grice's conversational principle. In R. Craig & K. Tracy (Eds.), *Conversational Coherence*. Beverly Hills, CA: Sage.

Murray, J. C. (2011). Do bears fly? Revisiting conversational implicature in instructional pragmatics. *The Electronic Journal for English as a Foreign Language (TESL-EJ)*, *15*(2). Retrieved January 30, 2013 from http://www.asian-efl-journal.com/march03.sub3.php

Nakai, F. (1999, August). *Japanese communication features: a survey*. Paper presented at the 12th World Congress of Applied Linguistics, Waseda University, Tokyo, Japan.

Nakane, C. (1970). *Japanese Society*. Tokyo: Charles E. Tuttle.

Nishiyama, K. (1994, June). *Japanese negotiators: Are they deceptive or misunderstood?* Paper presented at the 23rd annual convention of the Communication Association of Japan., Tama city, Tokyo, Japan.

Nishiyama, S. E. (1995). Speaking English with a Japanese mind. *World Englishes*, *14*(1), 27-36. Nisugi, M. (1974). Images of spoken Japanese and spoken English. In J. Condon & M. Saito (Eds.), *Intercultural Encounters* (pp. 199-204). Tokyo: Simul Press.

Nisugi, M. (1974). Images of spoken Japanese and spoken English. In J. Condon & M. Saito (Eds.), *Intercultural Encounters* (pp. 199-204). Tokyo: Simul Press.

Nittono, M. (1999, August). *Strategic discourse-based hedges in Japanese: Their forms and functions in conversational interaction*. Paper presented at the Twelfth World Congress of Applied Linguistics, Waseda University, Tokyo, Japan.

Noguchi, R. (1987) The dynamics of rule conflict in English and Japanese conversations. *IRAL*, *25*(1), 15-24.

Nolasco, R. & Arthur, L. (1987). *Conversation.* Oxford: Oxford University Press.

Nonaka, K. (1996). Cross-cultural differences in involvement strategies: a case study of English and Japanese spoken discourse. *JACET Bulletin*, *27*(1), 143-168.

Nozaki, K. (1993). The Japanese student and the foreign teacher. In P. Wadden (Ed.), *A Handbook for Teaching English at Japanese Colleges and Universities* (pp. 27-33). New York: Oxford University Press.

Nunn, R. (2003). Intercultural communication & Grice's principle. *Asian EFL Journal*, *5*(1). Retrieved August 14, 2008 from http://www.asian-efl-journal.com/march03.sub3.php

Okabe, R. (1983). Cultural assumptions of East and West. In W. Gudykunst (Ed.), Intercultural Communication Theory. Current Perspectives (pp.21-44). London: Sage.

O'Keeffe, A. (2004). 'Like the wise virgins and all that jazz': Using a corpus to examine vague categorisation and shared knowledge. *Language and* Computers, *52*(1), 1-20.

Okushi, Y. (1990). Misunderstood efforts and missed opportunities: An examination of EFL in Japan. *Penn Working Papers in Educational Linguistics*, *6*(2), 65-75.

Peaty, D. (1987). False Beginners: Who they are and what to do with them. *The Language* Teacher, *11*(14), 4-5.

Reesor, M. (2002). The bear and the honeycomb: A history of Japanese English language policy. *NUCB JLCC (Nagoya University of Commerce and Business Journal of Language, Culture and Communication*, *4*(1), 41-52.

Ricento, T. (1987). *Aspects of coherence in English and Japanese expository prose*. Unpublished doctoral dissertation, UCLA, Los Angeles, USA.

Robinson, M. (1992). Introspective methodology in interlanguage pragmatics research. In G. Kasper (Ed.), *Pragmatics of Japanese as native and target language, Volume 3,* (pp. 29-84). Honolulu: University of Hawaii.

Roger, A. (2008). Teaching the speaking skill to Japanese students part 1: Construct & practice. *The Journal of Kanda University of International Studies, 20*(1), 1-26.

Rose, K. (1992). Method and scope in cross cultural speech act research: A contrastive study of requests in Japanese and English. Unpublished doctoral dissertation, University of Illinois, Urbana-Champaign, USA.

Rose, K. (1996), American English, Japanese and directness: more than stereo-types. *JALT Journal*, *18*(1), 67-80.

Ross, S. (1998). Divergent frame interpretations in oral proficiency interview interaction. In R. Young & A. Weiyun He (Eds.), *Talking and Testing: Discourse Approaches to the Assessment of Oral Proficiency* (pp. 333-353). Philadelphia: John Benjamins.

Sasagawa, Y. (1996). Ibunka no kanten kara mita Nihongo no aimaisei - zainichi gaikokujin ryuugakusei choosa yori [Ambiguity in Japanese communication - from the foreign students' viewpoint]. *Nihongo* Kyoiku, *89*(1), 52-63.

Sato, S. & Okamoto, S. (1999). Reexamination of Japanese "cooperative" communication style. In J. Verschueren (Ed.), *Pragmatics in 1998. Selected Papers from the Sixth International Pragmatics Conference, Volume 2* (pp. 518-527). Antwerp: International Pragmatics Association.

Sato, Y. (2008). Oral communication problems and strategies of Japanese university EFL learners. Unpublished doctoral thesis, University of Reading, Reading, UK.

Scollon, R. & Scollon, S. (1995). *Intercultural Communication: A Discourse Approach*. Oxford: Blackwell.

Scollon, R. & Scollon, S. (2001). *Intercultural Communication: A Discourse Approach (2nd edition)*. Malden, MA: Blackwell.

Spees, H. (1994). A cross-cultural study of indirectness. *Issues in Applied* Linguistics, *5*(2), 231-253.

Sperber, D. & Wilson, D. (1986). *Relevance: Communication and Cognition*. Oxford: Basil Blackwell.

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.

Susser, B. (1998), EFL's othering of Japan: Orientalism in English language teaching. *JALT Journal*, *20*(1), 49-82.

Takahashi, A. (1987). *A contrastive study of indirectness exemplified in L1 directive speech acts performed by American and Japanese*. Unpublished Masters' Thesis, University of Illinois, Urbana-Champaign, USA.

Takanashi, Y. (2004). TEFL and communication styles in Japanese culture. *Language, Culture and Curriculum*, *17*(1), 1-14.

Takemata, K. (1976). *Genkoo Shippitsu Nyuumon [An Introduction to Writing Manuscripts]*. Tokyo: Natsumesha.

Talib, I. (2009). Literary Stylistic Notes, No. 23. Retrieved May 18, 2008 from http://courses.nus.edu.sg/course/ellibst/lsl23.html

Townsend, J. & Danling, F. (1998). Quiet students across cultures contexts. *English Education*, *31*(1), 4-25.

Triandis, H. C. (1988). Collectivism vs. individualism. In G. Verma & C. Bagley (Eds.), *Cross-cultural studies of personality, attitudes, and cognition* (pp. 60-95). London: Macmillan.

Triandis, H. C. (1989). The self and social behavior in differing cultural contexts. *Psychological Review*, *96*(1), 506-520.

Triandis, H. C. (1994). *Culture and Social Behavior.* New York: McGraw-Hill.

Triandis, H.C., Bontempo, R., Villareal, M.J., Asai, M. & Lucca, N. (1988). Individualism and collectivism: Cross-cultural perspectives on self-ingroup relationships. *Journal of Personality and Social Psychology*, *54*(1), 323-338.

Vygotsky, L. (1962). *Thought and Language*. Cambridge: The M.I.T. Press.

Yamada, H. (1997). *Different Games, Different Rules: Why Americans and Japanese Misunderstand Each Other*. New York: Oxford University Press.

Yano, Y. (2001). World Englishes in 2000 and beyond. *World Englishes*, *20*(2), 119-131.

Yashima, T. (2002). Willingness to communicate in a second language: The Japanese EFL context. *The Modern Language Journal*, *86*(1), 54-66.

Yoshida, T., Matsumoto, D., Akiyama, T., Moriyoshi, N., Furuiye, A. & Ishii, C. (2003). Peers' perceptions of Japanese returnees. *International Journal of Intercultural Relations*, *27*(6), 641-658.

Yutani, Y. (1977). Current English: Translation of news articles and "nonsequence" of tenses. *Academic Bulletin: Kyoto Bulletin of Foreign* Studies, *18*(1), 52-63.

Reengaging Quiet EFL Students in CLT Classrooms: Teachers' Language Attitude and Students' Stereotype Threat

Ling-Hui Hsu

Ming Chuan University, Taiwan

Bio data

Ling-Hui Hsu has done most of her research in intercultural communication with a focus on intergroup relations and prejudiced communication. She joined the Department of Applied English at Ming Chuan University in Taiwan in 2009. Since then she has expanded her research scope to cover biased attitude (e.g., social bias, accent stereotype) in the context of EFL teaching. linghui@mail.mcu.edu.tw

Abstract

Many studies on communicative language teaching (CLT) in an EFL context have cited student reticence as a major hindrance to successful adoption of CLT (Chang & Goswami, 2011; Fang, 2011; Howard & Miller, 2008; Hu, 2010; Jeon, 2009; Li, 1998). By referring to internalized culture or ascribed personality traits to explain EFL students' quietness, we leave ourselves with limited solutions to the problem. It is worthwhile to look at the interplay of learner and teacher variables to find out why against the same cultural background, some students are forever silent and others ready to communicate. The current study identifies three subtle psychological factors, student stereotype threat, stigma consciousness, and teacher language attitude that purportedly relate to students' willingness to participate in class. A total of 192 students enrolled in the oral communication course at a local Taiwanese university were included as the student participants. The six American instructors teaching these students in intact classes were invited as the teacher participants. To evaluate students' level of stereotype threat and stigma consciousness, the measures used were stereotype threat index (Aronson, Fried & Good, 2002) and stigma consciousness questionnaire (Pinel, 1999).

Address correspondence:
Department of Applied English, Ming Chuan University, #5 De-Ming Rd., Gui-Shan District, Taoyuan City, Taiwan

The instrument measuring teacher's language attitude was adapted from Cargile's and Giles' (1997) semantic differential scale testing the mood state of listeners when listening to either an audio-tape spoken with an American accent or with a Taiwanese accent. The study results revealed that students low in stereotype threat outperformed those high in stereotype threat in oral communication classes. In addition, students taught by instructors with a positive attitude towards accented English also fared better than those taught by instructors with a negative language attitude. More significantly, when taught by instructors with a positive language attitude, students high or low in stereotype threat did not differ in their performances. The results point to the important role of instructors who may engage nonparticipating EFL students with a good attitude toward diverse speech styles.

Keywords: communicative language teaching, EFL, stereotype threat, stigma consciousness, language attitude, student reticence

Introduction

Communicative language teaching (CLT) as a methodology of teaching English as a second language was initiated in the late 1980s (Chang & Goswami, 2011). Since its inception, CLT has been heralded as a better way of teaching and learning a foreign language (Hiep, 2007; Savignon, 1991). Starting from 1989 into mid-1990s, communicative language teaching (CLT) as a teaching pedagogy has been officially adopted in Asian countries such as China, Japan, Korea and Taiwan where English is taught as a foreign language (Nishino, 2008; Yuasa, 2010; Zhang & Wang, 2012). Over the years, CLT has gained popularity in general and taken up the mainstream position in foreign language teaching as reflected in government policies and school curricula in various Asian countries (Howard & Millar, 2008; Tao, Yu, & Jin, 2011; Yuasa, 2010). Irrespective of this, CLT initiation and implementation in EFL settings has been bumpy with a myriad of outstanding problems (Chang & Goswami, 2011; Hato, 2005; Jeon, 2009). The question is not about whether to adopt CLT in EFL countries but to learn more as to how to implement such a teaching pedagogy optimally.

A review of related literature suggests that parallel problems have been cited in studies on different EFL countries over the past decade (Chang & Goswami, 2011; Howard & Miller, 2008; Jeon, 2009; Li, 1998; Sakui, 2004; Takako, 2012). The documented obstacles impairing successful implementation of CLT in EFL contexts can be summarized as follows:

First, some local English teachers lack the required English proficiency or sociolinguistic skills to conduct a communicative class.

Second, the grammar-oriented or form-based instruction is still vital to the exam-focused educational system in many Asian countries. To excel in school or national examinations, students need to master the mechanics of English rather than cultivate communicative competence. For instance, the educational reform in Taiwan since 2001 has refocused the target of English education on English teaching for communicative purposes. However, in reality performance in English standardized tests still serves as the measure of students' achievements and as a determinant for their continuing education and upward social mobility. Such a social belief has created the status quo of teachers continuing to "teach to the test" (Chang & Su, 2010, p. 265) although teachers in high schools and colleges in Taiwan have generally endorsed the merit of CLT teaching pedagogy (Liu, 2005).

Third, English classes in most Asian countries are too large to encourage group work among students or facilitate communicative teaching. For instance, in some colleges in China, there can be more than one hundred students in one English classroom (Hu, 2010, p. 79).

Fourth, most teachers are under time pressure to implement the rigid curriculum schedule.

Fifth, students are reticent and passive in engaging in communicative activities in class. Indeed, many EFL teachers have the shared experience of facing quiet students in a supposedly interactive class, which is all the more true for teachers with older students in college classrooms (Fang, 2011). In fact, as pointed out earlier, in various studies from 1998 to 2011, EFL students' resistance to participation has been cited as one of the major obstacles to successful adoption of CLT and as the number one hindrance in the study of Chang and Goswami (2011) on factors that affect CLT implementation in Taiwanese college English classes.

It takes time and more than academic effort to bring about substantial changes to such external factors as educational system, traditional practice or unique social condition like exam-driven learning or population constraint in individual EFL countries (Liu, 2005). As such, the current study tackles the more micro aspects of the existing barriers in adopting CLT in EFL countries. This is to suggest that the focus is on the key players of the teaching pedagogy – teachers and students. In spite of the support of national policies and school curricula, only teachers can decide on the agenda of each and every class session (Chang & Goswami, 2011). According to Lee and Ng (2010), teachers are the directors in a class who decide on students' opportunities to participate in class and should be included in reticence studies as a situational variable. Besides teachers, learners are crucial in the CLT process which has the objective of improving learners' communicative competence through engaging them in

meaningful interactions. As pointed out by Tao, Yu, & Jin (2011), teachers provide guidance, whereas students are the "real participants and controllers" in a CLT class (p. 692).

The current study differs in two aspects from other cited research on EFL student reticence in CLT classes in Asia. First, although believing such external factors as educational system or curriculum focus play a role in CLT results, the current study centers on the interaction between the entities that truly live through the experience of CLT teaching and looks at subtle psychological processes that might be shaping CLT effect during the interactions of these two CTL enactors (i.e., teachers and students). Second, the current study is different in its approach to treat students' unwillingness to communicate as a situation-based variable which might be triggered by specific situational factors. Arguably, not all quiet students are reserved due to some enduring personality traits. By attributing problems to social, cultural, or personality factors, we are left with a very limited number of solutions (Aronson, Fried, & Good, 2001). Rather, quite students might be "ability-stigmatized students" (Good, Aronson, & Inzlicht, 2003, p. 658) in English ability who suffer from the stereotype threat of living up to the stereotype of being quiet with poor English, and they are conscious about the stigmatized stereotype about self (Pinel, 1999). While it is agreed that teachers should encourage student participation by being warm and enthusiastic, it is all too perfect to assume that all EFL teachers show patient understanding to quiet students. Conceivably, quiet and often underachieved students would grow more withdrawn in reacting to teachers with a negative attitude who favor students with better English and a nicer accent; that is, teacher's negative language attitude towards accented English (Williams, Whitehead, & Miller, 1972).

Ultimately, the purpose of the study is to examine the protracted problem of reticent EFL students in CLT classes by exploring beyond the factors commonly cited in existent CLT literature. Specifically, the current study narrows in on interactants in a CLT class and the state of mind of teachers and students derived from their interactions. How EFL learners feel about themselves, their peers and teachers might influence their willingness to speak in a communication situation. It is posited that stereotyped or stigmatized students with lower English capability are likely to remain quiet, and their willingness to speak up could be further discouraged by teachers with a negative language attitude toward accented English speech. If the proposed psychological factors are found to be related to students' willingness to speak in class, the fact that these factors are situation-induced should provide CLT educators with the hope of potentially beneficial interventions.

Theoretical Background

Classroom interaction, language output in SLA

The essence of CLT (student-centered approach with a focus on building communicative ability) parallels the key implication of the output hypothesis for second language pedagogy -- providing learners with adequate opportunities for speaking in class (Swain, 1992). According to the output hypothesis, language acquisition may occur through language production (Swain, 1985). Further, the interaction hypothesis suggests that negotiated interaction in second language facilitates SLA (Long, 1996). Through interaction, learners are provided with the opportunities to modify their own output based on received comprehensible input and negative feedback from their conversational partners (Mackey, 2002). Language educators all over the world have long recognized target language producing as crucial in language learning (Izumi, 2003). Nevertheless, EFL students' resistance to speak up in classroom has been a persistent problem begging for more exploration.

Willingness to communicate

It is commonly believed that cultural values and personality traits such as being anxious, shy, and fearful to speak up in front of others might influence students' willingness to participate in class (Chang & Goswami, 2011; Jeon, 2009). Such factors were conceptualized by MacIntyre and colleagues (MacIntyre, Clement, Dörnyei, & Noels¸1998) as the enduring influences (versus situational influences) on willingness to communicate (WTC) in a second language. Enduring influences are the more stable variables that supposedly influence one's willingness to communicate over time and across situations. Often cited enduring factors are personality traits such as "communication apprehension, perceived communication competence, introversion-extraversion, self-esteem" that underlie language learners' WTC (MacIntyre, Clement, Dörnyei, & Noels¸1998, p. 546). Situational influences refer to the more immediate variables inherent in an environment or individual such as topic, participants, and the physical setting (Lee & Ng, 2009), as well as relationships between interlocutors, number of people involved and degree of formality in a communication situation (MacIntyre, Clement, Dörnyei, & Noels¸1998).

Much research has been done on WTC in L1 (McCroskey & Baer, 1985; McCroskey & Richmond, 1987, 1990) and in L2 (Clement, Baker, & MacIntyre, 2003; MacIntyre & Charos, 1996; MacIntyre, Clement, Dörnyei, & Noels¸1998). While many variables were found to be predictors of an individual's WTC, communication apprehension and perceived communication competence turned out

to be the most immediate determinants of WTC (MacIntyre, 1994; MacIntyre, Babin, & Clement, 1999). To improve WTC is to reduce communication apprehension and enhance perceived communication competence in students. Informed by an earlier study of McCroskey and Richmond (1987), MacIntyre and Charos (1996) suggested that communication apprehension was the "best predictor of WTC" (p. 217), and highly anxious people tended to rate themselves as lower in perceived competence.

In L2, communication apprehension is known as language anxiety (Horwitz, & Cope, 1986) composed of communication apprehension, test anxiety, and fear of negative evaluation. Fear of negative evaluation may exert a heavy toll on young students who are normally highly attentive to peer perception. Students receiving negative evaluation (e.g., teacher's negative feedback or classmates' ridicule) may form a self-perceived image of being students with low language ability. The image might become fixed as students continue to remain reticent out of fear and anxiety which negatively impact their language proficiency in the end. The fixed image may turn into a stereotype to haunt the quiet students who may then fear of being reduced to that stereotype. To avoid the dire prospect, the students may opt for disidentifying with the situations that arouse fear in them (e.g., withdrawing from class participation). These students are suffering from what Steele and his colleagues called "stereotype threat" (Steele & Aronson, 1995; Spencer, Steele, & Quinn, 1999), one of the main foci of the current study.

Stereotype Threat

Individuals experiencing stereotype threat are in a predicament in which they bear the extra pressure or apprehension of confirming a negative stereotype in association with their group membership (Rosenthal, Crisp, & Suen, 2007; Spencer, Steele, & Quinn, 1999; Steel & Aronson, 1995). A stereotyped individual is self-conscious and anxious when asked to perform a stereotype-relevant task. In the first study on stereotype threat, Claude Steele and Joshua Aronson (1995) concluded that making racial stereotype (i.e., Blacks lack intellectual ability) salient or relevant would negatively impact performance of Black students (versus White counterparts) on a verbal ability test. Various studies on stereotype threat have been conducted in different settings ever since, mostly are race-based (Aronson, Quinn, & Spencer, 1998; Aronson, Fried, & Good, 2002; Chateignier, Dutrevis, Nugier & Chekroun, 2009) and gender-based (Spencer, Steel, and Quinn, 1999; Rosenthal & Crisp, 2006; Brown & Panel, 2003). Although no literature on stereotype threat was found to directly address the issue of EFL

learning, the induced anxiety and stereotyped membership entailed in the definition of stereotype threat led this researcher to make the following assumption.

It is believed that disenfranchised students in EFL classes might be victims of stereotype threat. Stereotyped individuals are concerned about fulfilling a negative stereotype true of one's group and true of oneself (Wout, Danso, Jackson, & Spencer, 2008). That is, quiet students (versus actively participatory students) are supposedly anxious in a CLT class due to their stereotyped group membership of performing poorly in a communication class. In school settings, where evaluation of ability is routine, intellectual inferiority stereotypes have become inevitable (Aronson, Fried, & Good, 2002). Indeed, school is a place where ability is judged and individuals are categorized, which is all the more true for educational institutions in Asian countries. Students know clearly which group membership (quiet versus active) they and others hold. Such a group membership is significant since being participatory and particularly, being able to speak like a native, is generally believed to be an indication of excellence in a CLT-oriented EFL class.

A significant implication of the stereotype threat phenomenon in communicative language teaching classrooms lies in the stereotyped targets' responses to stereotype threat. Stereotyped students may devalue a specific domain that brings about a self-evaluative threat (Aronson, Fried, & Good, 2002; Spencer, Steele, & Quinn, 1999). Conceivably, students who see themselves as part of the group with lower language capabilities may experience perpetual anxiety of confirming the negative stereotype of underperformance (e.g., speaking poorly in a class) anticipated both by others and themselves. In the end, they become marginalized and withdrawn in class, rendering a successful implementation of interactive teaching programs impossible.

Stigma Consciousness

In a study on the possibility of internalized inferiority in stereotyped individuals, Aronson and colleagues (1999) argued that repeated exposure to one's group membership stereotype "breeds an awareness of stigma" (p. 31). Stigma consciousness is a response to experienced stereotype threat. According to Pinel (2002), stigma consciousness refers to the phenomenon in which stereotyped individuals differ in the extent they believe their stereotyped status pervades their interactions with outgroup members. In much of the literature on stereotype threat, stereotyped groups are also stigmatized groups suffering from stigmatization (Good, Aronson, & Inzlicht; 2003; Lawrence & Crocker, 2009). Pinel (1999) found that stereotyped individuals differ in how often and how much they

expect to be prejudiced or discriminated against. That is, they vary in level of stigma consciousness as dubbed by Pinel. High stigma conscious individuals are chronically conscious of their stigmatization whereas people low in stigma consciousness pay rare attention to their stigmatized status.

Of particular relevance to the current research is the study conducted by Smith, Kausar, and Holf-Lunstad (2007) to investigate how differences in stigma consciousness influence Pakistani women in science versus non-science fields. Smith et al. (2007) concluded that high stigma conscious individuals are more inclined to withdraw from situations in which they are potentially stereotyped, a similar reaction found in individuals suffering from stereotype threat. In fact, some social psychologists pointed out that stigmatized individuals tend to disengage either psychologically or physically from stigmatizing domains as a function of self-protection. Instances in which the academically disadvantaged choose to disidentify with school can possibly attest to such an argument (Major & Schmader, 1998; Schmader, Major, Eccleston, & McCoy, 2001).

In light of the literature discussed so far, the review may well facilitate a new look at the student variable concerning EFL students' reticence in CLT classes in Asia. Rather than being culturally reserved, they might be suffering from stereotype threat and stigma consciousness to remain quiet out of fear and consciousness about their stereotyped selves. The next section explores a psychological mechanism underlying behaviors of language teachers, the director in a class as mentioned earlier. Subtle verbal or behavioral cues of teachers might inspire or quiet anxious EFL students.

Teacher Language Attitude

Besides conveying semantic-referential information, language also provides hearers with linguistic or paralinguistic cues to make inference of personal or social characteristics of the speaker (Cargile & Giles, 1997). Attitudinal studies on language attitude were started in the 1960s. Research in this regard has mostly yielded the same result in which foreign-accented speech is evaluated less positively than native speech, and regional varieties of English are ranked lower on socio-economic traits than standard English (Anisfled, Bogo, & Lambert, 1962; Bourhis and Giles,1976; Kristiansen and Giles,1992, Lambert, Hodgson, Gardner, & Fillenbaum, 1960; Mugler, 2002; Tsurutani & Selvanathan, 2013; Tucker & Lambert, 1969). The fact that earlier experiments in language attitude have come to similar conclusions has provided little motivation for new research to be done in the same field (Hiraga, 2005). However, this does not mean dissipating importance of language attitude as a psychological construct. To the contrary, it has been a robust sociolinguistic phenomenon true for different people

and languages. In a recent study on language attitude, Akomolafe (2013) lamented over the relative neglect given to the issue of accent discrimination and the plight of foreign-accented speakers.

In the context of language teaching, teachers' perceptions of a language and speakers of that language have an impact on students' language acquisition (Cargile, Giles, Ryan, & Bradac, 1994) and teachers' view of the students (Dooly, 2005). For example, in earlier studies, teachers were found to have different academic expectancies for students in reaction to students' speech styles or dialect features (Williams, Whitehead, & Miller, 1972). In the study of Williams and Naremore (1974), the academic ratings teachers assigned to children of various ethnic groups conformed to ethnic and language stereotyping on the part of the teachers. Teachers' attitudes towards dialects were found to reflect teachers' expectations for students' academic competence, with teachers rating dialect speakers lower on an oral recitation than standard English speakers (Pringle, 1980). More recently, Walker, Shafer and Iiams (2004) found that teachers' attitudes towards linguistically diverse student (LDS, a student who speaks a language other than English) were largely neutral, with some reporting strongly negative. A more frustrating finding in their study is that 51 percent indicated no interest in professional development training even if it was available.

To integrate the three psychological constructs discussed earlier and apply them to the context of a CLT classroom in an EFL environment, it is posited that language teachers with a negative language attitude might worsen the predicament of stereotyped or stigmatized students. By holding a negative language attitude, a language educator might unknowingly aggravate vulnerability of these students to the self-imposed pressure of living up to a negative stereotype of inferior language capability.

The more recent argument on WTC being a dynamic system (Kang, 2005; MacIntyre & Legatto, 2011; Riasati, 2012) makes it all the more legitimate to examine such psychological factors as learner's stereotype threat and teacher's language attitude in studies on EFL learners' reticence. WTC is a dynamic and complex phenomenon in constant fluctuation. EFL learners may experience momentary change in WTC due to dwindling self-confidence or perceived negative reaction of their interlocutors (MacIntyre & Legatto, 2011, p. 150). For example, they may experience stereotype threat in the air from being self-conscious about being a quiet student with lower English speaking capabilities. Or they may not receive deserved attention or approval from teachers because of their strong first language accent. Hence, the two research questions to be proposed and tested in the current study are listed as follows.

RQ1: Do EFL students who suffer from stereotype threat or stigma consciousness (ability-stigmatized students) tend to remain quiet in CLT classes?

RQ2: Do teachers who hold unfavorable attitudes towards accented English speech produce more quiet students in CLT classes?

Method

Overview

To answer the above two research questions, two groups of participants – student participants and teacher participants, were recruited from a university in northern Taiwan. For student participants, two measures were taken: stereotype threat index (adapted from Aronson, Fried, & Good, 2002) and stigma consciousness level (adapted from Pinel, 1999). The teacher participants had to complete a semantic differential scale which included pleasure and arousal subscales (adapted from Cargile & Giles, 1997) to determine their language attitude. The current study is exploratory in nature in that it tests a relatively new way of approaching the persistent problem (Babbie, 2004) of students' reticence in English classrooms. This is to suggest that the study attempts to explain variation in EFL students' participatory communication efforts in CLT classes by looking at identified psychological processes experienced both by teachers and students. It is hoped that the method developed in the current study will be feasible for undertaking further related studies. The following sections detail the design of the study and measuring instruments used in the current study.

Instruments

Index of Stereotype Threat

To measure their perceptions of stereotype threat, student participants were asked to respond to two statements (adapted from Aronson, Fried, & Good, 2002). This measure was chosen because it was a more straightforward and simplified version than the 8-statement measure used in the first study on stereotype threat by Steele and Aronson (1995). In addition, the two-statement measure of stereotype threat index has been adopted in several other studies (Cohen & Garcia, 2005; Steele, Spencer, Davies, Harber, & Nisbett, 2001; Tomasetto, Alparone, & Cadinu, 2011). The two original items used to measure race-based stereotype threat were modified to fit the academic context and specifically to measure participants' perceived apprehension of participating in discussions in the English oral communication class. The student participants were required to indicate their degree of agreement on a

7-point Likert scale anchored with "strongly disagree (1)" and "strongly agree (7)," to the following two statements. "People make a judgment about my intellectual ability based on how often I talk in class" and "People make a judgment about quiet students based on how well they perform in class discussions." A composite index of stereotype threat was formed by averaging across the two items.

A median split was performed to categorize participants into high and low stereotype threat individuals (for all participants, $M = 3.59$, $SD = 1.39$, $Md = 3.5$ in a range of 1.0-7.0). This allowed for retention of the entire sample. In addition, Ward and Rana-Deuba (1999) argued that the categorical approach combined with the technique of median split is the most appropriate option when the purpose of the investigation is to make explicit or direct comparisons across categories (e.g., between the two categories of high and low stereotype in the current study). As a result, there were 106 low stereotype threat individuals ($M = 2.55$ in a range of $1.0 - 3.5$, $SD = .77$) and 86 high stereotype threat individuals ($M = 4.87$ in a range of $4.0 - 7.0$, $SD = .80$; no participants' average scores would fall between 3.5 and 4.0 since there were two items in the measure). The Cronbach's alpha of the scale is .617.

The coefficient alpha of .617 is moderate considering the conventionally acceptable reliability benchmark of .70 (Nunnally, 1978; Voss, Stem, & Fotopoulos, 2000). However, alpha value is closely related to the number of items in a scale (Cronbach, 1951; Cortina, 1993; Spiliotopoulou, 2009), and the effect is especially noticeable for scales with lower than seven items (McKennell, 1978; Swailes & McIntyre-Bhatty, 2002). Therefore, the mean inter-item correlation, which is independent of test length (Cronbach, 1951), was calculated. As suggested by Clark and Watson (1995) and Cox and Ferguson (1994), the average inter-item correlation falls within the range of .15 to .50. The calculated mean inter-item correlation of the stereotype threat index in the current study is .45.

Stigma Consciousness Questionnaire (SCQ)

The stigma consciousness questionnaire was a 10-item scale adapted from Pinel's study (1999) on behavioral consequence of stigma consciousness for women. Stigma consciousness as a construct was identified and operationalized for the first time by Pinel in 1999. In the study, Pinel developed and validated the stigma consciousness questionnaire to be a reliable and valid instrument for detecting individual differences in stigma consciousness. The SCQ has been widely used in stigma studies concerning gender, gay/lesbian, race (Lewis, Derlega, Clarke, & Kuang, 2006; Pinel, 2004; Pinel & Chua, 2000; Pinel, Warner, & Chua, 2005) and occupational stigma (Wildes, 2004). The modified ten

items in the current study were used to measure participants' level of consciousness concerning the stereotype of being a "quiet student in English class." Example questions are "Stereotypes about quiet students have not affected me personally," "When interacting with others, I feel like they interpret all my behaviors based on the fact that I am a quiet student in class," and "My being a quiet student in class does not influence how people (classmates, teachers) interact with me." Participants marked their degree of agreement to every item on a 7-point Likert scale, ranging from "strongly disagree (1)" to "strongly agree (7)." Seven of the 10 items were con-trait items to be reverse scored. Since the majority of the students (about 80%) scored on the low end of the scale (below the midpoint of the scale, $M = 3.23$, $SD = .80$, $Md = 3.2$ in a range of 1.1- 5.8), rather than grouping participants into high versus low stigma conscious individuals, the stigma consciousness score was analyzed as a continuous variable in the current study. Reliability analyses of the current scale yielded a Cronbach's alpha of .704, which is equal to the conventionally acceptable reliability value of .70 and close to the reported alpha of .74 in Pinel's (1999) original study to develop stigma consciousness questionnaire for women (p.116).

Materials – Stimulus Voices

The language attitude experiment is based on the research design in Cargile's and Giles' study (1997) to be detailed in the next section. For the accented English speech recording, a native Taiwanese female was selected for her clear voice quality. More importantly, she speaks English with a typical Taiwanese dialect feature. She was asked to read a short passage (around one minute) in English and her reading was audio tape recorded. The stimulus audio recording of an English speaker was extracted directly from the English Language Listening course material used at a Taiwanese local university. A direct use of the audio material was to ensure the maximum standard American accent manipulation. For the purpose of comparison, the extracted standard English recording was also a female voice. Contents of the passages included in both recordings were carefully selected so that they sounded fairly neutral (i.e., one was a lecture in Linguistics, the other a talk on traveling tips). With the traditional matched-guise technique (Lambert, Hodgson, Gardner & Fillenbaum, 1960), the most commonly used analysis method in studies on language attitudes; one passage was read by the same speaker used as different "guises." However, such a research design is not necessarily appropriate for all experiments (Hiraga, 2005). Same as in the studies conducted by Hiraga (2005) and Mugler (2002), the current study aimed to ensure accent authenticity (Taiwanese-accented English versus standard

American English), therefore two different speakers were used (a local Taiwanese and an American). Both recordings were transformed into digital files so that some teacher participants could complete the experiment online at their convenience.

Semantic Differential Scale

To measure teacher participants' language attitude, the current study adopts the semantic differential scale used in Cargile and Giles' (1997) experiment. The scale contained both pleasure and arousal subscales. Since attention to message contents was required to measure participants' arousal level, to lessen demand on participants, only the pleasure subscale was used. In most studies on language attitude, the commonly measured evaluative reactions to different accents or languages were listeners' perceptions of speakers' personality traits or socio-economic status (Cargile & Giles 1998; Hiraga, 2005; Lambert, Hodgson, Gardner & Fillenbaum, 1960; Mugler, 2002; Podberesky, Deluty & Feldstein, 1990). However, listeners' evaluation of speaker's personality or socio-economic status was not relevant to the current study. Rather, it is believed that listeners' feelings (i.e., affective reactions) about a heard accent were more valuable information in teacher-student interactions. Therefore, the current study focused on the perceived mood state of listeners (Cargile & Giles, 1997) or teachers in response to the listened audio-recordings with either Taiwanese-accented English or standard American English.

In Cargile and Giles' (1997) experiment, listeners were measured for their affective reactions or mood state when listening to either an audio-tape spoken with an American accent or with any variety of a Japanese accent. The Cargile and Giles' study showed that American listeners rated American-accented speakers with higher degrees of pleasure than they did with Japanese-accented speakers. More importantly, the strength of the Japanese accent was not related to reported levels of pleasure. In other words, American listeners in Cargile and Giles' (1997) experiment demonstrated less pleasurable emotions as long as the speech heard was spoken by an outgroup member rather than an ingroup member regardless of the strength of accent.

Finally, teacher participants were required to complete two pleasure subscales which contained four questions each to measure their mood state after listening to either an audio-tape of the American-accented speaker or the Taiwanese-accented speaker. Each of the four questions started with the statement -- "Please indicate the extent to which each word describes your feelings at this moment." The four affective responses were: "satisfied-unsatisfied, hopeful-despairing, annoyed-pleased, happy-

unhappy." The teacher participants who gave higher pleasure ratings to American-accented recording were considered as individuals with a negative attitude towards accented English or a negative language attitude. Those who greeted Taiwanese-accented recording with more pleasure were categorized as individuals with a positive language attitude. One of the items was reverse coded. The Cronbah's alpha is .918. As it turned out, among the six teacher participants, half of them were classified as having a negative language attitude and the other half a positive language attitude. Correspondingly, in the current sample of 192 student participants, 93 of them were taught by teachers with a positive language attitude and 99 by teachers with a negative language attitude.

The Study

 Participants. The targeted participants were the English major freshmen enrolled in an oral communication course at a university in Taiwan. A total of 201 students participated in the first part of the study. Nine participants were excluded from the final analysis due to missing data, failure to put down names or signatures on the short consent form, or recording identical responses throughout the entire questionnaire. As a result, a total of 192 participants made up the final study sample. Regarding teacher participants, all of the six American instructors teaching the freshman oral communication course participated in the current study.

 Design and Procedure. The experiment employed a 2 X 2 Analysis of Covariance (ANCOVA) design with student stereotype threat (low or high) and teacher language attitude toward accented English (positive or negative) as between-participant factors. The variable of student stigma consciousness was included as a covariate. The dependent variable is the student performance in the oral communication course, which in the current study is the end-of-semester daily grades. The end-of-semester daily grades better reflect the average in-class participation for each student. Different from the midterm and final exam grades, the daily grades were made up of a more multi-faceted and ongoing evaluation of students throughout the semester. The midterm and final exam grades at the studied college were decided respectively by one sit-down interview between the instructor and each individual student. The end-of-semester daily grades, on the other hand, were composed of student performances in class such as oral reports and overall oral participation in class (e.g., voicing one's own opinions on a discussed topic, responding to questions posed by instructors or other students).

In-class paper-and-pencil surveys were administered to student participants. Students were reminded to put down their names and signatures on the short consent form, the first page of the questionnaire. The purpose of recording the names was to identify each survey participant and his or her daily grade earned in the oral communication course. Students were informed of the purpose of the study with a cover story but encouraged to draw on their experiences in oral communication courses. Following the short consent form were instruments measuring participants' stereotype threat index and stigma consciousness level, with two questions for the former and 10 questions for the latter.

Regarding the language attitude experiment conducted on oral communication instructors, a total of six instructors completed the survey, either in-person administered by the researcher, or online through email correspondence with the researcher. Half of them did the survey online. The language attitude questionnaire also started with a short consent form which included brief instructions of the study. Besides being assured of their anonymity in any publications associated with the study, they were required to put down their names to indicate their consent, which in fact was again for identification purpose. In the two-part survey, they were asked to first listen to audio-tape #1 (a short passage read by an American-accented female) and then complete the semantic differential scale measuring mood states. Following that, they listened to audio-tape #2 (a short passage read by a Taiwanese-accented female) and then completed an identical semantic differential scale.

Results

Testing research questions

All 192 qualified respondents were included in the final analysis. A 2 (stereotype threat: high or low) X 2 (language attitude: positive or negative) analysis of covariate was conducted with stigma consciousness scores as covariate. The dependent variable was the overall student performance in the oral communication course (i.e., end-of-semester daily grade). The results revealed a main effect of stereotype threat, $F(1, 187) = 4.76$, $p < .03$, $\eta^2 = .03$, and a main effect of language attitude, $F(1, 187) = 20.99$, $p < .001$, $\eta^2 = .01$, and a significant interaction between stereotype threat and language attitude, $F(1, 187) = 4.18$, $p < .04.$, $\eta^2 = .02$. However, the covariate of stigma consciousness did not significantly predict the dependent variable, $F(1, 187) = .15$, $p < .69$. Neither a significant predictor of the dependent variable (no observed main effect between stigma consciousness and students' daily grades) nor a moderator between the two independent factors of stereotype threat and language attitude and the dependent variable, the variable of stigma consciousness was dropped from further analysis.

More analyses were conducted to examine the preliminary findings listed above. An independent factorial ANOVA was conducted on the independent variables of stereotype threat and language attitude and the dependent variable of oral communication course grade (i.e., end-of-semester daily scores). Results of the two-way independent ANOVA showed a significant main effect of stereotype threat, $F(1, 188) = 4.63$, $p < .03$, $\eta^2 = .03$, a significant main effect of language attitude, $F(1, 188) = 21.33$, $p < .001$, $\eta^2 = .10$, and a significant interaction between stereotype threat and language attitude, $F(1, 188) = 4.16$, $p < .04$, $\eta^2 = .02$.

The results indicated that high stereotype threat students and low stereotype threat students differed significantly in their mean end-of-term daily oral communication scores ($M = 79.10$, $SD = 14.80$ for the former; $M = 82.78$, $SD = 7.10$ for the latter). In addition, the mean daily oral communication scores of students taught by instructors with a positive language attitude ($M = 84.70$, $SD = 4.80$) differed significantly from the mean daily oral communication scores of students taught by instructors with a negative language attitude ($M = 77.79$, $SD = 14.33$). Further analysis on the significant interaction between the two independent factors of stereotype threat and language attitude revealed some interesting findings.
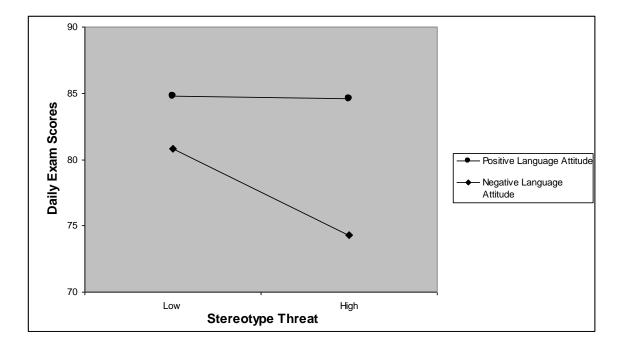
Pairwise comparisons (Winer, Brown, & Michels, 1991) were performed to examine differences in daily grades of the factor, *stereotype threat*, within the two levels of the factor, *language attitude*. A Bonferroni correction was applied to control for inflation of alpha. The comparisons (see Table 1) indicated that if taught by instructors with a negative attitude towards accented English, students low in stereotype threat performed significantly better in oral communication ($M = 80.79$) than students high in stereotype threat ($M = 74.33$), $F(1, 188) = 9.13$, $p < .003$, $\eta^2 = .05$. Students low in stereotype threat and taught by instructors with a positive attitude towards accented English performed moderately better in oral communication course ($M = 84.77$) than students low in stereotype threat but taught by instructors with negative language attitude ($M = 80.79$), $F(1, 188) = 3.72$, $p < .055$, $\eta^2 = .02$.

Table 1

Interaction of Student ST and Teacher LA on Student Performance in Oral Communication Class

	Student Stereotype Threat (ST)	
Teacher Language Attitude (LA)	High	Low
Positive	84.60	84.77
Negative	74.33	80.79

Also, students high in stereotype threat and taught by instructors with a positive language attitude performed significantly better in the oral communication course ($M = 84.60$) than students high in stereotype threat and taught by instructors with a negative language attitude ($M = 74.33$), $F(1, 188) = 20.02$, $p <. 001$, $\eta^2 = .11$. However, if taught by instructors with a positive language attitude, students high in stereotype threat and students low in stereotype threat did not differ significantly in their performances in oral communication classes ($M = 84.60$ for the former; $M = 84.77$ for the latter), $F(1, 188) = .006$, $p < .93$ (see Figure 1).

Figure 1. Estimated marginal mean oral communication daily grades as a function of student stereotype threat and teacher language attitude

Discussion

The current study looks into the phenomenon of student resistance to speaking up in ELF classrooms, a repeatedly accused perpetrator of the unsatisfactory result of CLT practice in Asia identified by various scholars throughout the last decade. Rather than totally endorsing the idea that students in Asia are culturally and intrinsically quiet, the current exploratory study identifies some subtle psychological processes either students or teachers go through during their communicative interactions. The state of mind of teachers and students, the CLT enactors, might help provide a new look into why some

students are more participatory and others more withdrawn while holding the cultural factor steady. The three psychological mechanisms in focus are stereotype threat and stigma consciousness on the part of students, and language attitude on the part of instructors.

The data analysis results revealed that in general, students low in stereotype threat performed significantly better in oral communication classes than those high in stereotype threat; and students taught by instructors with a positive language attitude outperformed those taught by instructors with a negative language attitude. However, students' level of stigma consciousness did not turn out to be a predictor of student performance in oral communication classes, defined operationally as end-of-semester daily grades. As mentioned earlier, end-of-semester daily grades in the current study reflected students' grades in a more holistic way to include such activities as in-class reports and in-class participation throughout the entire semester. In fact, the average level of stigma consciousness for this sample of participants was quite low (M = 3.23), with more than 80% of the participants scoring below the scale midpoint.

According to the definition of stigma consciousness, one is high in stigma consciousness when one is chronically aware of one's stereotyped status and believes the stereotype to pervade all aspects of one's social experiences (Pinel, 1999). Apparently, in the current study, a great number of the participants reported suffering from stereotype threat, a situational threat which surfaces when one's performance in a given situation may induce a negative stereotype about oneself or one's social group. For example, in an oral communication class, one may fear being negatively perceived as a quiet student of inferior English capability or belonging to such a group. However, the participants did not seem to have internalized the stereotyped status they might have experienced in oral communication classes. Nor did they allow the stigmatized thought to spill over into other aspects of their school experiences. It is speculated that college curriculum and campus life might provide multiple possibilities for students to create dynamic self-images. A student may self-perceive or be seen as "on the margin" in oral communication classes. However, this same student may be the "center of attention" in other classes, musical and artistic contests, or sports tournaments. Consequently, college students are less likely to see themselves as being reduced to one single stereotype across all situations.

The above explanation is plausible considering related research on learner self-beliefs by Sarah Mercer (2011). According to Mercer (2011), learner self-beliefs as a complex dynamic system should be better understood from a complexity perspective approach. A learner's self-concept is composed of multiple layers of self-beliefs in specific domains. Some are more broadly defined domains such as

FLL whereas others are more tightly defined domains such as English reading ability (p. 336). Therefore, it can be inferred that one learner may hold a more positive self-concept in one particular domain but not in the others. However, it is also possible that EFL learners who do not judge themselves positively in the domain of English speaking ability might do so in other domains of English learning.

In addition, of much relevance to the findings of the current study is Mercer's proposition on the dynamic nature of self-concept. A self system is not only holistic, complex, but is also in constant change depending on contextual changes in the environment. Therefore, it is possible that a quiet EFL student with a high level of stereotype threat in one class may venture to speak out in another due to a changing element in the context, a more encouraging teacher with a more positive language attitude. This idea of a multidimensional and dynamic self-concept along with a significant finding in the current study bode well for CLT in EFL classes in Asia.

The more meaningful finding in this study lies in the interaction of the two independent factors, student stereotype threat and teacher language attitude. The interaction shows an interesting pattern which requires a deeper examination of teacher-student dynamics in a CLT class. The results indicated that students low in stereotype threat performed better than students high in stereotype threat. However, this superiority in performance did not hold if students, both high and low in stereotype threat, were taught by instructors with a positive language attitude. This finding points out the importance of instructors' role and their attitude towards language and students' speech styles in the process of interactive teaching. Unarguably, the optimal situation is to eliminate students' stereotype threat and have them taught by instructors with a positive language attitude as demonstrated in the current study.

In every troubled CLT classroom, there might be quiet students suffering from stereotype threat and frustrated instructors who unknowingly hold a negative language attitude. These psychological barriers are self-generated and potentially intervenable. For instance, Byrnes, Kiger, and Manning (1996, 1997) studied teachers' attitudes towards English language learners (ELL) in the mainstream classes in the States. They found that the participation in carefully organized and formal training as one of the major reasons to bring about the most positive teacher attitudes. In spite of the importance of professional development training, many ELL teachers did not receive any pre-service training on working with ELLs (Walker, Shafer, & Iiams, 2004). Also, in literature on CLT in Asian EFL classrooms, a distinct lack of related training among teachers has been cited as one of the obstacles for successful implementation of CLT in Asia (Howard & Millar, 2008; Jeon, 2009; Li, 1998).

Training is necessary to modify teachers' attitudes and introduce teachers to strategies in second language teaching (Byrnes, Kiger, & Manning, 1997). The following suggested training is adopted from the development training for ELL teachers in general by Walker and colleagues (2004) and the proposed intervention for improving CLT teaching in the Japanese EFL context (King, 2011).

First, cultural training should be provided to teachers to help them identify their own cultural biases and negative attitudes. This is fairly important in Asian EFL context because increasingly, local English teachers in Asia are foreign trained (Howard & Millar, 2008). In addition, native English speaking teachers are found to teach in oral communication classes (Sakui, 2004) with a smaller number of students as those in the sampled university in the current study.

Second, pre-service teacher training should be provided, and attention should be paid to design context-specific training programs. This is important for CLT pedagogy in Asia. Many English teachers in Asia reported not knowing how to conduct CLT appropriately, and the specific social and cultural constraints in individual countries also pose barriers to successful implementation of CLT (Howard & Millar, 2008; Jeon, 2009; Li, 1998).

Third, teachers should pay close attention to the quantity and quality of their own talks in class to ensure maximum comprehension of input in learners.

Forth, after all, English is studied as a second language for ELL learners and EFL learners. Overly difficult and complex materials might tune EFL learners out.

Fifth, teacher silence might be used as a useful strategy to counteract student reticence (King, 2011). Longer teacher waiting time might ensure more and better responses since L2 learners need time to search for the right vocabulary and to form appropriate responses (Shrum, 1985; Tobin, 1987, as cited in King 2011).

Sixth, teachers should present themselves as approachable and supportive persons and create a collaborative learning environment so as to increase students' willingness to try out the target language, and promote more interactions between teachers and students and students and students.

Finally, mere awareness of an inclination to disfavor accented English serves as the first step in improving EFL teachers' attitudes toward diverse linguistic and paralinguistic variations in students' speech styles.

Limitations and implications for future research

First, only native English speaking teachers were included in the current study. The student participants were recruited from intact classes taught by these six American English oral communication teachers. For future related studies in the context of Taiwan, to reflect the real situation (i.e., most EFL teachers are local nonnative English teachers, Liu, 2005) and take into account that local Taiwanese teachers are likely to hold different attitudes towards English accents than native English speaking teachers, it is viable to include local Taiwanese EFL teachers. Second, with an inclusion of non-native English speaking teachers, Taiwanese bilingual speakers with different degrees of accent might be used in stimuli recording. The rationale is local teachers might be more sensitive to a diversity of accent although in Cargile and Giles' (1997) study, it was the existence of an accent rather than the strength of accent in the stimuli recordings that influenced the emotional reactions of American listeners.

Conclusion

The current study has started off with a perplexing and persistent personal observation of having two distinct groups of students in an oral communication class. Almost like having an assigned role, each and every student faithfully played his or her own part of either speaking up actively or remaining quiet regardless of any topics under discussion or activities in action. Although there is abundant EFL literature, which has examined this significant issue and identified several reasons for student reticence in class (Chang & Goswami, 2011; Howard & Miller, 2008; Hu, 2010; Jeon, 2009; Li, 1998; Sakui, 2004; Takako, 2012), this researcher has opted for looking beyond the external factors such as institutional constraints, curriculum focus, and educational values and cultural practice so as to address the internal mindset and affective experience of teachers and students, the two fundamental entities in a CLT class.

With repeated enactment of a chosen role (either being an outspoken or quiet student), students know perfectly well to which group they belong. That is, every student knowingly belongs to his or her self-perceived and other-perceived category. Hence, a stereotype has been formed in the mind of self and in the eyes of other interactants (students and teachers) in a communication class. What is significant in such a scenario of an enacted role and a created membership stereotype is that this stereotype is supposedly negative, with quiet students normally considered as under achievers in a language class, which is further reinforced by their lower grades. It is believed that quiet students shall

experience fear each and every time they are situated in a condition where their negative stereotype is likely to be revoked, a psychological threat identified as stereotype threat (Steel, 1997; Steel & Aronson, 1995). Repeated experience of stereotype threat might bring about internalized and constant awareness of one's stigma self to form another psychological response, stigma consciousness (Pinel, 1999). A logical assumption is that stereotyped or stigmatized students would become more reserved if taught by teachers with a negative language attitude who might express negative affect when listening to accented English spoken by students with a lower English capability.

Although the sampled students have reportedly suffered from stereotype threat but have not internalized the threat to become a constant stigma, the felt stereotype threat has significantly and negatively impacted the students on their performance in oral communication classes (e.g., receiving lower grades from being non-participatory in class discussions). However, students high in stereotype threat may find themselves scaffolded by teachers with a positive language attitude who are receptive of accented English and various speech styles. Rather than assuming that silence is a norm in Asian EFL classes which does not allow for much intervention, the finding in the current study is believed to be welcomed by most EFL teachers who are devoted to helping second language learners navigating through the long winding road of English learning. The current exploratory study can contribute to the field of CLT and EFL in proposing a different way of examining the prolonged problem of EFL student reticence. Ultimately, a carefully orchestrated communicative teaching program will work wonders only when the conductors and players are psychologically ready to engage themselves in the undertaking.

References

Akomolafe, S. (2013). The invisible minority: Revisiting the debate on foreign-accented speakers and upper mobility in the workplace. *Journal of Cultural Diversity*, *20*, 7-14.

Anisfeld, M., Bogo, N., & Lambert, W. E. (1962). Evaluational reactions to accented English speech. *Journal of Abnormal and Social Psychology*, *65*, 223-231.

Aronson, J., Lustina, M. J., Good, C., Keough K., Steele, C. M., & Brown, J. (1999). When White men can't do math: Necessary and sufficient factors in stereotype Threat. *Journal of Experimental Social Psychology*, *35*, 29-46.

Aronson, J., Quinn, D., & Spencer, S. (1998). Stereotype threat and the academic underperformance of minorities and women. In J. Swim and C. Stangor (Eds.), *Prejudice: The target's perspective*. Academic Press.

Aronson, J., Fried, C. B., & Good, C. (2002). Reducing the effects of stereotype threat on African America college students by shaping theories of intelligence. *Journal of Experimental Social Psychology, 38*, 113-125.

Babbie, E. (2004). *The practice of social research* (10th ed.). Wadworth/Thomson Learning: Belmont, CA.

Bourhis, R. Y., & Giles, HO. (1976). The language of cooperation in Wales: A field study. *Language Sciences*, *42*, 13-16.

Brown, R. P., & Pinel, E. C. (2003). Stigma on my mind: Individual differences in the experience of stereotype threat. *Journal of Experimental Social Psychology, 39,* 626-633.

Byrnes, D., Kiger, G., & Manning, L. (1996). Social psychological correlates of teachers' language attitudes. *Journal of Applied Social Psychology, 26*(5), 455-467.

Byrnes, D. A., Kiger, G., & Manning, M. L. (1997). Teachers' attitudes about language diversity. *Teaching and Teacher Education*, *13*, 637-644.

Cargile, A, C, & Giles, H. (1997). Understanding language attitudes: Exploring listener affect and identity. *Language and Communication*, *17*, 195-217.

Cargile, A. C. & Giles, H. (1998). Language attitudes toward varieties of English: An American-Japanese context. *Journal of Applied Communication Research*, *26*, 338-356

Cargile, A. C, Giles, H., Ryan, E. & Bradac, J. (1994). Language attitudes as a social process: A conceptual model and new directions. *Language and Communication*, *14*, 211-235.

Chang, M. & Goswami, J. S. (2011). Factors affecting the implementation of communicative language teaching in Taiwanese college English classes. *English Language Teaching*, *4*, 2011.

Chang, C. & Su Y. (2010). Educational reform in Taiwan: Beliefs about EFL teaching and learning. *The International Journal of Learning*, *17*, 265-278.

Chateignier, C., Dutrevis, M., Nugier, A., & Chekroun, P. (2009). French-Arab students and verbal intellectual performance: Do they really suffer from a negative intellectual stereotype? *European Journal of Psychology of Education, 17*, 219-234.

Clark, L. A. & Watson, D. (1995). Constructing validity: Basic issues in objective scale development. *Psychological Assessment, 7*, 309-319.

Clement, R., Baker, S., C., & MacIntyre, P. D. (2003). Willingness to communicate in a second language: The effect of context, norms, and vitality. *Journal of Language and Social Psychology*, *22*, 190-209.

Cohen, G. L., Garcia, J. (2005). "I am us" : Negative stereotypes as collective threats. *Journal of Personality and Social Psychology*, *89*, 566-582.

Cox, T. & Ferguson, E. (1994). Measurement of the subjective work environment. *Work and Stress, 8*, 98-109.

Cortina, J. M. (1993). What is coefficient alpha? An examination of theory and applications. *Journal of Applied Psychology, 78*, 98-104.

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, *16*, 297-334.

Dooly, M. (2005). How aware are they? Research into teachers' attitudes about linguistic diversity. *Language Awareness*, *14*, 97-111.

Fang, Y. (2011). A brief comment on communicative language teaching. *Journal of Language Teaching and Research*, *2*, 428-431.

Good, C., Aronson, J., & Inzlicht, M. (2003). Improving adolescents' standardized test performance: An intervention to reduce the effects of stereotype threat. *Applied Development Psychology*, *24*, 645-662.

Hato, Y. (2005). Problems in top-down goal setting in second language education: A case study of the action plan to cultivate Japanese with English abilities. *JALT Journal*, *27*, 33-52.

Heip, P. H. (2007). Communicative language teaching: Unity within diversity. *ELT Journal, 61*, 193-208.

Hiraga, Y. (2005). British attitudes towards six varieties of English in the USA and Britain. *World Englishes*, *24*, 289-308.

Howard, J. Millar, S. (2008). Ongoing challenges with language curriculum innovations in Asia: A South Korean case study. *Pacific Asian Education*, *20*, 59-75.

Horwitz, E. K., Horwitz, M.B., & Cope, J. (1986). Foreign language classroom anxiety. *The Modern Language Journal*, *70*, 125-132.

Hu, W. (2010). Communicative language teaching in the Chinese environment. *US-China Education Review*, *7*, 78-82.

Izumi, S. (2003). Comprehension and production processes in second language learning: In search of the psycholinguistic rationale of the output hypothesis. *Applied Linguistics*, *24*, 168-196.

Jeon, J. (2009). Key issues in applying the communicative approach in Korea: Follow up after 12 years of implementation. *English Teaching*, *64*, 123-150.

Kang, S. J. (2005). Dynamic emergence of situational willingness to communicate in a second language. *System*, *33*, 277-292.

King, J. E. (2011). *Silence in the second language classroom*. Unpublished doctoral dissertation, the University of Nottingham.

Kristiansen, T. & Giles, H. (1992). Compliance-gaining as a function of speech: Public requests in varieties of Danish. *International Journal of Applied Linguistics*, *2*, 17-35.

Lambert, W. E., Hodgson, R. C., Cardner, R. C. & Fillenbaum S. (1960). Evaluational reactions to spoken languages. *Journal of Abnormal and Social Psychology*, *60*, 44-61.

Lawrence, J. S., & Crocker, J. (2009). Academic contingency of self-worth impair positively- and negatively-stereotyped students' performance in performance-goal settings. *Journal of Research in Personality*, *43*, 868-874.

Lee, W., & Ng, S. (2010). Reducing student reticence through teacher interaction strategy. *ELT Journal*, *64*, 302-313.

Lewis, R. J., Derlega, V. J., Clarke, E. G., & Kuang, J. C. (2006). Stigma consciousness, social constraints, and lesbian well-being. *Journal of Counseling Psychology*, *53*, 48-56.

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

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

Long, M. (1996). The role of linguistic environment in second language acquisition. In W. Ritchie & T. Bhatia (Eds.), *Handbook of second language acquisition* (pp. *413-468)*. San Diego: Academic Press. pp. 413–468.

MacIntyre, P. D. (1994). Variables underlying willingness to communicate: A causal analysis. *Communication Research Reports*, *11*, 135-142.

MacIntyre, P. D., Babin, P.A., & Clement, R. (1999). Willingness to communicate: Antecedents and consequences. *Communication Quarterly*, *47*, 215-229.

MacIntyre, P. D., & Charos, C. (1996). Personality, attitudes, and affect as predictors of second language communication. *Journal of Language of Social Psychology*, *15*, 3-26.

MacIntyre, P. D. Clement, R., Dörnyei, A., & Noels, K. A. (1998). Conceptualizing willingness to communicate in a L2: A situational model of L2 confidence and affiliation. *The Modern Language Journal*, *82*, 545-562.

MacIntyre, P. D., & Legatto, J. J. (2011). A dynamic system approach to willingness to communicate: Developing an idiodynamic method to capture rapidly changing affect. *Applied Linguistics*, *32,* 149-171.

Mackey, A. (2002). Beyond production: Learners' perceptions about interactional processes. *International Journal of Educational Research*, *37*, 379-394.

Major, B., & Schmader, T. (1998). Coping with stigma through psychological disengagement. In J. Swim & C. Stangor (Eds.), *Prejudice: The target's perspective* (pp. 219-241). New York: Academic.

McCroskey, J. C., & Baer, J. E. (1985). Willingness to communicate: The construct and its measurement. Paper presented at the annual convention of the Speech Communication Association, Denver, CO.

McCroskey, J. C., & Richmond, V. P. (1987). Willingness to communicate. In J. C. McCroskey &J. A. Dally (Eds.), *Personality and interpersonal communication* (pp. 129-156). Newbury Park, CA: Sage.

McKennell, A. (1978). Attitude measurement" use of coefficient alpha with cluster and factor analysis. In J. Brynner & K. M. Stribley (Eds.), *Social research: Principles and procedures*. Milton Keynes: Open University Press.

Mercer, S. (2011). Language learner self-concept: Complexity, continuity, and change. *System*, *39*, 335-346.

Mugler, F. (2002). Attitudes to accents in English: A Pacific study. *Te Reo*, *45*, 65-89.

Nishino, T. (2008). Japanese secondary school teachers' beliefs and practices regarding communicative language teaching: An exploratory survey. *JALT Journal*, *30*, 27-51.

Nunnally, J. C. (1978). *Pscyhometric theory* (2nd ed.). New York: McGraw-Hill.

Pinel, E. C. (1999). Stigma consciousness: The psychological legacy of social stereotypes. *Journal of Personality and Social Psychology, 76*, 114-128.

Pinel, E. C. (2002). Stigma consciousness in intergroup contexts: The power of conviction. *Journal of Experimental Social Psychology, 38*, 178-185.

Pinel, E. C. (2004). You're just saying that because I'm a woman: Stigma consciousness and attributes to discrimination. *Self and Identity*, *3*, 39-51.

Pinel, E. C., & Chua, P. P. (2000). Standing strong in a prejudice world. Unpublished manuscript.

Pinel, E. C., Warner, L. R., Chua, P. P. (2005). Getting there is only half the battle: Stigma consciousness and maintaining diversity in higher education. *Journal of Social Issues*, *61*, 481-506.

Podberesky, R., Deluty, R. H., & Feldstein, S. (1990). Evaluations of Spanish-and Oriental-accented English speakers. *Social Behavior and Personality*, *18*, 53-63.

Pringle, C. A. (1980). The effect of teachers' attitudes toward dialects on their expectations for students' academic competence. Unpublished doctoral dissertation, Kent State University.

Riasati, M. J. (2012). EFL learners' perceptions of factors influencing willingness to speak English in language classrooms: A qualitative study. *World Applied Sciences Journal, 17*, 1287-1297.

Rosenthal, H. E. S. & Crisp, R. J., (2006). Reducing stereotype threat by blurring intergroup boundaries. *Personality and Social Psychology Bulletin*, *32*, 501-511.

Rosenthal, H. E. S., Crisp, R. J., & Suen, M. W. (2007). Improving performance expectancies in stereotypic domains: Task relevance and the reduction of stereotype threat. *European Journal of Social Psychology*, *37*, 586-597.

Sakui, K. (2004). Wearing two pairs of shoes: Language teaching in Japan. *ELT Journal*, *58*, 55-163.

Savignon, S. J. (1991). Communicative language teaching: State of the art. *TESOL Quarterly*, *25*, 261-277.

Schmader, T., Major, B., Eccleston, C., & McCoy¸ S. (2001). Devaluing domains in response to threatening intergroup comparisons: Perceiving legitimacy and the status-value asymmetry. *Journal of Personality and Social Psychology*, *80*, 736-753.

Shrum, J. L. (1985). Wait-time and the use of target or native languages. *Foreign Language Annals*, *18*, 305-313.

Smith, J. L., Kausar, R., Holt-Lunstad, J. (2007). Stigma consciousness in the classroom: A comparison of Pakistani women's motivation and well-being in science and non-science fields of study. *Sex Roles, 57*, 3-13.

Spencer, S. J., Steele, C. M., & Quinn, D. M. (1999). Stereotype threat and women's math performance. *Journal of Experimental Social Psychology, 35*, 4-28.

Spiliotopoulou, G. (2009). Reliability reconsidered: Cronbach's alpha and pediatric assessment in occupational therapy. *Australian Occupational Therapy Journal, 56*, 150-155.

Steel C. M. (1997). A threat in the air: How stereotypes shape intellectual identity and performance. *American Psychologist*, *52*, 613-629.

Steele, C., M., & Aronson, J. (1995). Stereotype threat and the intellectual testbperformance of African Americans. *Journal of Personality and Social Psychology*, *69*, 797-811.

Steele, C.M., Spencer, S.J., Davies, P.G., Harber, K., & Nisbett, R. (2001). African American college achievement: A "wise" Intervention. Unpublished manuscript, Stanford University.

Swailes, S. & McIntyre-Bhatty, T. (2002). The "Belbin" team role inventory: Reinterpreting reliability estimates. *Journal of Managerial Psychology, 17*, 529-536.

Swain, M. (1985). Communicative competence: Some roles of comprehensible input and comprehensible output in its development. In S. Gass & C. Madden (Eds.), *Input in Second Language Acquisition*. Rowley, MA.: Newbury House.

Swain, M. (1992). The output hypothesis: Just speaking and writing aren't enough. *Canadian Modern Language Review*, *50*, 158-164.

Takako, N. (2012). Modeling teacher beliefs and practices in context: A multimethods approach. *The Modern Language Journal*, *96*, 380-399.

Tao, Y., Yu, J., & Jin, Z. (2011). Communicative approach in college English teaching. *Sino-US English Teaching*, *8*, 689-693.

Tobin, K. G. (1987). The role of wait-time in higher cognitive learning. *Review of Educational Research*, *57*, 69-95.

Tomasetto, C., Alparone, F. R., & Cadinu, M. (2011). Girls' math performance under stereotype threat: The moderating role of mothers' gender stereotypes. *Developmental Psychology*, *47*, 943-949.

Tsurutani, C, & Selvanathan, E. (2013). Influence of generational cohort and experience with non-native speakers on evaluation of speakers with foreign-accented speech. *International Journal of the Sociology of Language, 224*, 43-62.

Tucker, R. G. & Lambert, W. E. (1969). White and Negro listeners' reactions to various American-English dialects. *Social Forces*, *47*, 463-468.

Voss, K. E., Stem, D. E., & Fotopoulos, S. (2000). A comment on the relationship between coefficient alpha and scale characteristics. *Marketing Letters, 11*, 177-191.

Walker, A., Shafer, J., & Iiams, M. (2004). "Not in my classroom": Teachers attitudes toward English language learners in the mainstream classrooms. *NABE Journal of Research and Practice*, *2*, 130-160.

Ward, C. & Rana-Deuba, A. (1999). Acculturation and adaptation revisited. *Journal of Cross-Social Psychology*, *30*, 422-441.

Wildes, V. J. (2005). Stigma in food service work: How it affects restaurant servers' intention to stay in the business or recommend a job to another. *Tourism and Hospitality Research*, *5*, 213-233.

William, F., & Naremore, R. C. (1974). Language attitudes: An analysis of teacher difference. *Speech Monographs*, *41*, 391-396.

Williams, F., J. L. Whitehead, & Miller, L. (1972). Relations between language attitudes and teacher expectancy. *American Educational Research Journal*, *9*, 263-277.

Winer, B. J., Brown, D. R., & Michels, K. M. (1991). *Statistical principles in experimental design* (3rd ed.). New York: McGraw-Hill.

Wout, D., Danso, H., Jackson, J., & Spencer, S. (2008). The many faces of stereotype threat: Group- and self-threat. *Journal of Experimental Social Psychology, 44*, 792-799.

Yuasa, K. (2010). English textbooks in Japan and Korea. *Journal of Pan-Pacific Association of Applied Linguistics*, *14*, 147-158.

Zhang, Y., & Wang, J. (2012). The elaboration of cultivating learners' English communicative competence in China. *English Language Teaching*, *5*, 111-120.

The Effects of Video Dubbing Tasks on Reinforcing Oral Proficiency for Chinese Vocational College Students

Ms. Pin He and Dr. Sukhum Wasuntarasophit

Khon Kaen University, Thailand

Bio data

From 2011 to 2013, Pin He was a Master's student in English at the Faculty of Humanities and Social Sciences, Khon Kaen University. Her primary research interests include second language pedagogy, pronunciation teaching, and self-directed learning. hp-word@hotmail.com

Sukhum Wasuntarasophit, Ph.D., is an English lecturer at the Faculty of Humanities and Social Sciences, Khon Kaen University, Thailand. He has taught research report writing and translation for undergraduate students, and corpus linguistics for MA students. His interests include EAP/ESP vocabulary and corpus-based studies, as well as research in ELT. sukwas@kku.ac.th

Abstract

This study investigated the effects of video dubbing tasks on improving English oral proficiency. In this research, 34 Chinese EFL learners were assigned to complete a video dubbing task in four weeks. Speech samples were collected from two spontaneous speaking tests before and after the treatment. The participants' attitudes toward the task were derived from questionnaire and interview results. The findings show that 1) the participants' oral proficiency in comprehensibility, fluency, and accentedness improved; 2) the participants enjoyed the task and had a high perception of their progress and positive attitudes toward the task, though they met with problems and difficulties. We can thus conclude that video dubbing tasks are effective in improving oral proficiency.

Address correspondence:
Department of English, Faculty of Humanities and Social Sciences, Khon Kaen University, Khon Kaen, Thailand

Keywords: video dubbing tasks, effects, oral proficiency, Chinese students

Introduction

Video has been widely used as an instructional tool for language teaching in recent years (Vanderplank, 2010) as it provides authentic language input in simulated and communicative contexts (Danan, 2004; Neri, Cucchiarini, Stirs & Boves, 2002). Besides that, video is also invaluable in promoting language production (Vanderplank, 2010). Video dubbing as a task is a language teaching technique requiring learners to replace the existing soundtrack of a video clip with their own (Burston, 2005), combining meaningful language input and output practice. However, there have been only a few studies on this topic (Burston, 2005; Chiu, 2012; Danan, 2010; He & Wasuntarasophit, 2013; Yachi & Karimata, 2008). Among them, He & Wasuntarasophit (2013) investigated the effects of video dubbing tasks on improving sentence stress for Chinese vocational college students, and found that the overall effects are significant. However, there has been no study done to investigate the effects of this technique on improving oral proficiency. To fill the gap, this study investigated the effects of video dubbing tasks on reinforcing oral proficiency for Chinese vocational college students and measured the students' attitudes toward the technique.

Literature Review

Pronunciation teaching and learning theories

Dalton & Seidlhofer (1994) proposed three "implicit-explicit" ranges of pronunciation teaching procedures (i.e. exposure, exercise and explanation) to guide pronunciation teaching and training. The three ranges are introduced in this section.

Exposure is the basic element for successful language acquisition as it establishes a target model for learners (Neri et al., 2002). Learners are supposed to hear and learn target sound features while processing the language for communication purposes (Dalton & Seidlhofer, 1994). Meaningful and communicative spoken language exposure is vital in pronunciation training. To be meaningful, the input should be relevant to the learner's needs and stimulate intrinsic motivation (Neri et al., 2002). Comprehensible input hypothesis suggests that language acquisition takes place when input is comprehensible; comprehensibility requires that the content and form of input be slightly more difficult than the learner's existing language competence (Krashen, 1982). In addition, to meet different learning styles, input should be presented in various forms: written, aural, and audio-visual (Neri et al., 2002). Although crucial in language learning, exposure is not enough; exercise in language use is also necessary.

Exercise offers perception and production practice of particular pronunciation features so as to help learners deeply process input gained through exposure (Dalton & Seidlhofer, 1994). According to the Output Hypothesis, by producing output, learners can test their hypothesis on the second language (L2) and notice the weakness of their language performance (Swain, 2000), which in turn helps consolidate the learner's existing knowledge. Moreover, according to cognitive psychology, exercise helps to internalize L2 knowledge to become more automatic (Lightbown & Spada, 1993) and thereby enhance fluency (Neri et al., 2002). Repeated practice reinforces the transfer of controlled knowledge to become automatic (Schmidt, 2002). Although imitation and repetition practice has been criticized in recent years as it fails to transfer language knowledge to actual communicative language use (Jones, 1997), many scholars point out the indispensable role of imitation and repetition in pronunciation training (Dalton & Seidlhofer, 1994; Ding, 2007; Morley, 1991). In line with this, Takeuchi, Ikeda & Mizumoto (2012) found that repetitive and cognitive reading aloud in L2 increases brain activation. To go beyond "imitation drills", Jones (1997) suggests creating meaningful contexts for practice to facilitate pronunciation acquisition through meaningful associations.

In the process of exposure and exercise, some learners who are talented in language learning might be able to "pick up" some pronunciation features. However, there are still many learners who need their attention to be drawn to specific language features by way of explanation of how to pronounce the respective sounds (Dalton & Seidlhofer, 1994). According to cognitive psychology theories in L2 acquisition, nothing can be learnt without "noticing" (Schmidt, 2002). By noticing, Schmidt intimates a conscious effort devoted to the understanding of target language features. The importance of explanation is also emphasized in the teaching approach of "Focus on Form" which suggests drawing students' attention to language features that appear incidentally in input and output while the students' focus is on meaning and communication (Long, 1991, as cited in Sheen, 2002).

When designing pronunciation training techniques, factors such as the learner's age, aptitude, motivation, attitude, personal characteristics, exposure to the target language, etc. should also be considered (Celce-Murcia, Brinton & Goodwin, 1996). Although hard to control, some factors can be manipulated to achieve better results in designing pronunciation training techniques (Schmidt, 2002). According to "affective filter theory", a learner's state of mind is the "filter" that determines whether and how much the learner can learn from the input (Lightbown & Spada, 1993). Being tense, anxious, or bored might make the input unacceptable and useless for acquisition. Language acquisition occurs when learners are motivated to learn (Schmidt, 2002). Motivation is "a combination of the learner's

attitudes, desires, and willingness" to devote their effort to learn L2 (Richards & Schmidt, 2002, p. 343). It is a dynamic process which involves three phases: choice motivation, executive motivation, and retrospective motivation (Schmidt, 2002). A learner's attitudes, linguistic and coping confidence, and initial beliefs are the determinant factors to generate choice motivation. Once the learner starts an activity or task, executive motivation is triggered. This type of motivation is appraised by the extents of the novelty and attractiveness of the learning experience, the significance of goal or need, coping potential, etc. Aside from that, a learner's performance and self-evaluation, teachers' and others' feedback and comments also play significant roles in building the learner's retrospective motivation.

Video dubbing tasks

Video dubbing tasks are projects in which students substitute their own voice for the original soundtrack of a video (Burston, 2005). Using audio-visual media, this technique contains a large number of listening and watching activities, which meet the learning preferences of Chinese students at tertiary levels (Life, 2011). Describing how to manage the projects at the technique level, Burston (2005) provided implementation suggestions and discussed the benefits and potential challenges of this task. Yachi & Karimata (2008) proposed an online video dubbing system. Danan (2010) developed a translating and dubbing task for military students with the students rating the task highly. The above three studies discussed the pedagogical benefits of video dubbing tasks on general language learning. For enhancing oral proficiency, Chiu (2012) used film dubbing projects to improve intonation and found that most participants considered the project to be effective. He & Wasuntarasophit (2013) found that video dubbing tasks were effective in improving Chinese students' sentence stress in reading aloud and answering questions tasks, but not as much in picture-describing tasks. Thus, it can be concluded from previous studies that video dubbing tasks are valuable teaching techniques. According to Burston (2005) and Danan (2010), apart from the motivational value, this technique is considered suitable for all language proficiency levels and effective for all language skills' development.

More specifically, Chiu (2012) identified the pedagogical benefits of dubbing tasks on reinforcing intonation by surveying students' perceptions and opinions. However, no empirical evidence has been provided to identify the effectiveness of this technique on oral proficiency. The study is devoted to fill this research gap.

To make the learning process more complete and fit the current study context, the video dubbing task was adapted, based on pronunciation teaching and learning theories described previously.

According to the three "implicit-explicit" ranges of pronunciation teaching procedures (Dalton & Seidlhofer, 1994), video dubbing tasks in previous studies worked very well in regards to exposure and exercise; however, further explanation was lacking. Hence, a major adaptation in this study was the addition of explanation and other various awareness-raising activities. Reviewing previous studies on pronunciation (Aufderhaar, 2004; Dlaska & Kerkeler, 2013; Ingels, 2011; Lu, 2010; Tanner & Landon, 2009; Wang & Munro, 2004), the researchers included activities used to reinforce the "noticing" of target features. The learners compared their pre-recorded dub to the model, collected teachers' feedback on their performance, and marked the location of target features from the model input. All of these strategies were adopted for this study's video dubbing task.

Oral proficiency and sentence stress

Oral proficiency has received a lot of attention in pronunciation teaching and assessment in recent years (Derwing, Munro & Wiebe, 1998; Derwing, Munro & Thomson, 2008; Derwing, Thomson & Munro, 2006; Seferoglu, 2005; Tanner & Landon, 2009; and Yates, Zielinski & Pryor, 2011). Reviewing these studies, the goals of improved speech fluency, comprehensibility, and accentedness have been the most common focus. Comprehensibility is a listener's perception of an utterance (Derwing & Munro, 2005). Accentedness is defined as a "listener's perception of how a speaker's accent is different from that of the first language (L1) community" (Derwing & Munro, 2005, p. 385). Foreign accent affects communication as it reduces intelligibility (Derwing et al., 2006). Fluency is "fast, smooth and accurate performance" in speaking (Kormos & Denes, 2004, p. 161).

Oral proficiency is affected by a number of factors with sentence stress being an important one (Celce-Murcia et al., 1996). Stress marks the rich and important information in a spoken sentence, while unstressed marks can be predicted from context (Pennington & Richards, 1986). Sentence stress errors may cause misunderstanding and thereby affect comprehensibility (Kang, Rubin & Pickering, 2010). Sentence stress also plays an important role in speech fluency and accentedness (Kang, 2010; Kormos & Denes, 2004). However, Chinese EFL learners usually have problems with this pronunciation feature. The commonly seen ones include speaking English with no prominent stress (Deterding, 2006, 2010; Gao, 2005; Hide, 2002; Tian, 2010), misplacing sentence stress (Chun, 1982, as cited in Chen, Robb, Gilbert & Lerman, 2001), and failing to make a sufficient distinction of duration in stressed and unstressed syllables (Chen et al., 2001). Hence, the present study has hoped to

reinforce Chinese EFL learners' speech comprehensibility and fluency, while reducing accentedness by focusing on sentence stress through the aforementioned video dubbing task.

Research Questions

1. To what extent can the video dubbing task reinforce oral proficiency for Chinese vocational college students?

2. What are the participants' attitudes toward the video dubbing task on reinforcing oral proficiency?

Method

Participants

The participants enrolled in this study comprised 34 Chinese EFL learners (34 female; mean age: 18.48 years, age range: 17-21 years), all 1st year English major students from a Chinese vocational college and native speakers of Mandarin Chinese. These participants were chosen by a purposive sampling technique. Following are the reasons that they were selected. First, they were reported to have problems with sentence stress according to their teacher's observations. The participants' poor performance in the pre-test of this study confirmed this. Second, they were enrolled in the course English Speaking to which the video dubbing task was assigned as extracurricular work. Lastly, these students showed interest in the study and voluntarily agreed to do the video dubbing task after their teacher introduced it to them.

Since the experiment was conducted in the last month of the semester, all of the participants' English courses had been finished before the treatment. Considering the EFL environment in China, they had limited exposure to English out of class (Chen & Goh, 2011). Hence, the effect and validity of this study were relatively guaranteed by isolating it from any other input.

The treatment: video dubbing task

The video dubbing task in this study required students to work in pairs, substituting their own voices for those on the original soundtrack of a 26-minute English video. The video chosen was the 8th episode from the 3rd season of the American TV sitcom Friends which has been commonly used in EFL studies (Ahn, 2011; Al-Surmi, 2012; Washburn, 2001). To ensure the language input was comprehensible for the participants, RANGE software (Nation & Heatley, 2002) was used to analyze the words in the scripts. The results of the analysis showed that this episode covered nearly 95% of the

3,500 basic words in the word list of the New Senior High School English Curriculum Standard (New Senior High, 2011). According to Liu & Nation (1985), 95% coverage of known words is necessary for learners to guess unknown words correctly within a text. Furthermore, English scripts were provided to the students and the video was English captioned and Chinese subtitled, making up the language input (Danan, 2004).

The task was to be completed in four weeks as an after-class assignment. The students worked in pairs dubbing the selected video and were required to practice at least an hour per day during the four-week treatment. In the study of Danan (2010), students took 20 hours to translate and dub a 20-minute TV series episode. Hence, four weeks were perceived to be enough for the participants to dub the 26-minute video without translation. Adapted from the models described by Burston (2005) and Danan (2010), the video dubbing task in this study contained seven main steps: preparation, task assignment, pre-recording, teacher's feedback and instruction, annotation, practice and rehearsal, and final recording. More details of these steps are provided in Table 1.

Table 1

The procedures of the video dubbing task

Step	Activities
1	- Students (Ss) watch the video in class. - Teacher (T) gives an introduction to the background story of *Friends* and the main characters in the episode.
2	- T assigns the task. Ss get the video and scripts from T. - Ss receive a *Video dubbing task instruction sheet* from T, in which task requirements, scoring criteria and suggestions are provided.
3	Ss practice dubbing for pre-recording after class within 1 week.
4	- T gives a 30-minute instruction about sentence stress in class. The instruction is given in Chinese and English. - T gives feedback and suggestions on Ss' performance in the pre-recording. - T asks Ss to listen to their own pre-recordings and compare their performance to the original soundtrack for self-correction and self-improvement.
5	T assigns the annotation task and requires the annotated scripts to be submitted within 1 week.
6	Ss practice and rehearse dubbing the video.
7	Ss record their final versions of the video dub and send their work to T.

Data collection of speech samples

In order to measure the effectiveness of the video dubbing task on reinforcing comprehensibility, fluency, and accentedness, the participants' speech sample recordings were collected from a pre-test and a post-test. These two tests have been commonly used in collecting speech samples for testing oral proficiency (Derwing et al., 1998, 2006; Tanner & Landon, 2009; Underhill, 1987). The Answering Questions Test required the participants to make a one-minute speech by answering questions on the topic of friendship, which correlated with the content of the video episode of Friends about the relationships between six good friends. In the Story-Telling based on Pictures Test, the participants needed to make a speech for at least one minute first telling a story based on a sequence of four pictures, and then answering a question related to the pictures.

There were minor differences between the two tests though both of them were speaking tests. According to Skehan (1996), the difficulty of a test is affected by three factors: code complexity (lexical and syntactic difficulty), cognitive complexity (information processing and familiarity), and communicative demand (time pressure). Approved by three experts, the two tests of this study were parallel in "code complexity" and "communicative demand", whereas the cognitive complexity of the Story-Telling based on Pictures Test was higher than the Answering Questions Test. Not only was the latter test more relevant to the content of the video in the task, but also referred to the participants' opinions within their own contexts which they were more familiar with. In contrast, the Story-Telling based on Pictures Test required more information processing from the students' imaginations, and evaluated the participants' oral proficiency in a stranger context.

In the two tests, the instructions and the questions were written in both English and Chinese to ensure that their performance on the tests was not affected by their English reading ability (Payne & Whitney, 2002). In each test, the participants were given 15 minutes for preparation before they started speaking. The two types of speaking tests (i.e. Answering Questions Test and Story-Telling based on Pictures Test) remained the same in both the pre-test and the post-test in order to measure progress (Derwing et al., 2006; Munro & Derwing, 2008).

Rating the Collected Speeches

After the participants' speeches were recorded, their performances in the two speaking tests were evaluated by two native speakers (NSs). Both of them were American lecturers teaching at Khon Kaen University (KKU) and reported that they did not speak any local dialects of American English. Before

rating the data, the raters were instructed in how to employ the rating criteria. To ensure they understood how to rate, they listened to and evaluated three warm-up recordings for practice. The speeches were judged in terms of comprehensibility, fluency, and accentedness by two NSs using a six-point scale (0-5) based on their subjective impression. This kind of subjective rating method has been widely used in the studies of prosodic improvement in L2 acquisition (Aufderhaar, 2004; Derwing et al, 1998; Ducate & Lomicka, 2009; Hardison, 2005; Seferoğlu, 2005; Tanner & Landon, 2009). After the speeches were obtained, one of the NS marked all of the recordings, and another NS marked 15% of the data. The data were randomly selected from both the pre-test and post-test results. The inter-rater reliability Pearson coefficient (r) was .82, which shows that the agreement between the two raters was acceptable (Kazdin, 1982, as cited in Ingels, 2011).

Before the experiment, all of the tests had been piloted by two Chinese EFL learners, who were at a similar English level to the participants in this study. One of the two raters piloted the four scoring criteria to check their reliability. It was found that the criteria were suitable for scoring.

Data Collection of the Participants' Attitudes

After the post-test, a questionnaire and an interview were conducted to study the participants' attitudes toward the video dubbing task. Attitudes in this study refer to the opinions and feelings that the participants hold concerning treatment (Ducate & Lomicka, 2009). Based on the previous studies (Ducate & Lomicka, 2009; Harwood & McMahon, 1997; Tanner & Landon, 2009), the following topics related to attitude were investigated: level of effort devoted, perceived progress, level of attractiveness toward the task, difficulties encountered, and recommendations for improving the task.

The questionnaire (see Appendix) contained two parts: part 1 comprised 6 questions directing students to choose the most suitable corresponding answers; part 2 had 14 statements which the students responded to using a 5-point Likert scale ranging from 1 (strongly agree) to 5 (strongly disagree) (Tanner & Landon, 2009). The questionnaire investigated the students' level of effort devoted and level of attractiveness toward the task. The students also provided feedback on their perceived progress and difficulties encountered.

Semi-structured interviews were arranged after the questionnaire was conducted to acquire more in-depth information regarding the student's experience doing the video dubbing task. The questions (see Appendix) were developed based on Ducate & Lomicka (2009) and Tanner & Landon (2009), and

investigated the students' gains in the task, positive and negative aspects of the task, difficulties encountered, and recommendations for the future improvement of the task.

After these three research instruments were constructed, the content validity and wording of the questionnaire and interview were checked and approved by two experts from KKU. The Chinese versions were checked by an expert from China.

Results

Results of tests

To answer research Question 1 (*To what extent can the video dubbing task reinforce oral proficiency for Chinese vocational college students?*), both descriptive and inferential statistics were used to analyze the data from the study. Following Seliger & Shohamy (1995), a paired-sample t-Test was utilized to compare the means obtained from the test scores as the design of this study was a one-group pre-test + post-test design, and to measure the extent of the oral proficiency progress the participants made after doing the video dubbing task. The scores of comprehensibility, accentedness, and fluency in each of the Answering Questions Test and Story-Telling Based on Pictures Test range from 0-5. The combination of these three scores is the sum score (0-15) for oral proficiency. The statistical results of the two tests are shown in Table 2 and 3 respectively.

Table 2

Effects of the video dubbing task for the Answering Questions Test

Test		$\overline{x}$	S.D.	Sig.
Comprehensibility	Pre-test	2.79	0.94	.021
	Post-test	3.15	0.95	
Fluency	Pre-test	2.56	0.74	.023
	Post-test	2.85	0.65	
Accentedness	Pre-test	2.47	0.50	.037
	Post-test	2.74	0.66	
Sum	Pre-test	7.79	2.12	.005
	Post-test	8.76	2.04	

Table 2 shows positive effects of the video dubbing task in reinforcing oral proficiency. The mean scores of comprehensibility, fluency, and accentedness of the pre-test and the post-test of the Answering Questions Test showed a significant difference (p< .05). The p-values for the Answering Questions Test were: comprehensibility, .021; fluency, .023; accentedness, .037; sum, .005. These results indicated that the participants' spoken English had become smoother with less of a foreign accent and could be understood more easily in the post-test when compared to the pre-test. It implies that the video dubbing task was effective in improving the participants' comprehensibility, fluency, and accentedness in the Answering Questions Test.

Table 3

Effects of the video dubbing task for the Story-Telling based on Pictures Test

Test		$\overline{x}$	**S.D.**	**Sig.**
Comprehensibility	Pre-test	2.56	1.23	.006
	Post-test	3.09	0.75	
Fluency	Pre-test	2.12	1.00	.003
	Post-test	2.68	0.72	
Accentedness	Pre-test	2.35	0.69	.023
	Post-test	2.65	0.64	
Sum	Pre-test	6.97	2.68	.001
	Post-test	8.38	1.84	

Table 3 shows more positive effects of the video dubbing task in reinforcing oral proficiency. The mean scores of comprehensibility, fluency, and accentedness from the pre-test and the post-test of the Story-Telling Based on Pictures Test showed a significant difference (p< .05). The table shows that the p-values for the Story-Telling Based on Pictures Test are: comprehensibility (.006); fluency (.003); accentedness (.023). The sum p-value of these three features is .001. This implies that the effect of the treatment was significant on improving participants' speaking comprehensibility, fluency, and accentedness. The results have indicated that, after taking the video dubbing task, the participants' speech was less hesitant and less accented, and their speech could be understood with less difficulty for NSs.

In sum, the results from the tests revealed that the post-test scores of all areas were significantly higher than the pre-test scores at the .05 level. The video dubbing task was effective in improving the participants' comprehensibility, fluency, and accentedness. In both of the two tests, the participants made greater improvement in comprehensibility and fluency than that of accentedness. Besides this, the participants performed better in the Answering Questions Test than in the Story-Telling based on Pictures Test. The average scores of the pre-test and post-test of the Answering Questions Test were higher than those of the Story-Telling based on Pictures Test. Meanwhile, the sum p-value results of the Story-Telling based on Pictures Test was lower than that of the Answering Questions Test, which signifies that the participants made better progress in the Story-Telling based on Pictures Test.

Findings from the questionnaire and interview

To answer Research Question 2 (*What are the participants' attitudes toward the video dubbing task on reinforcing oral proficiency?*), 14 of the 34 participants were interviewed by random selection. The participants' responses in part 2 of the questionnaire are presented in Table 4. In the investigation, the participants indicated the amount of time and effort that they put into completing the video dubbing task and the level of perceived pronunciation progress gained. Beneficial aspects of the task, difficulties encountered, and recommendations for improving the task were also provided. The table shows that the average score ($\overline{x} = 3.82$) of the 14 items is at the higher end of the Likert scale. The findings from the questionnaire and interview showed that the participants' general attitudes toward the video dubbing task were positive.

Table 4

The participants' responses about attitudes toward the video dubbing task (N=34)

Items	$\overline{x}$	S.D.	Level
1. I did my **best** in the video dubbing task.	4.00	0.77	High
2. I could have put more time and **effort** into completing the task.	4.35	0.81	Very high
3. I have made progress with my English **pronunciation** by doing the task.	3.82	0.79	High
4. I have learned how to pronounce **sentence stress** by doing the task.	3.76	0.65	High
5. I found myself sounding more **like a native speaker** when speaking English.	3.09	0.83	Moderate
6. I found I could speak more **fluently** in English.	3.76	0.69	High
7. I have more **confidence** now when speaking English.	3.53	0.82	High
8. I **enjoyed** doing the task.	3.97	1.02	High
9. I like the **video** in the task.	3.88	1.00	High
10. I enjoyed **imitating** the speech of the characters in the video.	4.03	0.96	High
11. By **comparing** my own speech in the pre-recording with the original one in the video, I discovered I had pronunciation problems.	3.94	0.85	High
12. I found the **teacher's instruction** to be very helpful in completing the task.	3.76	0.78	High
13. After **marking** the location of sentence stress for the video scripts, I could put stress on sentences more correctly while practicing the dubbing.	3.82	0.93	High
14. I found it very helpful to improve my pronunciation by practicing with a **partner.**	3.79	.770	High
Average	3.82	.838	High

Level of effort devoted. The average amount of effort and time that the participants devoted in doing the task was not satisfying. Part 1 of the questionnaire reveals that 91.7% students used less than 20 hours, 50% used less than 10 hours of practice in the video dubbing task, which is far less than the required amount of time (30 hours). The high and very high agreement of statements 1 and 2 in Table 4 indicates that the participants had tried their best in terms of their own ability in dubbing, while they

felt that the amount of time and effort they put into doing the task was not enough. The reasons given by the participants during the interview included laziness, lack of self-control, and premature satisfaction with their performance after several rounds of practice.

Perceived progress. The participants had a high level of perceived improvement in their general pronunciation ability, sentence stress, fluency, and speaking confidence. Table 4 shows that the mean scores of perceived progress (i.e. statements 3, 4, 6, 7) were at a high level. In Part 1 of the questionnaire, 27 participants (79.4%) answered "yes" to the question "apart from sentence stress, have you made progress in other aspects of pronunciation by doing the dubbing task?" The mainly referred improved areas included intonation, pauses, liaison, speech rate, and word pronunciation. The findings from the interview also confirmed these results.

Apart from the gains in linguistic knowledge and skills, three (21.4%) students reported feelings of self-consciousness with their problems in pronunciation and an improvement of learning strategies as a result during the interview. An example comment:

> *Student-22:* "I used to think my pronunciation was good. I realized my problems after I watched the video. The intonation (in the video) was so different from mine. … I will watch more English videos in the future. "

Level of attractiveness toward the task. The participants liked the video and enjoyed the task. The activities in the task were evaluated to be beneficial for reinforcing sentence stress. The statements 8 to 14 in Table 4 show that the mean scores of each statement were at a high level.
Example comments from the interview are as follows:

> *Student-20:* "Imitating the English in the video was very funny and made me happy. I feel accomplished after I was able to speak English like a NS."
> *Student-21:* "I was not able to follow the speech rate. After the teacher pointed out our problems and assigned us to mark the location of sentence stress in the scripts, I felt much better doing the task."

Five (35.7%) participants showed an interest and willingness to do more video dubbing tasks in the future while discussing initiatives during the interview. Student-34: "We did gain a lot of benefits from the task, and hope the teacher can assign us more tasks like this one." The teacher confirmed this by saying, "I want to carry out more video dubbing tasks. My students like the task very much. They expressed their willingness and interest: 'teacher, please choose more English movies for us, we want to get more practice.'"

Difficulties encountered. The video dubbing tasks are perceived as boring if they are too long or difficult, or if the speech in the video is too fast. Question 6 in part 1 of the questionnaire required the participants to identify problems and difficulties they encountered when doing the video dubbing task. They could choose more than one option for this question. Results show that 28 participants (82.4%) selected the statement that "it was boring to do the dubbing practice and rehearsal", 17 (50%) felt that "the video was too long", 12 participants (35.3%) agreed that "I am not a good voice actor, so it is hard for me to emote while dubbing lines", 6 participants (17.6%) agreed that "it was hard to follow and imitate the speech in the video because the speech rate was too fast". Besides these five options, the participants added that "it was hard to catch the actors' rhythm and intonation", "I could not persist in the practice due to laziness." The participants expressed similar problems in the interview. Here are some example comments:

> *Student-28:* "Though the video was very interesting, after several rounds of practice, I became bored."
>
> *Student-34:* "I was very upset and annoyed when I started to read after the characters in the video because I was not able to follow their speech rate at all."

Recommendations for improving the task. Suggestions that the participants mentioned in the interview included: using a lower speech rate video (4 students: 28.6%); changing the video to be a shorter one with fewer lines (4 students: 28.6%); adding more people to each group (3 students: 21.4%); receiving more supervision and feedback from the teacher (2 students: 14.3%).

Discussion

The research results show that the video dubbing task was effective in reinforcing oral proficiency for Chinese vocational college students. The participants enjoyed doing the task and perceived that they made progress after doing the task, though they had some difficulties and problems with it. Possible reasons for the results are discussed and focused on the aspects of the effects and weaknesses of the video dubbing task.

Effects

The positive results of the video dubbing task are consistent with the findings of Danan (2010), Chiu (2012), and He & Wasuntarasophit (2013). They can be attributed to the three "implicit-explicit" ranges of pronunciation teaching theory: exposure, exercise, and explanation (Dalton & Seidlhofer,

1994), which comprise the major part of the task. The selected episode of Friends provided abundant meaningful and comprehensible spoken language, which helped to stimulate exposure to native speaking sounds and language acquisition for the students (Burston, 2005; Danan, 2010). With the purpose of speaking as close to the original speech as possible, the participants did plenty of imitation and repetition practice in contextualized scenarios, during which they had to pay attention to sentence stress and other pronunciation features of the characters' speech patterns (Burston, 2005). The pronunciation features were processed more deeply, which led to the improvement of speech comprehensibility and accent. With the time constraints of the video, the students were forced to produce their lines quickly without hesitation (Danan, 2010), fluency thus improved. Furthermore, in the video dubbing task, the students received explicit learning strategies from the teacher's instruction and other awareness-raising activities. With feedback from the teacher and peers and the students' self-correction, reflective practice was effective in promoting awareness and self-learning.

The high motivation value of the video dubbing task was another element that contributed to the effects of this technique. In the initial phase, the explicitly set goal of sentence stress and score incentives triggered the students' motivation to do the task (Wu & Wu, 2008). During the task, the novelty and entertainment value brought about by watching the funny video and imitating the lines, the cooperation with their partner, and the teacher's supervision all maintained the students' executive motivation. Pre-recording tasks and the teacher's feedback further strengthened retrospective motivation.

In addition, the video dubbing task offered a stress-free and friendly learning environment. First, the task fit the general character of Chinese students: introverted, quiet and shy (Liu & Jackson, 2008). Recording enabled the participants to feel relaxed and free to speak English behind the scenes compared to speaking in public and also offered them opportunities to perform constant self-assessment or peer-review (Barry, 2012). They could devote a lot of effort in dubbing and redubbing to achieve their best performance. Cooperating with peers also reduced the students' anxiety and increased their motivation (Danan, 2010; Wu & Wu, 2008). Lastly, the video dubbing tasks enabled learners to gain individual control of the learning process (Burston, 2005; Danan, 2010). Considering individual differences of pronunciation proficiency, the students enjoyed great freedom in controlling the amount of practice according to their needs. It was found in the interview that the goal of speaking English like a NS from dubbing the video brought a sense of achievement to the students. The increase

in self-efficacy enhanced their learning beliefs and benefited their future English learning to some extent (Ellis, 2008).

The results show that by focusing on sentence stress, the video dubbing task was effective in improving oral proficiency. As mentioned before, previous research (Dalton & Seidlhofer, 1994; Derwing & Munro, 2005; Derwing et al., 1998; Kang, 2010; Kormos & Denes, 2004) found that sentence stress plays an important role in speech comprehensibility, fluency, and accentedness. Apart from sentence stress, many participants in the study noticed progress in other pronunciation features like intonation and pronunciation, though these features were not explicitly taught during the instruction phase. These reinforced pronunciation features might have also contributed to the overall improvement in oral proficiency.

It was found that the participants made a greater improvement in speech comprehensibility and fluency than in accentedness. A similar result was found by Derwing et al. (2006) in which Slavic learners made significant progress in fluency but only a small improvement in accentedness. In the current study, the participants might have seen a more urgent need to speak fluently and clearly than to be less accented. It is also suggested that correct accentedness might take a longer time to develop than comprehensibility and fluency.

It is worth mentioning that the participants performed better in the Answering Questions Test than in the Story-Telling based on Pictures Test during both the pre-test and the post-test. This might have resulted from the fact that the cognitive complexity of the Answering Questions Test is less demanding than that of the Story-Telling based on Pictures Test (Skehan, 1996). However, compared to the Answering Questions Test, the participants made better progress in the Story-Telling based on Pictures Test, the topics of which were less related to them. This shows that the participants' improvement in oral proficiency enabled them to handle speaking contexts that they were not very familiar with.

Weaknesses

Despite the effects of the video dubbing task on reinforcing oral proficiency and the students' positive attitudes toward the task, there were negative aspects of the technique shown from the students' feedback. The possible reasons are analyzed as follows:

First, the participants felt bored of the repeated read-after practice. Stress and tired feelings were caused by the fast speech rate of the actors, the length of the video, and too many lines and roles to dub. Since the speech rate was not just "slightly" higher than the participants' existing level of English

proficiency, the language exposure in the video was not consistent with Krashen's "i+1" criteria (Schmidt, 2002). Dubbing for such a fast speech rate was too demanding for these participants. For the reading aloud activity, the materials exceeded learners' L2 ability, causing low brain activation (Takeuchi et al., 2012). The fast speech rate also reduced the learners' listening comprehension, which might not only prevent learners from receiving language in a meaningful way, but also frustrate and demotivate learners (Zhao, 1997).

Second, although the participants found the video and the task to be very interesting, after several rounds of practice, they became bored with the repetition and imitation exercises. This might be because after they were familiarized with the content of the video and scripts, the input exposure and output exercises lost their novelty and meaningfulness. The task thereby turned into mechanical practice like a "drill" commonly found in the audio-lingual approach (Rodgers & Richards, 1986), which could have generated their bored feelings.

In addition, educational psychology claims that students have an attention span of around 15 to 20 minutes during any learning activity (McGoey & DuPaul, 2000). However, the video in this study had a length of 26 minutes, which meant the entire time for dubbing practice took at least 26 minutes. The participants were forced to concentrate for too long beyond the period of appropriate focused attention (Rieber, 1990). According to Krashen's affective filter hypothesis, learners tend not to learn when negative emotion exists (Krashen, 1982). These negative feelings toward the video dubbing task might somewhat reduce its advantages.

Implications and Conclusion

Future improvements of the video dubbing task

In response to the participants' complaint about the fast speech rate, the first suggestion for improvement is to replace the current video with one of a lower speech rate. Alternatively, inspired from the research of Wasuntarasophit (1997), the researchers propose to present the soundtrack of the video at three different speech rates during different periods of the task: a slow rate for the beginning period, an intermediate rate after students become familiar with the lines, and a normal rate after they can master the scripts. Second, before assigning the video dubbing task, the preparation time should be extended to give students more time to familiarize themselves with the video and reduce their anxiety. The third recommendation is to add more people to one group; with fewer lines to dub, students can concentrate more on pronunciation and thus understand the role of the characters more and act out their

emotions better. Dubbing fewer lines could also enable students to put more focus on pronunciation features. Alternatively, setting individualized learning goals for students of different English proficiency levels within the same class (Wu & Wu, 2008) could be more effective. Another suggestion is to divide the 26-minute video into several sections, providing 2 to 3 weeks to complete each section in order to reduce students' fatigue and maintain their attention and interest in the practice. Finally, it is suggested to assign more pre-recordings for students to complete before the final recording. By doing so, students can not only be prevented from laziness, but also receive more feedback from the teacher on their specific problems.

Implications for language pedagogy

There are several pedagogical implications for the use of video dubbing tasks in pronunciation teaching and training. First, with target pronunciation features as explicitly set learning goals, video dubbing tasks could be used and integrated into a pronunciation instruction curriculum or be customized according to the students' typical problems and needs in their pronunciation development. Second, when selecting a video, teachers should take into account the learners' proficiency levels; with not only the content and linguistic difficulties, but also the speech rate being considered. For lower level students, videos which are around 10 minutes and have fewer lines are recommended. Third, individualized guidance, supervision, and encouragement to maintain executive motivation should be given to students who lack independent learning ability. Lastly, though the video dubbing task in the current study contained the teacher's explicit instruction on the focused pronunciation features, the participants perceived that they also made progress on the pronunciation features that were not taught. This implies that teachers for whom it is not convenient to teach pronunciation can assign video dubbing tasks as a self-learning project for students to do after class.

Directions for future studies

A few limitations of this study should be mentioned for those wishing to direct further studies. First, this study had a relatively small sample size (34 female students). Future studies should enroll more participants and particularly male learners to add validity and generalizability to the present results. Second, considering the students involved in this study had limited competence in English pronunciation, future studies should be conducted with students of different English levels to better find out the effects of video dubbing tasks on the learner's level of proficiency. In addition, further

research could also try to identify the influence of the learner's level of effort devoted, motivation, and attitudes on the effects of the technique. Since the video dubbing task in this study comprised a one-session short-term intensive training, another direction for future study is to explore the effects of a less intensive longer-term task with more sessions focused on reinforcing pronunciation ability. Finally, with regard to research methods, it is suggested to change the 1-5 Likert rating scale in the questionnaire to an even numbered scale, such as 0-5, or 1-4 to avoid participants giving a neutral response to the statements (Gambrell, Palmer, Codling & Mazzoni, 1996).

Conclusion

Despite the limitations mentioned above, the findings from this study have provided the first empirical evidence that video dubbing tasks can be effective in improving EFL learners' oral proficiency. Video dubbing tasks offer meaningful language input and output via interesting stimuli and include the challenging tasks of acting as voice actors, teacher's instruction and feedback, and engaging in other awareness-raising activities focused on the specific pronunciation feature of sentence stress. All of these motivated the students to do the task and improved their oral proficiency of comprehensibility, accentedness, and fluency. Referring to the imitation and repetition exercises in video dubbing tasks, on the one hand they were seen to help to improve pronunciation ability; on the other hand, they were seen to reduce the meaningfulness of the video and the task, and consequently students might feel bored and tired performing the drill-like practice.

Corresponding authors' contact details:

Ms. Pin He,

Email: hp-word@hotmail.com

Mailing address: the Faculty of Humanities and Social Sciences, Khon Kaen University

Tel.: +8615727774710

Affiliations: Student of M.A. in English at the Faculty of Humanities and Social Sciences, Khon Kaen University.

Sukhum Wasuntarasophit, Ph.D,

Email: sukwas@kku.ac.th

Mailing address: the Faculty of Humanities and Social Sciences, Khon Kaen University

Tel.: +66810474875

Affiliations: Lecturer of English at the Faculty of Humanities and Social Sciences, Khon Kaen University.

References

Ahn, K. (2011). Conceptualization of American English native speaker norms: A case study of an English language classroom in South Korea. *Asia Pacific Education Review, 12*(4), 691-702.

Al-Surmi, M. (2012). Review: Quaglio (2009). Television dialogue: The sitcom Friends vs. natural conversation. Philadelphia: John Benjamins. *Corpora, 7*(1), 117-120.

Aufderhaar, C. R. (2004). *The influence of using discourse analysis techniques on the filtered speech of authentic audio text to improve pronunciation.* Ph.D. Dissertation in Education, University of Cincinnati.

Burston, J. (2005). Video dubbing projects in the foreign language curriculum. *CALICO Journal, 23*(1), 77-92.

Barry, S. (2012). A video recording and viewing protocol for student group presentations: Assisting self-assessment through a Wiki environment. *Computers & Education, 59*(3), 855-860.

Celce-Murcia, M., Brinton, D. M., & Goodwin, J. M. (1996). *Teaching pronunciation: A reference for teachers of English to speakers of other languages.* New York: Cambridge University Press.

Chen, Y., Robb, M. P., Gilbert, H. R., & Lerman, J. W. (2001). A study of sentence stress production in Mandarin speakers of American English. *Acoustical Society of America, 109*(4), 1681-1690.

Chen, Z., & Goh, C. (2011). Teaching oral English in higher education: Challenges to EFL teachers. *Teaching in Higher Education, 16*(3), 333-345.

Chiu, Y. H. (2012). Can film dubbing projects facilitate EFL learners' acquisition of English pronunciation? *British Journal of Educational Technology, 43*(1), 24-27.

Dalton, C., & Seidlhofer, B. (1994). *Pronunciation [volume in the Scheme for Teacher Education].* Oxford: Oxford University Press.

Danan, M. (2004). Captioning and subtitling: Undervalued language learning strategies. *Meta, 49*(1), 67-77.

Danan, M. (2010). Dubbing projects for the language learner: A framework for integrating audiovisual translation into task-based instruction. *Computer Assisted Language Learning, 23*, 441–456.

Derwing, T. M., & Munro, M. J. (2005). Second language accent and pronunciation teaching: A research-based approach. *TESOL Quarterly, 39*(3), 379-397.

Derwing, T. M., Munro, M. J., & Wiebe, G. (1998). Evidence in favor of a broad framework for pronunciation instruction. *Language Learning, 48*(3), 393-410.

Derwing, T. M., Thomson, R. I., & Munro, M. J. (2006). English pronunciation and fluency development in Mandarin and Slavic speakers. *System, 34*(2), 183-193.

Derwing, T. M., Munro, M. J., & Thomson, R. I. (2008). A longitudinal study of ESL learners' fluency and comprehensibility development. *Applied Linguistics, 29*(3), 359-380.

Deterding, D. (2006). The pronunciation of English by speakers from China. *English World-Wide, 27*, 175–198.

Deterding, D. (2010). Norms for pronunciation in Southeast Asia. *World Englishes, 29*(3), 364-377.

Ding, Y. (2007). Text memorization and imitation: The practices of successful Chinese learners of English. *System, 35*(2), 271-280.

Dlaska, A., & Krekeler, C. (2013). The short-term effects of individual corrective feedback on L2 pronunciation. *System, 41*, 25-37.

Ducate, L., & Lomicka, L. (2009). Podcasting: An effective tool for honing language students' pronunciation? *Language Learning & Technology, 13*(3), 66-86.

Ellis, R. (2008). Learner beliefs and language learning. *Asian EFL Journal, 10*(4), 7-25.

Gambrell, L. B., Palmer, B. M., Codling, R. M., & Mazzoni, S. A. (1996). Assessing motivation to read. *The Reading Teacher, 49*(7), 518-533.

Gao, L. L. (2005). Pronunciation difficulties analysis: A case study. *CELEA Journal, 28*(2), 76-84.

Hardison, D. M. (2005). Contextualized computer-based L2 prosody training: Evaluating the effects of discourse context and video input. *CALICO Journal, 22*(2), 175-190.

Harwood, W. S., & McMahon, M. M. (1997). Effects of integrated video media on student achievement and attitudes in high school chemistry. *Journal of Research in Science Teaching, 34*(6), 617-631.

He, P., & Wasuntarasophit, S. (2013), Reinforcing sentence stress for Chinese EFL learners through video dubbing tasks, *Proceedings of the 9th International Conference on Humanities and Social Sciences (ICHUSO),* Khon Kaen University, 667-680.

Hide, O. (2002). Interlanguage phonology: Implications for a remedial pronunciation course for Chinese learners of English. *Antwerp Papers in Linguistics, 100*, 17-46.

Ingels, S. A. (2011). *The effects of self-monitoring strategy use on the pronunciation of learners of English.* Ph.D. Dissertation in Educational Psychology, University of Illinois at Urbana-Champaign.

Jones, R. (1997). Beyond "listen and repeat": Pronunciation teaching materials and theories of second language acquisition. *System, 25*, 103-112.

Kang, O. (2010). Relative salience of suprasegmental features on judgments of L2 comprehensibility and accentedness. *System, 38*(2), 301-315.

Kang, O., Rubin, D. O. N., & Pickering, L. (2010). Suprasegmental measures of accentedness and judgments of language learner proficiency in oral English. *The Modern Language Journal, 94*(4), 554-566.

Krashen, S. (1982). *Principles and practice in second language acquisition.* Pergamon: Oxford.

Kormos, J., & Dénes, M. (2004). Exploring measures and perceptions of fluency in the speech of second language learners. *System, 32*(2), 145-164.

Life, J. (2011). Motivation and EFL university students in north-east Asia. *Asian EFL Journal, 13*(3), 11-41.

Lightbown, P. M., & Spada, N. (1993). *How languages are learned.* Oxford: Oxford University Press.

Liu, M., & Jackson, J. (2008). An exploration of Chinese EFL learners' unwillingness to communicate and foreign language anxiety. *The Modern Language Journal, 92*(1), 71-86.

Liu, N., & Nation, I. S. P. (1985). Factors affecting guessing vocabulary in context. *RELC Journal, 16*(1), 33-42.

Lu, J. (2010). *CASTLE: A computer-assisted sentence stress teaching and learning environment.* Ph.D. Dissertation in Computer Science at Massey University, Manawatu, New Zealand.

McGoey, K. E., & DuPaul, G. J. (2000). Token reinforcement and response cost procedures: Reducing the disruptive behavior of preschool children with attention-deficit/hyperactivity disorder. *School Psychology Quarterly, 15*(3), 330-343.

Morley, J. (1991). The pronunciation component of teaching English to speakers of other languages. *TESOL Quarterly, 25*, 481–520.

Munro, M. J., & Derwing, T. M. (2008). Segmental acquisition in adult ESL learners: A longitudinal study of vowel production. *Language Learning, 58*(3), 479-502.

Nation, I. S. P., & Heatley, A. (2002). *Range: A program for the analysis of vocabulary in texts [Computer software].* Wellington, NZ: LALS, Victoria University of Wellington, New Zealand. Downloadable from http://www. victoria. ac. nz/lals/staff/paul-nation/nation. Aspx. (Retrieved 11 November 2012).

Neri, A., Cucchiarini, C., Strik, H., & Boves, L. (2002). The pedagogy-technology interface in computer assisted pronunciation training. *Computer Assisted Language Learning, 15*(5), 441-467.

New Senior High. (2011, n.d.). *New senior high school English curriculum standard.* Retrieved November 4, 2012, from http://www.pep.com.cn/ge/jszx/jtj/kb/kbx/

Payne, J. S., & Whitney, P. J. (2002). Developing L2 oral proficiency through synchronous CMC: Output, working memory, and interlanguage development. *CALICO Journal, 20*(1), 7-32.

Pennington, M. C., & Richards, J. C. (1986). Pronunciation revisited. *TESOL Quarterly, 20*(2), 207-225.

Richards, J.C., & Schmidt, R. (2002). *Longman dictionary of applied linguistics and language teaching.* 3rd ed. Harlow: Longman.

Rieber, L. P. (1990). Using computer animated graphics with science instruction with children. *Journal of Educational Psychology, 82*(1), 135-140.

Rodgers, T. S., & Richards, J. C. (1986). *Approaches and methods in language teaching.* Cambridge: Cambridge University Press.

Schmidt, N. (Ed.). (2002). *An introduction to applied linguistics.* New York: Hodder Headline Group.

Seferoğlu, G. (2005). Improving students' pronunciation through accent reduction software. *British Journal of Educational Technology, 36*(2), 303-316.

Seliger, H. W., & Shohamy, E. (1995). *Second language research methods.* Oxford: Oxford University Press.

Sheen, R. (2002). "Focus on form" and "focus on forms". *ELT Journal, 56*(3), 303-305.

Skehan, P. (1996). A framework for the implementation of task-based instruction. *Applied Linguistics, 17*(1), 38-62.

Swain, M. (2000). The output hypothesis and beyond: Mediating acquisition through collaborative dialogue. In J.P. Lantolf (Ed.), *Sociocultural theory and second language learning* (pp. 97-114). Oxford: Oxford University Press.

Takeuchi, O., Ikeda, M., & Mizumoto, A. (2012). Reading aloud activity in L2 and cerebral activation. *RELC Journal, 43*(2), 151-167.

Tanner, M. W., & Landon, M. M. (2009). The effects of computer-assisted pronunciation readings on ESL learners' use of pausing, stress, intonation, and overall comprehensibility. *Language Learning & Technology, 13*(3), 51-65.

Tian, Z. J. (2010). *Norm orientation of Chinese English: A sociohistorical perspective.* Ph.D. Dissertation in English, Free University of Berlin.

Underhill, N. (1987). *Testing spoken language: A handbook of oral testing techniques.* Cambridge University Press.

Vanderplank, R. (2010). Déjà vu? A decade of research on language laboratories, television and video in language learning. *Language Teaching, 43*(1), 1-37.

Wang, X., & Munro, M. J. (2004). Computer-based training for learning English vowel contrasts. *System, 32*(4), 539-552.

Wasuntarasophit, S. (1997). *An investigation into the use of training in pronunciation features and gradually increasing rates of delivery to help students improve their listening comprehension.* Master's Thesis in Applied Linguistics, King Mongkut's University of Technology Thonburi.

Washburn, G. N. (2001). Using situation comedies for pragmatic language teaching and learning. *TESOL Journal, 10*(4), 21–26.

Wu, W. C. V., & Wu, P. H. N. (2008). Creating an authentic EFL learning environment to enhance student motivation to study English. *Asian EFL Journal, 10*(4), 211-226.

Yachi, M., & Karimata, E. (2008). Online video dubbing system for language education. *Proceedings of World Conference on E-Learning in Corporate, Government, Healthcare, and Higher Education, 1*, 4029-4034.

Yates, L., Zielinski, B., & Pryor, E. (2011). The assessment of pronunciation and the new IELTS pronunciation scale. *IELTS Research Reports, 12*, 1-46.

Zhao, Y. (1997). The effects of listeners' control of speech rate on second language comprehension. *Applied Linguistics, 18*(1), 49-68.

APPENDIX   Questionnaire and Interview

English version:

Questionnaire

Name : _____

Student ID : _____

Instructions:

1. Please choose the response that best describes your attitudes toward the video dubbing task. You can choose only one option for each question except the ones specified with "(You can choose more than one option)". There are three parts to this questionnaire. Please answer all the questions.

2. The "native speakers" in this questionnaire refer to people who speak English as their mother tongue.

I. Please choose the answer(s) that best describes your situation.

1. Please calculate the amount of time you devoted in completing the video dubbing task with reference to your Journal Entries. (h: hours)

□ less than 5 h   □ 5-10 h   □ 10-15 h   □ 15-20 h   □ more than 20 h

2. Please calculate how many practice times you had in completing the video dubbing task with reference to your Journal Entries. (t: times)

□ ≤ 3 t   □ 4-6 t   □ 7-9 t   □ 10-12 t   □ 13-15 t   □ ≥16 t

3. Please rank the following types of practice according to the time and effort you devoted. Please specify the figure in the □, 1 refers to the most, 2-second; 3-third...

 (Please identify the type of practice not referred to below.)

 □ watch video to understand the plot and story of the video

 □ read script   □ practice dubbing by reading the scripts aloud

 □ repeat and imitate after the speech in the video

 □ rehearse dubbing with partner   □ _____

4. Apart from the sentence stress, have you made progress in other aspects of pronunciation by doing the task?

 □ No.

 □ Yes. (Please identify the improved aspects)_____

5. Apart from the sentence stress, have you found any other problems in your pronunciation by doing the task?

□ No.

□ Yes. (Please identify the problems)_____

6. Have you encountered any of the problems below while you were doing the task? (you can choose more than one option)

□ hard to record the voice   □ the video is too long

□ boring to do the dubbing practice and rehearsal

□ hard to follow and imitate the speech in the video because the speech is too fast

□ I am not a good voice actor, so it is hard for me to emote while dubbing

□ _____

II. Please read the sentences and check the answer which fits your opinion most.

5-strongly agree   4-agree   3-neutral   2-disagree   1-strongly disagree

	1	2	3	4	5
1. I did my best in the video dubbing task.					
2. I could have put more time and effort into completing the task.					
3. I have made progress with my English pronunciation by doing the task.					
4. I have learned how to pronounce sentence stress by doing the task.					
5. I found myself sounding more like a native speaker when speaking English.					
6. I found I could speak more fluently in English.					
7. I have more confidence now when speaking English.					
8. I enjoyed doing the task.					
9. I liked the video in the task.					
10. I enjoyed imitating the speech of the characters in the video.					
11. By comparing my own speech in the pre-recording with the original one in the video, I discovered I had pronunciation problems.					
12. I found the teacher's instruction to be very helpful in completing the task.					
13. After marking the location of sentence stress in the video scripts, I could put stress on sentences more correctly while practicing dubbing.					
14. I found it very helpful for improving my pronunciation by practicing with a partner.					

Key questions asked in the interview

English version:

Key questions

1. Please describe the process you went through completing the video dubbing task. Which activity or steps took you the most amount of time?

2. Have you encountered difficulties in doing the task? If yes, what were the difficulties?

3. What have you learned or gained from doing the task?

4. What are the positive and negative aspects of the task? (You can talk about the steps or activities in the task that you liked or disliked.)

5. Could you please give suggestions and recommendations for the future improvement of the task?

Towards integrating Computer-based Testing (CBT) into the EFL curriculum: Iranian EFL teachers' perspectives on challenges and obstacles

Reza Dashtestani

University of Tehran

Bio data

Reza Dashtestani is a doctoral candidate in Applied Linguistics at the University of Tehran. He is a member of the reviewers or editorial boards of several journals related to educational technology. Also, he is the associate editor of the TESOL International Journal. His areas of research include computer-assisted learning, online and distance education, digital literacies, and interactive learning environments. rdashtestani@ut.ac.ir

Abstract

Experts in language testing have shown interest in the use of computer-based tests in educational contexts (Brown, 1997; Chapelle & Douglas, 2006; Fulcher, 2003). However, limited research has been undertaken to explore EFL teachers' perceptions on the use of computer-based tests. To fill in the gap, this mixed-method study examined the perspectives of 247 Iranian EFL teachers on the implementation of computer-based testing (CBT) in Iranian EFL contexts. Three instruments were used for data collection. The quantitative data were collected using a questionnaire survey to investigate the EFL teachers' perspectives on the implementation of CBT. Follow-up in-depth interviews were conducted with 68 EFL teachers to provide supplementary and qualitative data. A total of 11 EFL syllabi were further analyzed to explore any use of computer-based tests in Iranian EFL courses. Findings indicated that the teachers held positive attitudes toward the implementation of CBT. Several challenges to the implementation of CBT in Iranian language teaching contexts were identified and discussed. The results showed that computer-based tests are not used and developed by

Address correspondence:
Faculty of Foreign Languages and Literatures, University of Tehran, Karegar-e-Shomali st. adjacent to Faculty of Physical Education, Tehran, Iran

Iranian EFL teachers. The participants proposed several strategies and measures which might facilitate the incorporation of CBT in language testing. The findings have implications for the incorporation of CBT in the EFL curriculum.

Introduction

Over recent years, educational authorities and researchers have become more and more interested in the use of computers in educational assessment and testing (Conole & Warburton, 2005; Sorana-Daniela & Lorentz, 2007). Similarly, in language teaching contexts, computer-based testing (CBT) has offered considerable benefits for EFL (English as a Foreign Language) teachers and students (Chapelle, 2007; Guiran & Shuli, 2010). Computer-based tests are suggested to be more beneficial than paper-based tests since CBT provides accurate scoring, increased security of the test, and ease of administration (Alderson, 2000; Brown, 2007; Wang & Shin, 2009). Specifically, the use of CBT in language testing can improve assessment in different regards. For instance, Brown (1997) points out that CBT would enhance the efficiency of language testing, facilitate the immediacy of feedback provision, and decrease the possibility of cheating.

Generally, CBT is defined as a test that is administered on test takers' personal computers or computers available at educational institutions (Wei, 2007). Chapelle and Douglas (2006) point out that "a computer adaptive test selects and presents items in a sequence based on the test taker's response to each item" (p.7), and it is widely used in educational measurement and testing. If a test taker chooses a correct response to a question, a more difficult item will be presented by the computer. In contrast, if an examinee provides a wrong response to a question, an easier item will be presented to the test taker.

The term CBT was used for the purposes of this study since the term is related to the use of computers in testing. CBT might be adaptive or non-adaptive. Despite an abundance of studies on students' perceptions of CBT both in EFL and other educational contexts (e.g., Adair, Jeager, & Pu, 2012; Al-Amri, 2008; Dashtestani & Sharifi, 2012; Salimi, Rashidy, Salimi, & Amini Farsani, 2011; Shi, 2012; Yurdabakan & Uzunkavak, 2012), very limited attention has been given to the issue of teachers' perceptions of and attitudes toward CBT. In general, the results of the studies conducted on students' attitudes toward CBT have revealed the positive attitudes of most students toward the use of CBT in EFL and other educational contexts.

Concerning the implementation of CBT in EFL contexts, Chapelle and Douglas (2006) assert that language teachers play an important role in the implementation of CBT in language testing contexts. They believe that language teachers should be accountable for preparing language learners to take computer-based tests. Language teachers are also supposed to have the competence to select, use, and develop appropriate and well-designed computer-based tests to meet the needs of their learners. Possibilities of response analysis, feedback, and record keeping are the merits that CBT can provide for language teachers. Likewise, it is essential that language teachers have the ability to evaluate computer-based tests if these tests are supposed to be incorporated into the EFL curriculum.

Literature review

Despite a plethora of studies on the issue of teachers' attitudes toward the use of technology and the implementation of computer-assisted learning (CAL) in educational settings, there is a dearth of research on teachers' attitudes toward the implementation of CBT. In general, the findings of research on teachers' attitudes toward the use of technology in education suggest that teachers perceive the use of technology as beneficial and effective and adopt positive attitudes (e.g., Albirini, 2006; Dashtestani, 2012; Ismail, Almekhlafi, & Almekhlafy, 2010; Kim, 2002; Lau & Sim, 2008; Park & Son, 2009; Simonsson, 2004). Previous research has also revealed that there are several challenges and obstacles to the use of technology in educational settings (e.g., Dashtestani, 2012; Kim, 2002; Park & Son, 2009; Simonsson, 2004). Lack of facilities, teachers' lack of knowledge of technology use, lack of support to use technology, and traditional approaches to teaching are the most considerable limitations of the use of technology in education.

Regarding technology use in the Iranian EFL context, Dashtestani (2012) examined Iranian EFL teachers' perspectives on the use of technology in EFL instruction. He reported that Iranian teachers held positive attitudes toward the use of computers in EFL instruction, and identified several limitations and challenges. One perceived benefit of the use of technology in EFL instruction was to improve the quality of assessment and testing through using computers in language testing. However, lack of computer-based facilities and teachers' low computer literacy levels were considerable challenges to technology use in EFL instruction of Iran. Park and Son (2009) analyzed the perceptions of EFL teachers on the integration of technology in language learning. They pointed out that their teachers believed that the use of technology improves the quality of language teaching and learning considerably. Teachers' limited knowledge of CALL and computers discouraged EFL teachers from

using technology in their teaching. The teachers further reported that access to authentic and sufficient teaching resources provided tremendous opportunities for them to teach more effectively. Similarly, in Syria, Albirini (2006) conducted a study on Syrian teachers' attitudes toward the use of technology in their instruction. He argued that even though the teachers perceived the use of technology as beneficial and effective in general, their attitudes can be affected by their computer literacy and cultural beliefs. Lau and Sim (2008) also conducted a study on attitudes of Malaysian teachers towards computer-assisted learning (CAL). They argued that the teachers had positive attitudes toward the use of technology and used technology for teaching. The teachers who had received instruction on the use of technology for educational purposes had more positive perceptions than those who had not received any instruction.

In terms of teachers' perceptions of CBT, Laborda and Litzler (2011) investigated challenges of training teachers for the implementation of CBT in language teaching contexts. They further assessed teachers' attitudes toward the implementation of CBT after receiving training concerning CBT and the use of technology in education. The teachers had limited knowledge of testing and assessment and the use of technology in language testing prior to the training. After the training, the teachers developed positive attitudes toward CBT since the training course improved their ability to use computer-based tests considerably. Similarly, Laborda and Royo (2009) assessed the effect of training on implementing CBT on EFL teachers' attitudes. A six-month course was held to improve the attitudes and competence of 26 language teachers to use and develop computer-based tests. The results of the questionnaire study confirmed that the majority of the language teachers had positive attitudes toward training on the use and designing of computer-based tests. The findings further demonstrated that despite teachers' positive reactions to CBT, the teachers were not able to use technology for language testing purposes. To conclude, Laborda and Royo (2009) argued that many language teachers might not be familiar with or prepared to develop and use computer-based tests. Furthermore, information and communications technology (ICT) competence required for the use of technology in language testing is different from the ICT competence required for the implementation of computer-assisted language learning (CALL) and teaching. Therefore, it is equally important to hold specific training courses to prepare language teachers to use computers for language testing.

Jamil, Tariq, and Shami (2012) examined the perceptions of university teachers on CBT versus paper-based testing (PBT). A total of 111 university teachers in Pakistan participated in the study. They reported that the majority of teachers adopted positive attitudes toward CBT, while few teachers were

positive about PBT as well. Findings also revealed that those teachers who had experience of using computer-based tests or those who had participated in computer training courses had more positive attitudes than the other groups of teachers. Thus, it can be concluded that teacher training for the use of computers in testing can improve teachers' attitudes toward CBT. Using a mixed methods study, Broughton, Robinson, and Hernandez-Martinez (2013) analyzed the perceptions of 13 university lecturers on the use of computers in assessment. The lecturers pointed out both advantages and disadvantages of the implementation of CBT. Increase in students' motivation, provision of feedback for a large number of students, and socialization of learning were the perceived merits of CBT. The disadvantages of the implementation of CBT included the poor quality of feedback and teachers' resistance to the change to computer-based tests. The participants asserted that CBT can be an effective method of assessment if its drawbacks are identified and accommodated. In Korea, Joo (2007) undertook a study on comparing students' and teachers' attitudes toward a computerized oral test (COT) and face-to-face interview (FTFI). The findings indicated that the teachers perceived the use of COTs as effective and beneficial. Test fairness and practicality were the most significant merits of COTs, while lack of interaction and validity were the major drawbacks of COTs to the research participants. In general, the teachers showed more positive attitudes toward COTs than the students. Koppel and Hollister (2003) assessed the suitability of the use of CBT compared to PBT. A comparison of students' and teachers' perspectives on CBT showed that the teachers had more positive attitudes than the students and believed that CBT offers several benefits such as time-efficiency, simple and convenient use, and improved test validity.

As for students' perspectives on the use of computer-based tests, Tella and Bashorun (2012) investigated the attitudes of 2,209 Nigerian students from ten faculties on the use of computer-based tests. A majority of the students preferred CBT to PBT due to the fact that CBT was perceived to increase their performance. Slow speed of the Internet, data loss, limited computer skills, and lack of computer-based facilities were the major obstacles to the implementation of CBT. In Iran, Dashtestani and Sharifi (2012) explored the perceptions of Iranian EAP students of the use of web-based tests to assess academic vocabulary learning. They concluded that the use of technology in assessing academic vocabulary improved students' attitudes toward testing and learning academic vocabulary. Furthermore, the students showed high levels of self-efficacy in the use of computers and web-based tests. Salimi et al. (2011) compared EFL students' attitudes toward CBT and PBT. They pointed out that the students preferred the use of computer-based tests to paper-based tests and had positive attitudes toward CBT.

To conclude, the analysis of the previous literature on teachers' attitudes toward the implementation of CBT shows that there is a general consensus on the efficiency and benefits of implementing CBT in educational contexts. However, it appears that teachers are aware of the limitations and challenges of the implementation of CBT. The results of comparative studies suggest that teachers prefer the use of computer-based tests over paper-based tests. The findings of research on teachers' attitudes toward CBT are in accordance with the results of research on teachers' attitudes toward the use of technology in education.

Rationales for conducting this study

Over the years, experts in language testing have directed their attention toward the implementation of CBT (e.g., Banerjee, 2003; Chapelle, Chung, Hegelheimer, Pendar, & Xu, 2010; Chapelle, Jamieson, & Hegelheimer, 2003; Choi, Kim, & Boo, 2003). According to Choi et al. (2003), basically two reasons have inspired language testers to become interested in the implementation of CBT in language teaching, i.e. controversy over the use of Item Response Theory (IRT) in language teaching and the potential of CBT in increasing interactivity in language testing settings. It has been suggested that CBT can be an influential tool for language testers due to its comparability with PBT (Choi, Kim, & Boo, 2003; Jamil, Tariq, & Shami, 2012).

More specifically, even though a considerable body of research has been undertaken to examine the validity, reliability, and efficiency of CBT for language testing (e.g., Fulcher, 2003; Malabonga, Kenyon, & Carpenter, 2005; Lokan & Fleming, 2003; Ockey, 2007; Stricker, 2004), limited attention has been directed toward the issue of perceptions of language teaching stakeholders, including language teachers, towards the implementation of CBT. Atkins and Vasu (2000) point out that teachers' attitudes toward technology would have a direct impact on the future use of technology in educational contexts. In language teaching contexts, Jones (2001) argues that technology will not be used in language teaching contexts unless teachers and students have positive attitudes toward it in their teaching and learning experiences. Considering the teachers' pivotal role in language teaching contexts (Dashtestani, 2012; Jones, 2001; Richards, 2001), their attitudes toward the use of technology should be regarded as important and influential. To date, limited attention has been directed toward the issue of EFL teachers' perceptions of the use of computer-based tests both in Iran and other countries (Broughton, Robinson, & Hernandez-Martinez, 2013; Jamil, Tariq, & Shami, 2012; Laborda & Litzler, 2011).

Research questions

This study was conducted to unravel the perceptions of Iranian EFL teachers on the use of CBT in language teaching contexts. Considering the issues and problems discussed earlier, the following research questions were formulated in this study:

1. What are the attitudes of Iranian EFL teachers toward the implementation of CBT in EFL courses?
2. What is the current situation of the implementation/use of CBT in Iranian EFL courses?
3. What are the perceptions of Iranian EFL teachers on possible challenges, strategies and   measures to include CBT in language teaching contexts?

The study

Research design, instruments and data analysis

This study is conducted using a mixed-method design in order to obtain validated and rich data. In addition, this research design was used since methodological triangulation is one of the most influential research designs in the educational research (Best & Kahn, 2006). More importantly, Cohen, Manion, and Morrison (2000) encourage researchers in the field of education to conduct research using multi-methods approaches. The exclusive use of one method might distort the description of a situation or behavior. The use of multi-methods approaches would enhance both reliability and validity of the findings of a research study. The combination of qualitative and quantitative approaches to data collection in the current study provided the researcher with confirmatory and supplementary data. Therefore, three instruments including questionnaires, interviews, and syllabus analysis were employed in the current study.

The development of the survey

"Surveys are one of the most common methods of collecting data on attitudes and opinions from a large group of participants" (Mackey & Gass, 2005, p. 92). The quantitative data of the current study were collected using a Likert-item questionnaire. Initially, the development of the survey came from the critical review of the relevant literature on the use, merits, drawbacks, and challenges of CBT for language teaching and other educational contexts and fields (e.g., Brown, 1997; Chapelle, 2007; Chapelle & Douglas, 2006; Conole & Warburton, 2005; Guiran & Shuli, 2010; Sorana-Daniela & Lorentz, 2007). Input to design the questionnaire was also received by conducting interviews with 30

Iranian EFL teachers prior to the development of the questionnaire about the use and challenges of administering CBT in the Iranian EFL contexts. The items were then written with regard to the input received from the above-mentioned sources. A jury of three professors of education, three professors of EFL, and two professors of educational technology were invited to evaluate the content validity of the survey. A checklist containing Likert-item and open-ended questions regarding the suitability of the content and appropriateness of the items of the survey for the aims of the present study was submitted to the jury to check the content validity of the items. The items of the questionnaire were amended and improved after five consulting sessions with the same jury. Some items were added or removed during this stage of survey development based on the received feedback and advice from the jury of the experts.

Specifically, a Likert-item questionnaire was employed in this study (Appendix A). The questionnaire comprised four sections, i.e. teachers' perspectives on the implementation of CBT in EFL courses (items=10, Cronbach's Alpha= 0.83), EFL teachers' perceptions of challenges to including/using CBT in EFL courses (items=11, Cronbach's Alpha=0.92), the current use of computer-based tests in the Iranian EFL courses (items= 2), and EFL teachers' perceptions of the strategies to include CBT in EFL courses (items=8, Cronbach's Alpha=0.88). Accordingly, a Cronbach's Alpha index more than 0.70 was regarded as satisfactory for establishing the reliability of the survey (Hair, Anderson, Tatham, & Black, 1995).

The development of the interview

According to Mackey and Gass (2005), interviews are appropriate tools to investigate factors which are not directly observable including attitudes, perceptions, and opinions. In order to provide triangulated data, a second instrument, i.e. an in-depth interview, was included in the study. The focus of the interview questions was the same issues in the questionnaire phase of the study. The same literature and panel of experts was used to develop the questions of the interview. The purposes of the study were explained for each participant before his/her voluntary participation in the interview study. The following interview questions were used in the study:

1. What do you think about the use of computers in language testing?
2. What do you think about your competence/ability to develop/use computer-based tests in your language teaching?

3. What do you think are the possible challenges to the use/development of computer-based tests?

4. What kinds of tests do you use to assess your students?

5. What kinds of computer-based test development/use activities have you done?

6. What do you think are the possible strategies to include CBT in language teaching contexts?

Syllabus analysis

Handelsman, Rosen, and Arguello (1987) point out three criteria for syllabus analysis. These criteria include the description of the course, objectives of a course, and activities used in a course. Since the focus of this study was language testing, a checklist was prepared to check the availability or use/non-use of certain aspects of CBT in different elements of the syllabus, i.e. course description, course objectives, and activities, of each language teaching institution. The checklist comprised two sections. The first section included some items related to the use/non-use of any kind of computer based tests in the syllabus, including computerized adapted tests, web-based language tests, e-portfolio assessment, and multimedia-assisted tests. The items of the second section were designed to track down the use of computers in the assessment of different language skills and sub-skills, i.e. speaking, listening, reading, listening, pronunciation, vocabulary, and grammar. The syllabus analysis was used to identify the use of any kind of computer-based tests and to assess the current state of testing approaches in the Iranian EFL courses. The purposes of the study were explained to the educational supervisors of each institution under study. After ensuring anonymity and ethical aspects of the study, 11 educational supervisors agreed to submit their syllabus to the researcher.

The 11 syllabi were chosen from the 16 language teaching institutions from which the participants of the study were selected. These institutions were selected based on cluster sampling. These institutions were located in the capital of Iran (Tehran) and Alborz which is the nearest big city to the capital. All these institutions had a high rank and more teaching facilities and resources compared to other language teaching institutions located in other provinces.

Data analysis and procedures

The teachers who participated in the present research were invited to complete the questionnaires face-to-face. SPSS 16 was used to provide the mean and standard deviation of the responses to each questionnaire item. The interview data were recorded and then transcribed. The interview results were

analyzed through identifying the emerging themes and content analysis. For inter-rater reliability, two raters coded the data and the common emerging themes were reported.

Sampling

Of the 300 questionnaires administered, 247 were returned. Some questionnaires were not considered since the responses were incomplete or unclear. The teachers who participated in the study were selected from 16 language teaching institutions in the provinces of Tehran and Alborz in Iran. Cluster sampling was used to select and include the participants in the study. The institutions were categorized into several clusters and simple random samples of the groups were selected. The average age of the participants was 33.7. Both females (n=151) and males (n=96) were included in the study. All the participants had attended teacher training/education courses prior to their employment in their language teaching institution. The EFL teachers had taught English for an average of 10.4 years. They had used computers for an average of 8.7 years. All the participants had regular access to the Internet and computers. Most participants perceived their levels of English proficiency as advanced (n=211) and a few of them perceived their English proficiency level as upper-intermediate (n=36). The EFL teachers taught different levels of proficiency to adult EFL learners. A total of 136 language teachers had studied English-related majors while 111 teachers had studied majors not relevant to English at university.

After the administration of the survey, all participants were invited to take part in a 30-minute interview. Of all 247 participants, 68 participants accepted the invitation to participate in the interview phase of the study. Extreme care was exercised to ensure the voluntary participation of the participants in the interviews.

Findings

EFL teachers' perspectives on the implementation of CBT in language teaching

Triangulation of interview and questionnaire data

Generally, the results of the survey on EFL teachers' perceptions of the use of CBT in EFL courses showed that the teachers perceived the implementation of CBT as positive (total mean=2.956). Specifically, storage efficiency, interactivity, ease of administration, improved reliability and impartiality, provision of instant feedback, creation of motivation for students, and time and location flexibility were the perceived benefits of the implementation of CBT (Table 1).

Ease of design, and cost-effectiveness were not perceived as advantages of CBT by the participants.

Table 1

Perspectives on the implementation of CBT in EFL courses

Questionnaire items	Mean	Standard deviation
Little physical space is required to store answer scripts of CBT	3.51	0.82
CBT can be interactive by the use of multimedia	3.42	0.72
CBT is easy to be administered	3.26	0.97
Using CBT improves the reliability of assessment	3.09	0.83
Instant feedback can be given to students using CBT	3.01	0.98
Using CBT improves impartiality	2.99	0.84
The use of CBT is motivating for EFL students	2.73	1.06
CBT can be administered everywhere at anytime	2.65	0.92
Designing a CBT is cost-effective	2.16	0.83
CBT is easy to be designed	1.79	0.89

The results of the interviews were in line with the results of the survey. One point that was mentioned both in questionnaires and interviews was that CBT may improve interactivity and enable EFL teachers to design and use interactive testing tasks and items. Specifically, in the interviews, the teachers reported that videos, audios, and pictures can be used in CBT and this makes CBT suitable for assessing EFL learners' English proficiency. Likewise, in the questionnaire the teachers pointed out that the use of multimedia in CBT can improve the level of interactivity in the EFL classroom.

> *It's a good idea to use CBT in EFL contexts. Computer-based tests are useful for assessing EFL learners. I believe that one of the most beneficial aspects of CBT is that the use of these types of tests enables EFL teachers to use pictures, videos, different fonts, etc. This property can make the test more appealing and motivating for EFL learners.* (Teacher19)

Well, CBT can improve the quality of testing and assessment in EFL. A lot of different audio-visual features which do not exist in paper-based tests can be used in a computer-based test. I think this is a great merit. (Teacher 24)

One new perceived merit of CBT, which was not reported in the questionnaires, was that CBT facilitates the process of scoring for EFL teachers.

I suppose computers will help EFL teachers to score the papers more easily. This will save EFL teachers a lot of time and energy. Personally, scoring and test paper correction is a big challenge for me. (Teacher 31)

There will be fewer human errors in scoring the test when computers are used in language testing. Thus, more accurate and fair scores will be reported to our students. (Teacher 8)

Most EFL teachers participated in the interviews considered the use of CBT as useful and effective for language teaching contexts. It appears that the teachers preferred CBT to PBT and traditional approaches to language testing. Similar findings were identified in the questionnaires accordingly.

CBT can be a good testing tool for EFL teachers. I believe the use of CBT is more contributive than the use of traditional testing methods and procedures. (Teacher 39)

We can use computers in different aspects of language teaching. Testing and assessment are not a different issue and computers would be used in this aspect of language teaching as well. (Teacher 44)

My opinion is that in the future there won't be any paper-based test and CBT will be used commonly in different educational fields. (Teacher 4)

The results of the interviews and questionnaires indicated the positive attitudes of EFL teachers toward the implementation of CBT. The use of multimedia and increase in interactivity were the perceived benefits which were reported both in interviews and questionnaires. However, ease of scoring was a new merit of CBT which was reported only in the interviews.

EFL teachers' perceptions of challenges of including/implementing CBT in language teaching

Triangulation of interview and questionnaire data

Based on the values shown in Table 2, the EFL teachers who took part in this study perceived several challenges to the implementation of CBT in EFL courses, i.e. lack of computer-based facilities, lack of training/education on the implementation of CBT, lack of access to computer-based tests, their lack of technological and testing knowledge, lack of support to implement CBT, and high cost of developing computer-based tests.

However, the EFL teachers did not perceive students' resistance to CBT and the lack of CBT security as important challenges to the implementation of CBT in EFL courses.

Table 2

Perspectives on challenges to the implementation of CBT in EFL courses

Questionnaire items	Mean	Standard deviation
There is lack of technological facilities to use CBT	3.43	0.73
Lack of training/education on how to use/develop computer-based tests	3.37	0.88
Computer-based tests are not accessible to EFL teachers	3.34	0.63
EFL teachers lack technological knowledge on how to develop computer-based tests	3.33	0.83
Lack of support from educational authorities to help teachers use/develop computer-based tests	3.20	1
EFL teachers lack testing knowledge to develop computer-based tests	2.94	0.94
EFL teachers lack knowledge how to use computer-based Tests	2.85	1.02
Computer-based tests are expensive to be developed	2.75	0.96
CBT cannot be used for testing different language skills	2.56	1.05
There is lack of test security when computer-based tests are used	2.38	0.91
Some students might show resistance to CBT	2.03	0.85

A finding echoed both in interviews and questionnaires was that the majority of EFL teachers mentioned that they did not know how to develop and use computer-based tests for EFL purposes. As

reported in the questionnaires, the teachers did not have the required testing and technological knowledge to develop and use computer-based tests.

> *To be honest, I don't know how to use and develop computer-based tests. I reckon that it's a difficult task to develop computer-based tests.* (Teacher 51)

> *I don't have any knowledge about developing and designing computer-based tests. Of course, developing CBT requires certain types of knowledge and we need training on how to develop these tests.* (Teacher 62)

> *I know very little about the use of computers and technology in language teaching and testing. Generally, in my view, the use of computers in language testing is different from using it in language teaching. When it comes to language testing, more complicated skills are needed.* (Teacher 23)

Another finding in both interviews and questionnaires was lack of computer-based facilities, which was perceived to be the other major challenge to the inclusion of CBT in language testing. Most EFL teachers reported that there are not sufficient technological facilities to use or develop computer-based tests for language teaching.

> *In my opinion, the major challenge is to provide adequate and high-quality computer-based facilities to use computers in language testing. As it is clear, considering the current facilities and capacities, it seems that it's not possible to use computers in language testing in the Iranian EFL context.* (Teacher 14)

> *I've travelled to other countries and have seen their computer-based equipment. We cannot compare their facilities with ours. So I think the first obstacle is to equip language teachers with up-dated technologies and computer-based affordances.* (Teacher 26)

Likewise, based on the results of the interviews and questionnaires, suitable computer-based tests related to EFL were not easily accessible and available to EFL teachers. Many teachers stated that they did not know how to have access to new and relevant computer-based tests which they need.

> *Unfortunately, in our country [Iran] it's very hard to find and use computer-based tests since most of these tests are not free or are not developed for the EFL use. We (EFL teachers) don't*

know how to find and select suitable computer-based tests which are compatible with our students' needs. (Teacher 23)

I really don't have any idea how I can find some good computer-based tests for my EFL classes. I guess we have few suitable computer-based tests which are available to language teachers. (Teacher 55)

A finding which was not reflected in the questionnaires but shown in the interviews was that some EFL teachers reported that there are not any well-designed local computer-based tests to be used for their EFL courses and assessment activities.

One serious problem with CBT is that the majority of computer-based tests are developed in other EFL or educational contexts and they may not be suitable enough for our specific EFL context. (Teacher 3)

The idea of using computers in language teaching is good, but we should be able to design context-based tests first. (Teacher 44)

Based on the results of the interviews, some EFL teachers had concerns that students and educational authorities were accustomed to the use of traditional and paper-based tests and their attitudes should be changed toward the use of computer-based tests.

It seems that at present most teachers, students, and educational authorities trust paper-based tests. I suppose it takes time to change their attitudes and make them interested in the use of computer-based tests. (Teacher 66)

To summarize, teachers' lack of knowledge of computers and CBT, lack of computer-based facilities, and lack of access to computer-based tests were the limitations identified in the questionnaires and interviews.

Current use/implementation of CBT in the Iranian language teaching context

Triangulation of interview, questionnaire, and syllabus analysis data

As mentioned, a total of 11 class syllabi were accessed with the permission of educational supervisors of each language teaching institution. The results of the syllabus analysis showed that no type of

computer-based test, including computerized adapted tests, web-based language tests, e-portfolio assessment, or multimedia-assisted tests, were included in the syllabi of the language teaching institutions. Paper-based mid-term and summative final exams were the only formal testing activities found in those syllabi. It was also revealed that computers were not used in testing any type of language skill in the Iranian language teaching institutions.

The results of the survey regarding EFL teachers' current use of computer-based tests showed that a large number of EFL teachers never use computer-based tests in their EFL courses (n=228, 92.3 %). A few EFL teachers (n=19, 7.7%) reported that they rarely make use of computer-based tests in their EFL courses. Furthermore, 99.2% (n=245) of the participants had not developed/adapted any computer-based test.

As reported in the survey, the interview participants stated that they did not use computer-based tests. They had not been involved in any kind of computer-based test development or adaptation. The EFL teachers reported that they commonly use paper-based tests.

Based on the results of the three instruments, it is evident that CBT is not a part of language testing in the Iranian EFL curriculum. Neither the teachers nor the institutions use computer-based tests to assess students' achievement.

EFL teachers' perceptions of the strategies and measures to include CBT in language teaching

Triangulation of interview and questionnaire data

As Table 3 illustrates, the EFL teacher participants agreed on the importance of some strategies to include CBT in EFL courses, i.e. changing traditional approaches to testing, and evaluation, encouraging EFL teachers to use CBT, providing access to computer-based tests, improving computer-based facilities, providing software tools for the implementation of CBT, fostering EFL teachers' technological and testing competence, and training on developing/using CBT.

Table 3

Perceptions of the strategies to include CBT in EFL courses

Questionnaire items	Mean	Standard deviation
Changing current traditional approaches to testing and evaluation	3.70	0.74
Encouraging EFL teachers to use CBT in their assessment activities	3.63	0.69
Providing easy access to suitable computer-based tests for EFL teachers	3.58	0.62
Improving computer-based facilities to use computer-based tests	3.52	0.61
Providing EFL teachers with software tools required for CBT	3.46	0.77
Enhancing EFL teachers' level of technological knowledge	3.29	0.89
Training EFL teachers to develop/use computer-based tests	3.27	0.69
Enhancing EFL teachers' level of testing knowledge	3.25	0.78

Similar to the questionnaire data, the results of the interviews about the participants' perceptions of the strategies to include CBT in EFL courses depicted that providing computer-based equipment is one of the important measures which might facilitate the inclusion of CBT in EFL courses.

At present, it seems that the first important obstacle to the inclusion of technology is lack of appropriate technological equipment. Solving the problem of lack of facilities will have a direct effect on the inclusion of technology in language testing and assessment as well. (Teacher 27)

Our EFL classes should have computers and other technological tools if we want to use computer-based tests in EFL courses. (Teacher 35)

Once again, the results of the interviews are in line with the results of the questionnaires in that the majority of teachers asserted that there should be some training for EFL teachers on how to use/develop computer-based tests.

What seems necessary is to enable us [EFL teachers] to use/develop computer-based test through planning some training courses or workshops. These courses and workshops will have implications for the inclusion of CBT in EFL contexts. (Teacher 46)

Well, CBT should be the focus of teacher training courses if it is supposed to be included in our language testing system. (Teacher 8)

Based on the results of the interview data, some teachers pointed out the necessity of attitudinal changes toward CBT and teacher encouragement to use computer-based tests as well.

To conclude, training teachers to use computer-based tests and improving computer-based facilities were similar to findings from the questionnaires and interviews.

Discussion

The study employed a mixed-method design to investigate the attitudes of Iranian EFL teachers toward the use of computer-based tests in language teaching contexts. The results of the study indicated that even though the language teachers show positive attitudes toward CBT and its implementation in EFL courses, such a practice is not seen in the actual teaching field. The finding of positive attitudes is commensurate with similar studies which have revealed the positive attitudes of teachers toward the implementation of CBT (Jamil, Tariq, & Shami, 2012; Laborda & Royo, 2009). Since language teachers play a very important role in the implementation and inclusion of CBT in language testing (Chapelle & Douglas, 2006), their positive perceptions of the incorporation of CBT in language testing are relevant and influential. Similarly, for successful implementation of CBT, it is crucial that teachers feel enthusiastic about the implementation of CBT. However, it should be noted that the positive attitude of language teachers is not the only factor to guarantee the successful implementation of CBT in language teaching. The positive reaction of language teachers should be taken into account by EFL educational authorities and teachers should be supported to use and develop computer-based tests in EFL courses. As Jones (2001) pointed out, it is not possible to use technology in EFL if language teachers do not have positive perspectives on the use of technology. Furthermore, this positive attitude of language teachers toward CBT might show that they are willing to have more knowledge and expertise in order to use and develop different types of computer-based tests. Obviously, positive attitudes should be combined with high computer literacy levels and testing knowledge to use and develop computer-based tests appropriately.

The discrepancy between the teachers' positive attitudes and their lack of using computer-based tests may be related to several factors which discourage teachers from using technology in their classrooms. It should be noted that while the adoption of positive attitudes is necessary for technology inclusion, it is not a sufficient condition. This discrepancy of perceptions and the actual use of computer-based tests can be related to a plethora of barriers which impede the inclusion of technology in EFL instruction (Dashtestani, 2012). These barriers might be directly associated with teachers'

factors, such as low computer literacy levels and lack of knowledge of developing and using computer-based tests, or institutional factors, such as lack of support from educational authorities, negative attitudes of educational authorities to technology, and rigid and inflexible curricula. Whatever the reasons for the discrepancies are, further research should be directed towards reducing the gap between teachers' perceptions of technology and their actual use of technology in EFL instruction (Dashtestani, 2014).

In addition, the findings were indicative of the perceptions of the language teachers on various obstacles and challenges, which might discourage language teachers from using and developing computer-based tests in language testing. These obstacles included lack of computer-based facilities, lack of training on developing computer-based tests, lack of accessibility, teachers' lack of technological and testing knowledge, lack of support, and high cost of developing computer-based tests. One of the serious barriers to the implementation of CBT is EFL teachers' lack of testing and computer literacy and competence in order to develop and use computer-based tests in language testing. This finding is compatible with Laborda and Royo's (2009) findings, which showed that despite teachers' positive attitudes towards CBT, they did not have adequate computer literacy levels to make use of computer-based tests. Therefore, EFL authorities and providers are recommended to provide language teachers with specific training courses to enhance their levels of computer literacy and assessment competence. Alternatively, this computer literacy training can be included in mainstream teacher training/education programs. What seems necessary is to raise language teachers' awareness about their need for fostering their testing and computer competence since in the future the use of computer-based tests will be unavoidable in educational and EFL contexts. The other considerable challenge is lack of computer-based equipment and facilities in EFL courses. Lack of technological facilities might discourage teachers from using technology in their teaching and testing activities. Lack of facilities would also cause a change in teachers' positive attitudes toward technology inclusion. In terms of CBT, language teachers should be provided with adequate computer-based facilities, suitable and up-to-date CBT software tools, and easy access to technologies related to CBT implementation.

Concerning the current state of the use of CBT in the Iranian EFL courses, the findings of the three research instruments depicted that traditional approaches to language testing are commonly used for assessing EFL students' performance. This lack of CBT and technology use might be caused by the presence of challenges discussed earlier. Taking into consideration the facilitative role of technology in language testing, a change in traditional approaches to language testing should occur and teachers

should modify their testing methods through using electronic and technology-enhanced assessment (Jamil, Tariq, & Shami, 2012). To achieve this aim, educational authorities are supposed to remove the potential barriers to the use of technology in language teaching and testing. Admittedly, this change to CBT should be welcomed and accepted by all EFL stakeholders including language teachers, EFL students, course designers, and EFL test developers prior to its implementation.

The language teachers offered some strategies which might facilitate the inclusion of CBT in language testing. The first major strategy proposed by the Iranian EFL teachers was implementing changes in traditional approaches to testing. Implementing this change in the language testing system of Iran requires the attention and cooperation of various language testing stakeholders who can make decisions about the future of language testing. Change in approaches to language testing may be closely associated with educational stakeholders' attitudes toward the use of technology in education. Therefore, all decision-makers and planners should take appropriate measures in order to provide teachers with the opportunity to adopt innovative and technological approaches to language testing.

The other significant measure to facilitate technology inclusion in language testing is to provide EFL teachers with the required resources and facilities which are essential to implement CBT effectively. As previous research has shown, one serious barrier to the use of technology in EFL instruction is the lack of required technology-based facilities (Albirini, 2006; Dashtestani, 2012; Ismail, Almekhlafi, & Almekhlafy, 2010; Kim, 2002; Lau & Sim, 2008). Concerning CBT, it appears that EFL teachers need to have access to up-dated hardware and software tools. Without equipping our educational system with appropriate types of facilities, we cannot expect teachers to improve their teaching and testing quality and develop professionally.

As Mishra and Koehler (2006) argue, teachers' pedagogical knowledge and technology knowledge should be improved at the same time. In terms of implementing CBT, the findings indicated that the teachers believed that their computer literacy and testing knowledge should be fostered. It is important to hold workshops and teacher training courses in order to improve Iranian teachers' testing and computer knowledge. Apparently, a long-term training program should be considered to encourage and train language teachers to use computer-based tests effectively. Definitely, without taking these measures into account, the implementation of CBT would not be feasible. Most of these measures can be taken by EFL authorities to help language teachers include CBT in language testing. In addition, EFL authorities might be aware of more facilitative strategies and measures on how to include technology in language testing. Similar strategies on technology inclusion have been reported by

several EFL researchers and scholars (Hubbard, 2008; Kassen, Lavine, Murphy-Judy, & Peters, 2007; Son, Robb, & Charismiadji, 2011).

Finally, more context-based studies should be carried out in different educational contexts on how to implement CBT successfully and efficiently. Also, the perceptions of different EFL stakeholders should be investigated. Apparently, the effective implementation of CBT requires the positive perceptions of students, teachers and other EFL stakeholders. In addition, more research is required to examine the limitations of CBT for language teaching contexts.

Conclusion and implications

The present study aimed at investigating perceived challenges to the implementation of CBT in Iranian EFL teaching contexts. Based on the findings of this study, language teaching stakeholders, including course designers, authorities, providers, and teachers, should facilitate the inclusion of CBT in the language curriculum through taking various measures. The incorporation of CBT in the EFL curriculum would require different stakeholders of language teaching to cooperate in order to equip language teachers with the necessary skills of the implementation of CBT. Accordingly, language teaching stakeholders should strive to develop positive attitudes toward the implementation of CBT due to the affordances and merits which CBT would offer for language teachers and learners. Therefore, further research is required to provide insights into the perceptions of the other language teaching stakeholders. It might be worth mentioning that while the findings of this study would be limited to the specific context of Iran, the findings can have implications for other language teaching contexts. Research on CBT in language teaching context is at its early stages and further investigation is needed to unravel various parameters that should be taken into account when implementing CBT.

Although this study depicted the positive perspectives of language teachers on the implementation of CBT, further research should be undertaken to compare language teaching stakeholders' preference of CBT over PBT. Although the findings of this study indicate that the implementation of CBT would be beneficial, it does not mean that CBT is necessarily more efficient than PBT. Both CBT and PBT can be advantageous with regard to the particular purposes they serve. Thus, it is recommended that language teachers take into consideration the specific purposes of their unique testing and teaching contexts and learners prior to the implementation of CBT.

Furthermore, language teachers can have considerable effects on motivating their students to use technology for their learning. In terms of CBT, language teachers can train students to take computer-

based tests and make them aware of the benefits and merits of CBT. If teachers adopt negative attitudes toward the use of technology in language teaching and learning, students' motivation to use technology will be reduced. As a consequence, taking into account the findings of this study, language teaching supervisors and authorities can pave the way for the implementation of CBT in EFL courses and encourage teachers to prepare their students for the successful implementation of CBT.

One significant implication of this study would be the fact that language teachers need training to be able to implement CBT and use computer-based tests effectively and appropriately. Laborda and Royo (2009) argue that one major challenge to the implementation of CBT is teachers' lack of familiarity with computers and technology. Teachers' computer literacy is an integral condition for successful implementation of CBT, especially when it is used for high stakes tests. As a result, EFL authorities and supervisors are invited to hold workshops and training programs on fostering language teachers' computer literacy levels prior to the implementation of CBT in language teaching contexts. Enhancing teachers' computer literacy would also increase their confidence and motivation to use computers and implement CBT.

More importantly, poorly-developed computer-based tests might discourage language teachers and students from their use. The attitudes of teachers and students toward CBT would depend to a great extent on the quality of computer-based tests. Therefore, language test developers and course designers are strongly recommended to take several measures to ensure the reliability and validity of computer-based tests. The mere use of technology in language teaching will not guarantee effective learning. It is essential to take heed of the way technology is used in language teaching and learning.

Limitations of the study

The conduction of this study was not without limitations. First of all, the teachers who participated in this study were those who were employed and worked in language teaching institutions. Other teachers who worked at universities and schools could provide valuable insights into the issue of this research study. Also, the researcher wished to include perceptions of EFL students on the implementation of CBT while it was not feasible since EFL students were not supposed to have the knowledge of testing and using computers in language teaching.

References

Adair, D., Jaeger, M., & Pu, J.H. (2012). *Assessing student attitudes using a computer-aided approach.* Paper presented at AAEE conference. Melbourne, Australia.

Al-Amri, S. (2008). Computer-based testing vs. paper-based testing: A comprehensive approach to examining the comparability of testing modes. *Essex Graduate Student Papers in Language & Linguistics, 10*, 22-44.

Albirini, A. (2006).Teachers' attitudes toward information and communication technologies: the case of Syrian EFL teachers. *Computers & Education, 47*(4), 373–398.

Alderson, J.C. (2000). *Assessing reading*. Cambridge: Cambridge University Press.

Atkins, N. E., & Vasu, E. S. (2000). Measuring knowledge of technology usage and stages of concern about computing: A study of middle school teachers. *Journal of Technology and Teacher Education, 8*(4). 279–302.

Banerjee, J. (2003). The TOEFL CBT (Computer-based test). *Language Testing, 20*, 111-123.

Best, W. B., & Kahn, J. V. (2006). *Research in education* (10th ed.). New York: Pearson Education Inc.

Broughton, S. J., Robinson, C. L., & Hernandez-Martinez, P. (2013). Lecturers' perspectives on the use of a mathematics-based computer-aided assessment system. *Teaching Mathematics and its Applications*, *32*(2), 88-94.

Brown, H. D. (2007). *Teaching by principles: An interactive approach to language pedagogy.* White Plains, NY: Pearson Education.

Brown, J. D. (1997). Computers in language testing: Present research and some future directions. *Language learning and Technology, 1*(1), 44-59.

Chapelle, C. A. (2007). Technology and second language acquisition. *Annual Review of Applied Linguistics, 27*, 98-114.

Chapelle, C. A., & Douglas, D. (2006). *Assessing language through computer technology.* Cambridge: Cambridge University Press.

Chapelle, C., Jamieson, J., & Hegelheimer, V. (2003). Validation of a web-based ESL test. *Language Testing, 20*(4), 409-439.

Chapelle, C. A., Chung, Y. R., Hegelheimer, V., Pendar, N., & Xu, J. (2010). Towards a computer-delivered test of productive grammatical ability. *Language Testing, 27*(4), 443-469.

Choi, I. C., Kim, K. S., & Boo, J. (2003). Comparability of a paper-based language test and a computer-based language test. *Language Testing, 20*(3), 295-320.

Cohen, L., Manion, L., & Morrison, K. (2000). *Research methods in education*. London: Routledge.

Conole, G., & Warburton, B. (2005). A review of computer-assisted assessment. *ALT Journal, 13,* 17-31.

Dashtestani, R. (2012). Barriers to the implementation of CALL in EFL courses: Iranian EFL teachers' attitudes and perspectives. *The JALT CALL Journal, 8*(2), 55-70.

Dashtestani, R., & Sharifi, A. (2012). *Web-based assessment of academic vocabulary: Iranian EAP students' attitudes, confidence and self-efficacy.* Paper presented at the 6th conference on issues in English language teaching in Iran, University of Tehran.

Dashtestani, R. (2014). English as a foreign language—teachers' perspectives on implementing online instruction in the Iranian EFL context. *Research in Learning Technology, 22*. doi:10.3402/rlt.v22.20142

Fulcher, G. (2003). Interface design in computer-based language testing. *Language testing, 20*(4), 384-408.

Guiran, G., & Shuli, Z. (2010). Research on the computer-based college English test system in China. In *Computer Science and Information Technology (ICCSIT), 2010 3rd IEEE International Conference, 7,* 587-589.

Handelsman, M. M., Rosen, J., & Arguello, A. (1987). Informed consent of students: how much information is enough. *Teaching of Psychology, 12*, 107–109.

Hair, J. F., Anderson, R. E., Tatham, R. L., & Black, W. C. (1995). *Multivariate data analysis*

Hubbard, P. (2008). CALL and the future of language teacher education. *CALICO Journal, 25*(2), 175–188. *with readings.* New York: Prentice-Hall.

Ismail, S. A. A., Almekhlafi, A. G., & Al-Mekhlafy, M.H. (2010). Teachers' perceptions of the use of technology in teaching languages in United Arab Emirates' schools. *International Journal of Research in Education, 27*, 37-56.

Jamil, M., Tariq, R.H., & Shami, P. A. (2012). Computer-based vs. paper-based examinations: perceptions of university teachers. *The Turkish Online Journal of Educational Technology, 11*(4), 371-381.

Jones, J. (2001). CALL and the responsibilities of teachers and administrators. *ELT Journal, 55*(4), 360-367.

Joo, M. J. (2007). The attitudes of students' and teachers' toward a Computerized Oral Test (COT) and a Face-To-Face Interview (FTFI) in a Korean university setting. *Journal of Language Sciences 14*(2), 171, 193.

Kassen, M. A., Lavine, R. Z., Murphy-Judy, K., & Peters, M. (2007). *Preparing and developing technology-proficient L2 teachers: CALICO Monograph Series Volume 6. San Marcos*, TX: Computer Assisted Language Institution Consortium (CALICO).

Kim, H. (2002). Teachers as a barrier to technology-integrated language teaching. *English Teaching, 57*(2), 35-64.

Koppel, N. B., & Hollister, K. K. (2003). Comparison of live versus paper-based assessment in computer application course. *Journal of Informatics Education Research, 5*(1), 39-50.

Laborda, J. G., & Litzer, M. F. (2011). Constraints in teacher training for computer-assisted language testing implementation. *International Education Studies, 4*(2), 13-17.

Laborda, J. G., & Royo, T. M. (2009). Training senior teachers in compulsory computer based language tests. *Procedia-Social and Behavioral Sciences*, *1*(1), 141-144

Lau, B. T., & Sim, C. H. (2008). Exploring the extent of ICT adoption among secondary school teachers In Malaysia. *International Journal of Computing and ICT Research, 2*(2), 19–36.

Lokan, J., & Fleming, M. (2003). Issues in adapting a computer-assisted career guidance system for use in another country. *Language Testing, 20*(2), 167-177.

Mackey, A., & Gass, S. M. (2005*). Second language research: Methodology and design.* Mahwah, NJ: Lawrence Erlbaum Associates.

Malabonga, V., Kenyon, D. M., & Carpenter, H. (2005). Self-assessment, preparation and response time on a computerized oral proficiency test. *Language Testing, 22*(1), 59-92.

Mishra, P., & Koehler, M. (2006). Technological pedagogical content knowledge: A framework for teacher knowledge. *The Teachers College Record, 108*(6), 1017-1054.

Ockey, G. (2007). Construct implications of including still image or video in computer-based listening tests. *Language Testing, 24*(4), 517-537.

Park, C. N., & Son, J.-B. (2009). Implementing computer-assisted language learning in the EFL classroom: Teachers' perceptions and perspectives. *International Journal of Pedagogies and Learning, 5*(2), 80-101.

Richards, J. C. (2001). *Curriculum development in language teaching.* Cambridge: Cambridge University Press.

Salimi, H., Rashidy, A., Salimi, E. A., & Amini Farsani, M. (2011). Digitized and non-digitized language assessment: A comparative study of Iranian EFL language learners. *International Proceedings of Economics Development & Research, 26*, 314-318.

Shi, F. (2012). Exploring students' anxiety in computer-based oral English test. *Journal of Language Teaching and Research*, *3*(3), 446-451.

Simonsson, M. (2004). Technology use of Hispanic bilingual teachers: A function of their beliefs, attitudes and perceptions on peer technology use in the classroom. *Journal of Instructional Technology, 31*(3), 257-266.

Son, J.-B., Robb, T., & Charismiadji, I. (2011). Computer literacy and competency: a survey of Indonesian teachers of English as a foreign language. *CALL-EJ, 12*(1), 26-42.

Sorana-Daniela, B., & Lorentz, J. (2007). Computer-based testing on physical chemistry topic: A case study. *International Journal of Education and Development using Information and Communication Technology, 3*(1), 94-95,

Stricker, L. J. (2004). The performance of native speakers of English and ESL speakers on the computer-based TOEFL and GRE General Test. *Language Testing, 21*(2), 146-173.

Tella, A., & Bashorun, M. T. (2012). Attitude of undergraduate students towards computer-based test (CBT): A case study of the university of Ilorin, Nigeria. *International Journal of Information and Communication Technology Education, 8*(2), 33-45.

Wang, H., & Shin, C. D. (2009). Computer-based & paper-pencil test comparability studies*. Test, measurement, and research services Bulletin, 9*, 1-6.

Wei, C.L. (2007). *A comprehensive guide to Internet-based TESL/TEFL*. Taiwan: Kuan Tang Press.

Yurdabakan, I., & Uzunkavak, C. (2012). Primary school students' attitudes towards computer-based testing and assessment in Turkey. *Turkish Online Journal of Distance Education, 13*(3). Retrieved from https://tojde.anadolu.edu.tr/tojde48/articles/article_12.htm.

Questionnaire on attitudes of EFL teachers toward implementing CBT

Dear Participants,

The following questionnaire is part of a research project that investigates the perceptions of Iranian EFL teachers of the implementation of CBT. Your responses will be treated in strict confidence and individual teachers/schools will not be identified in any report or publication. Please answer all questions as accurately as you can.

Background information

Gender:

Age:

Name of teaching institute/university/organization:

Province:

Major:

Years of teaching English:

Years of using computers:

Access to computers: Limited ___   Moderate ___   Regular ___

Access to the Internet: Limited ___   Moderate ___   Regular ___

Level of English proficiency? Basic_____   Intermediate_____ Upper-intermediate_____ Advanced_____

Section 1: Perspectives on the use of CBT in EFL courses
_____

	Strongly disagree	Disagree	Agree	Strongly agree
1. CBT is easy to be designed				
2. Designing a CBT is cost-effective				
3. Instant feedback can be given to students using CBT				
4. Using CBT improves the reliability of assessment				
5. CBT is easy to be administered				
6. Using CBT improves impartiality				

7. CBT can be interactive by the use
 of multimedia

8.The use of CBT is motivating for EFL
Students

9. Little physical space is required to
store answer scripts of CBT

10. CBT can be administered everywhere
at anytime
_____

Section 2: Current use of CBT in the Iranian EFL courses

_____

	Never	Rarely	Sometimes	Frequently	Always

_____

1. Use of computer-based tests in your EFL courses

2. Your development/adaptation of any computer-based tests

_____

Section 3: Perspectives on challenges to the use of CBT in EFL courses

_____

	Strongly disagree	Disagree	Agree	Strongly agree

_____

1. EFL teachers lack testing
knowledge to develop
computer-based tests

2. EFL teachers lack technological
 Knowledge on how to develop
computer-based tests

3. EFL teachers lack knowledge
 how to use computer-based
tests

4. There is lack of technological
facilities to use CBT

5. Computer-based tests are not
 accessible to EFL teachers

6. There is lack of test security
 when computer-based tests
are used

7.Some students might show
 resistance to CBT

8. CBT cannot be used
for testing different
language skills

9. Computer-based tests
are expensive to be
developed

10. Lack of training/education
on how to use/develop
computer-based tests

11. Lack of support from educational
 Authorities to help teachers
 use/develop computer-based tests
_____

Section 4: Perceptions of the strategies to include CBT in EFL courses
_____

	Strongly disagree	Disagree	Agree	Strongly agree

_____

1. Improving computer-based
facilities to use
computer-based tests

2. Providing EFL teachers
with software tools
required for CBT

3. Training EFL teachers
to develop/use
computer-based tests

4.Enhancing EFL teachers'
level of technological
knowledge

5. Enhancing EFL teachers'
level of testing knowledge

6. Providing easy access to

suitable computer-based
tests for EFL teachers

7. Encouraging EFL teachers
to use CBT in their
assessment activities

8. Changing current
traditional approaches
to testing and evaluation
_____

Language Test Preparation in China:

A Reflective Perspective on a Communicative Approach

Sky Lantz-Wagner

University of Dayton

Bio data

Sky Lantz-Wagner holds an MA in TESOL from the Monterey Institute of International Studies and has presented research at the California TESOL Conference. He served as a Peace Corps volunteer for two years in China where he taught English at a university in the Guizhou Province. He is currently teaching in the Intensive English Program at the University of Dayton in Ohio. His research interests include classroom innovation, and intercultural communication.

skylantzwagner@gmail.com

Abstract

This study explores how communicative language teaching impacts high stakes test preparation in China. Using data from a test preparation course designed for the Test for English Majors Band 4 (TEM4), the current research seeks to answer the following questions: what communicative language teaching methods were implemented during the preparation course as documented in a teaching journal, how did students perceive the course, and what results did the students achieve on the TEM4 after taking the course. Data were collected from over 100 college sophomores at a small university in Guizhou Province, China. Reflective analysis of diary entries revealed some of the assumptions and principles with which the course was designed and implemented. A quantitative analysis showed that students who took part in the TEM4 test preparation course scored substantially higher than students who took the test without any structured preparation the previous year.

Address correspondence:
University of Dayton, 300 College Park,
Dayton, OH, 45469

Keywords: *test preparation*, *CLT, reflective methodology, innovation, diary study, China, TEM4*

Introduction

Overview

In recent decades, several reforms to the testing system in China have been proposed, but few have been widely adopted. One reform that did eventually gain governmental support was passed in 1992, when, in an attempt to move away from the traditional grammar-translation method that has dominated language teaching in China, the national State Education Development Commission (SEDC) (1992) mandated that English be taught for communicative purposes. The reform was a radical change that failed to take into account the social, political, and historical implications of implementing such a policy in many parts of the country. Supplanting the deeply seated, historically important system based on Confucian ideas of knowledge dissemination was bound to create controversy in many areas as to how and why the new system should be adopted.

In regards to changing the nature of language teaching in China, what appears to have happened is that cities and regions, such as Shanghai and Beijing, with greater access to financial resources, have been able to hire more highly qualified teachers from within China and from overseas. Those schools have adapted more quickly to the mandated change in language teaching (Qiang, Wolff, Hai, & Gregory, 2004). Rural areas, with fewer resources and with larger numbers of schools that are under control of provincial governments, have taken longer to adapt.

Washback

Standardized tests in China play a critical role throughout college life. For instance, Chinese English majors are required to take the Test for English Majors Band 4 (TEM4), a criterion-referenced, high-stakes exam administered during the fourth semester of a four-year degree (Cheng, 2008). The TEM4 contains six subtests: dictation, listening comprehension, grammar and vocabulary, cloze, reading comprehension, and writing. The listening comprehension subtest is subdivided into dialogues, passages, and news items; the writing subtest is subdivided into composition and note writing. The total test-taking time is 135 minutes. English majors in China also take the TEM8 during their eighth semester of college, even if they do not pass the TEM4. The likelihood of finding a job after graduation increases significantly with a passing score on either exam. Therefore, there is considerable pressure on students to prepare for the tests, even if their current proficiency level makes it unlikely that they will pass.

Pressure to do well on tests lies not only on students, but also on teachers. Thus, testing has a significant bearing on the classroom, a phenomenon known as *washback* or "the positive or negative impact of a test on … teaching or learning" (Richards & Schmidt, 2010, p. 634). Several studies have shown the washback effects of standardized tests such as the IELTS and TOEFL (e.g., Alderson & Hamp-Lyons, 1996; Raimes, 1990, respectively). These studies reported that teachers feel constrained by factors such as textbooks, time, and lack of experience with the test itself. Another study by Zhan (2008) measured the washback on Chinese learners as a result of taking the College English Test Band 4 (CET4), a criterion-referenced English test designed for non-English majors in China. The results showed washback was likely to occur as a result of test preparation for the CET4, but varied depending on the attitudes and study habits of the test takers.

Despite the large number of studies conducted on washback, it is still not entirely clear how testing affects teaching and learning. This may be due partly to the fact that many washback studies are qualitative in nature. The current study seeks to fill a gap in the literature by providing insight into the relationship of testing, teaching, and learning by combining qualitative reflection with quantitative evidence.

Test Preparation

Test preparation is defined by the National Council on Measurement in Education (2012) as "any of a number of activities in which a prospective test taker might participate, primarily for the purpose of optimizing their score on an upcoming test." There are many forms of test preparation, but most often test preparation focuses on the skills a teacher perceives as necessary to pass the test. For example, a multiple choice test format, common to most standardized tests, tends to measure mechanical skills such as rote memorization, but not higher-order thinking skills or communicative ability (Scouller, 1998).

On the other hand, communicative language tests employ communicative language teaching (CLT) principals to prepare students for real life communicative tasks. Testing communicative competence has posed challenging questions for many language teachers and researchers (McNamara, 2001). Many of these challenges stem from the large quantity of resources required to simulate real-life situations. Nevertheless, there is an effort among many English language teachers to increase the communicativeness of language assessment by using information gap activities, role-play situations, letter writing, and summarizing (Kitao & Kitao, 1996).

While examinations seem to limit the ability for teachers to innovate, especially in a Chinese context (e.g., Deng & Carless, 2010), teachers and researchers alike have begun to explore the possibility of adapting teaching practices to serve a dual purpose: to communicate effectively in the target language and to pass high stakes exams. Read and Hayes (2003) conducted a study in which they compared two IELTS preparation courses in New Zealand. One course offered "intensive, teacher-directed preparation for taking the test" while the other course adopted a more "student-centered methodology and [attempted] to work on the students' needs to develop their language competence" (p. 188). Students in the course that focused on language competence had more positive views of their course than their counterparts, and while there were no significant differences when comparing the test scores for the two courses, there was a larger average increase from pre-test to post-test scores for the students in the student-centered course.

In a Chinese context, Xie and Andrews (2013) explored the impact of test design and test uses on test preparation. Their research indicated that students who endorsed high-stakes exams were more likely to engage in multiple modes of test preparation and seek additional help when there was a perceived needed. Although their research did not focus on teaching, the implications for teaching are clear: washback often has unintended effects and focusing on what students perceive as necessary to pass a test may be able to positively impact preparation practices.

While using CLT principles for test preparation is not a new idea, *communicative language test preparation* (CLTP) is a term that has rarely, if ever, been used in professional literature. CLTP is rooted in the principles of CLT, a philosophy that places construction of meaning, interaction, and functional use at the center of language learning (Richards & Rogers, 1986), and is based largely on Canale and Swain's (1980) concept of communicative competence.

To clarify, CLTP is not preparation for a kind of communicative language test such as an information gap or role play. Rather, CLTP utilizes activities such as role plays, information gaps, pair work, and group discussions to prepare for non-communicative tests, in this case, the TEM4. One important factor of CLTP is that it does not attempt to supplant an existing system. Rather, CLTP recognizes the importance of working within the existing testing framework while focusing on the four language skills (speaking, listening, reading, and writing) in the preparation process. As with CLT, CLTP attempts to incorporate communicative, interactional activities that create opportunities for students to produce the target language.

In the fall of 2010, the researcher began serving as a Peace Corps volunteer, teaching in a town in the Guizhou Province in southwestern China.The emotional impact of the TEM4 quickly became apparent as students expressed fears and concerns in both written letters and face-to-face conversations. At the beginning of the spring term in 2012, students were asked what they hoped to achieve by the end of the semester. The most common response was to pass the TEM4. From these conversations, a professional, social, and ethical responsibility (Shohamy, 2001) prompted the researcher to prepare students for what they considered to be such an important test. Despite a belief that the TEM4 did not meet Nevo and Shohamy's (1986) *Fairness Standard*, which takes into account the welfare of the test takers, the test was inevitable. Therefore, an attempt was made to create a fair, balanced test preparation course that could benefit students in multiple ways.

There are risks in independently developing an educational program in a foreign context. One risk is taking an Orientalist approach to language teaching (Said, 1978). Said's contention was that the Orient is habitually viewed as inferior to the West, leading Eurocentric nations to justify imposing their beliefs on nations in the East. Another risk, closely related to the first, is Othering students. Othering occurs when groups or individuals create an Us-Them representation of other groups (Palfreyman, 2005). Avoiding the aforementioned dangers requires understanding the teaching context to the best of one's ability. Thus, the current research included an exploration of the educational context. During the exploratory phase, the importance of Confucian ideals in regards to education became clear. Since the time of Confucius, testing in China has been unmistakably important. Tests are often used to determine which schools a student may attend or for which jobs a person is best suited. Because of the importance of tests for students, the researcher further sought todevelop a creative and culturally appropriate means to prepare students for a high-stakes exam.

There is much debate as to whether or not CLT is an appropriate teaching methodology in contexts where language learning has historically involved other forms of instruction (Holliday, 1994; Hu, 2005; Ouyang, 2012). However, as language is a semiotic system, CLT has potential in a wide variety of pedagogical environments. As with the above-mentioned Read and Hayes (2003) study, this research was undertaken to explore the application of CLTP on standardized test preparation. The current study seeks to answer the following questions:

(1) What communicative language teaching methods were implemented during the CLTP course, as documented in the teaching journal?
(2) How did students perceive the CLTP?

(3) What results did the students achieve on the TEM4 after taking the CLTP course?

Methods

Course Design

The CLTP course was designed and implemented by the researcher who, upon arrival in China, had completed the first year of a master's program in TESOL, which included courses in curriculum design, language assessment, language teacher education, and sociolinguistics.

The CLTP course began at the start of the spring semester in 2012, lasted eight weeks until the TEM4 was administered in mid-April, and included three separate sections of students. The course consisted of two parts: regular classroom instruction and weekly evening test practice. The regular classroom instruction time was devoted to improving communicative competence while the evening practice sought to familiarize students with test components.

Nine different CLT tasks were used throughout the course (see Table 1). The first was an interview in which partners took notes and later presented information in front of the class. The next was an ad hoc needs assessment (AHNA) (Shaw, 1982) in which students individually ranked items in a list, conferred with a small group about their list, and finally came to a consensus on the rankings. For the current study, the list items in the AHNA dealt with test-preparation topics, methods, and purposes.

Another CLT task was a jigsaw reading, in which students read news articles for homework. In the subsequent class, students who read the same article reviewed relevant information such as main ideas and key vocabulary. They then divided into "jigsaw" groups made up of students who were experts on their relative articles. In the jigsaw groups, each student explained his or her news story while the other group members took notes on graphic organizers designed for the class.

For a partner dictation, a reading passage that normally would have been read in class was printed and divided in two parts. One part was given to half the class, the other part to the other half. Students were then grouped with partners. Partner A read the passage aloud, including punctuation, while Partner B transcribed the dictation. The pair switched roles and then checked the work together. Partner dictation led to several instances of negotiation for meaning, for instance when students were unclear as to the pronunciation of certain words.

Other communicative tasks included a crossword puzzle, a vocabulary story, and communicative vocabulary review. The crossword puzzle was chosen to focus on geography vocabulary because many of the listening comprehension questions contained material from international current events. Students

had access to world maps and worked together to complete the puzzle. Students also worked together to create a story using vocabulary words learned during the course. Finally, in a group vocabulary review, students used notecards or notebooks to quiz one another on previously learned vocabulary.

The evening test practice included test-taking strategies and provided students with opportunities to complete the TEM4 section by section. Evening sessions usually started with students taking a section of a sample test, which was followed by answers to the sample questions, and concluded with a discussion of the answers. A complete practice test was offered to students three weeks prior to the actual test. The following table illustrates the nine different instances of CLT methods throughout the course.

Table 1

Instances of CLT methods

Week	Journal
1	Interview about holidays
2/20 – 2/24	Ad hoc needs assessment
2	School holiday
2/27 – 3/2	
3	Jigsaw reading
3/5 – 3/9	
4	Vocabulary review
3/12 – 3/16	Partner dictation
5	Crossword puzzle
3/19 – 3/23	
6	Story writing vocab review
3/26 – 3/30	(w/ pictures)
7	School holiday
4/2 – 4/6	
8	Jigsaw reading
4/9 – 4/13	Partner dictation
9	Group vocab review
4/16 – 4/20	

Teaching Diary

During the eight-week CLTP course, a teaching diary was kept as a means to reflect on the course and also to promote professional development. The same lesson was often taught two or three times to different sections, and keeping a diary made it easy to identify opportunities to improve subsequent lessons. It was also hoped that the contents of the diary could provide a framework for curriculum design and implementation if another teacher or other teachers wanted to continue the course at the school in the future.

According to Nunan and Bailey (2009), a diary is a form of introspective data collection, which involves "reporting on one's own thoughts, feelings, motives, reasoning processes, and mental states, often with a view to determining the ways in which these processes and states shape behavior" (p. 285). In the data collection, diary entries were written systematically as soon after the teaching event as possible. The resulting document contained 34 single-spaced, word-processed pages (19,845 words) including descriptions of procedures, feelings about certain events, and interpretations of various occurrences during class time.

The diary was kept for both the regular class meetings and the evening classes and was read repeatedly to identify activities that incorporated macro language skills. If the activities incorporated at least three skills, including speaking and listening, they were then considered examples of CLT and were measured using Liao's (2004) five criteria: (1) focus on communication, (2) interactional activities should be instances of real communication, (3) stress is placed on two-way (or multi-way) communication, (4) students have sufficient exposure to (but are not limited to) L2, and (5) all four language skills are emphasized.

At the end of the course, a *content analysis*, which Freidman (2012) explains, "involves coding data in a systematic way in order to discover patterns and develop well-grounded interpretations" (p. 191), was performed. The first step of the content analysis involved labeling and coding certain actions, events, and topics. The second step was to select the most frequent codes from the initial coding. Types of patterns included repeated key words, connected comments, and metaphors (Nunan & Bailey, 2009, pp. 416-417). Wordle (Appendix A) was also used to help identify word frequency and patterns in the diary. As patterns appeared in the data, the steps outlined by Lieblich, Tuval-Mashiach, and Zilber (1998) were followed in order to put impressions in writing, to decide which patterns or themes to focus on, and to trace the patterns or themes from their beginnings to their ends.

Questionnaire

In order to gauge students' perceived value of the course, a nine-item questionnaire (Appendix B) was created and administered at the end of the course but before the actual test. The questionnaire sought to elicit the students' perceptions of the course. Questions 1, 7, and 8 were in Yes/No format, and requested students to expand on their answers. Question 2 was a Likert-scale question that sought to gauge perceived level of preparation before the exam. Questions 3 and 6 were related to additional preparation students undertook independently. Questions 4 and 5 related to the student's expected performance on the exam. Finally, Question 9 asked students to make suggestions for improving the course in the future.

The questionnaire was originally intended only as a means of action research. The researcher's departure from China following the TEM4 meant that student feedback would not be able to be incorporated into future courses. It was hoped that the students would have found the course valuable for their language development or their preparation for the TEM4 or both. When learning of the questionnaire's possible research value, all students provided their informed consent to be part of the study. In addition, to protect their rights and privacy, no names of students were used.

Participants

Several months after the test was administered, the university received the test scores and sent them, along with scores from the previous year, to the researcher. The scores for both subsets of test takers were input into Microsoft Excel and descriptive statistics including mean, median, mode, and standard deviation were calculated.

It is important to note that English major students in China have two opportunities to take the TEM4. If students score lower than 60% on the exam, they are able to retake the test the following year; the test is only administered once a year. The participants in the current study were 105 second-year English majors in three separate classes. Of the 105 participants in the CLTP course, 102 took the TEM4. All 102 examinees were first-time test takers.

Findings

Communicative Language Teaching

The first research question was answered by analyzing the teaching diary that was kept during the CLTP course. A grounded approach (Linconln & Guba, 1985) to data analysis was employed. In a

grounded approach, categories emerge during the process rather than being determined at the onset. During the initial analysis, the two categories that emerged were communicative and non-communicative methods. During subsequent readings of the journal, two subcategories of communicative methods emerged: traditional and adaptive.

Questionnaire Responses

Answering the second research question required reviewing student responses to the questionnaire (see Appendix B) that was administered at the end of the CLTP course. Unfortunately, many problems were discovered while analyzing the questionnaire results. These problems, enumerated in the Discussion section, rendered much of the data unsuitable for analysis.

Due to the possible misunderstandings in completing the questionnaire, only the results of the first two questions are presented below because they are the two items most closely related to the research question and are the two least likely to have been misinterpreted. The first question asked, "Do you think your English has improved as a result of our test preparation course? Explain." There was an overwhelmingly positive response to this question; of the 105 respondents, 103 reported that their English improved because of the test preparation. When asked to explain, the most frequent answer was "dictation" (n=19). Other notable answers for explaining the perceived improvements in English were, "test skills," (n=8), "listening," (n=5,) and "confidence," (n=3).

The second question asked students to rate their level of perceived preparedness on a scale from 1 to 10, with 1 being "very unprepared" and 10 being "very well prepared."The responses were recorded separately for each of the three sections taught. The average level of perceived preparation was 5.39, 5.34, and 5.64, respectively.

Tests Scores

In 2011, 105 students from the university involved in the current study took the TEM4. The average score was 43.19 with a standard deviation of 6.69. The high score was 65, the low score 30, and four out of the 105 received a passing score of 60 or higher. In 2012, 102 students from the same university took the TEM4. The average score was 45.19 with a standard deviation of 7.47. The high score was 73, the low score 29, and six out of the 102 received a passing score. Table 2 lists the descriptive statistics for the student test scores both years.

Table 2

Descriptive statistics from 1ˢᵗ time test takers in 2011 and 2012

	2011	2012
Number	105	102
Mean	43.190	45.656
Median	42	45
Mode	40	44
High	65	73
Low	30	29
Passing	4	6
Std Dev	6.699	7.472

Maximum score = 100

Discussion

Innovation

When it was first introduced, CLT was an innovation in language teaching methodology. Markee (1997) provided one definition of innovation as it applies to language teaching curricula: *innovation* is "a managed process of development whose principal products are teaching (and/or testing) materials, methodological skills, and pedagogical values that are perceived as new by potential adopters" (p. 46). Much time has passed since the introduction of CLT into China. By Markee's definition, CLT is no longer considered an innovation. Nevertheless, every educational context is unique, which allows CLT to be applied in creative and innovative ways.

Although innovations are often viewed as a positive force in language education, there are dangers associated with certain innovations. According to Holliday (1994) innovations are inherently ethnocentric, especially when they are implemented without regard for the norms of the group that will adopt them. However, not all innovations are alike. Some innovations, such as CLT when it was introduced in China, are implemented from the top-down. With top-down innovations, there is typically minimal consensus from those who would eventually use the innovation. On the other hand, those who recognize a need in the existing system initiate other innovations at a more grass-roots level. As with many dichotomies in language teaching, innovation exists on a continuum. The test preparation course in the current study was designed predominantly from the bottom-up in the sense that course development was not a part of the administrative decision-making process at the university. Similarly, the content of the course was based on the perceived needs of the test takers, the consumers of the

innovation. In one respect, the course could be considered top down because it occurred during required, self-study evening classes. However, students were not required to take part in the CLTP course, even though everyone elected to do so.

Decisions about CLT

Much of the literature on CLT in China shows a divided perspective. Researchers such as Liao (2000, 2004) contend that CLT is the best option for language teaching in China at every level and in every context. Others such as Ouyang (2012) state that CLT has potential benefits in China, but must take into consideration "cultural contexts of an extensive power distance, hierarchically granted voices, identities as members of a collective, and their norms of interaction in a tacit consensus over long term consequence and hence internalization" (p. 282). In other words, it is important to note that cultural factors in the classroom and in the greater macro-environment may impact the effectiveness of an approach such as CLT.

CLT, as with any educational philosophy, carries a view of language teaching that may not be compatible in certain social contexts (Holliday, 1994). For example, Holliday's notion of the *learning group ideal* may clash with beliefs about student-teacher roles, autonomy, and identity, and may be resisted, or in some cases refused outright (e.g., Shamim, 1996). For example, students who are required to speak in front of classmates from whom they may feel socially distant run the risk of losing face and feeling humiliated (Ouyang, 2012). Similarly, experiences teaching at the university level in rural China revealed that students themselves were largely responsible for preparing for standardized tests and received little support from teachers or the administration. Due to the lack of guidance, students often deferred to the forms of study and preparation most comfortable to them, including memorization, repetition, and direct translation. Thus, the administrative, pedagogical, and social contexts were considered as carefully as possible in planning and implementing the test preparation course.

One important consideration was the nature of CLT as opposed to the nature of the TEM4. The latter focuses on mastery of discrete skills. Decoding sounds is measured by way of dictation, and reading comprehension is measured using multiple-choice questions. CLT, on the other hand, focuses on the integration of language skills, a reflection of the semiotic and dialogic nature of language. This consideration led to the conclusion that even if test scores did not increase, students might benefit from opportunities to communicate in English, both with the researcher and fellow classmates.

Teaching Diary Analysis

According to Bailey and Oschner (1983, p. 189), a teaching diary is a first-person, introspective account of language teaching that can describe factors such as emotions, strategies, and perceptions that are not normally accessible to outside observers. Diaries may be used for a variety of purposes, including professional development (Bailey, Curtis, &Nunan, 2001) and reflective teaching (Bailey, 2001). The diary kept was a combination of both professional development and reflective teaching, but was mainly a means of recording data to learn, develop, and improve.

The process of keeping a teaching diary allowed for reflection on the creation and progress of the course, and the diary itself facilitated action research and shed light on thought processes and decisions, such as what materials to use and ways to modify teaching methods or delivery, that were made during planning and instructional time. The diary also served as the basis for reflection on the social responsibility of a language tester (Shohamy, 2001) and the appropriateness of CLT in a context that has traditionally relied on other teaching methodologies.

The test preparation course was designed to focus on communication and therefore examples of CLT as defined by Liao (2000) were incorporated into the course. Of the nine instances of CLT that occurred during the course, all of them occurred during regularly scheduled classes (see Table 3). While analyzing the teaching diary, some unexpected patterns emerged from the examples of CLT. One pattern was that activities initially considered to be communicative did not meet one or more of the criteria of CLT. For example, one subtest of the TEM4 listening comprehension section requires students to listen to a brief dialogue then answer comprehension questions. In week two, a practice dialogue was transcribed onto a PowerPoint slide, and students read the slide before listening to an audio of the dialogue. This activity required speaking, listening, and reading, but was not bi- or multi-directional, and therefore was not considered an instance of CLT.

Another observation from the teaching diary was that the instances of CLT fell into two categories: traditional and adaptive. Traditional forms of CLT often occurred as structured pair work, including comparing answers to comprehension questions; group work, such as jigsaw readings; and class discussions, including feelings about certain topics such as global warming and illiteracy. Adaptive forms of CLT were the result of an attempt to make traditionally non-communicative preparation activities more communicative. For instance, one subtest of the TEM4 listening comprehension section was dictation. During the first several weeks of class, sample dictation passages were read while students listened and transcribed. Attempts to make every activity more communicative eventually led

to the implementation of partner dictation (Smith, 2010). Partner dictation changed the nature of the activity by incorporating reading and speaking into an activity that had previously only focused on listening and writing. It also changed the traditionally one-directional teacher-led dictation task into a bi-directional student-centered task.

Of the nine instances of CLT tasks, seven were traditional and three were adaptive. One instance of CLT methodology, a vocabulary review, began as a more traditional activity, but was modified after teaching the lesson once so that the subsequent activity was an adaptive form of CLT. Table 3 shows all instances of traditional and adaptive CLT methods that occurred during instructional time.

Table 3

Traditional and adaptive CLT methods

Week	Journal	Type	Skills
1 2/20 – 2/24	Interview about holidays Ad hoc needs assessment	Traditional Traditional	S, L, W S, L, R, W
2 2/27 – 3/2	School holiday	NA	NA
3 3/5 – 3/9	Jigsaw reading	Traditional	S, L, R, (W)
4 3/12 – 3/16	Vocabulary review Partner dictation	Traditional Adaptive	S, L, R
5 3/19 – 3/23	Crossword puzzle	Adaptive	S, R, W
6 3/26 – 3/30	Story writing vocab review (w/ pictures)	Traditional (Adaptive)	S, R, W
7 4/2 – 4/6	School holiday	NA	NA
8 4/9 – 4/13	Jigsaw reading Partner dictation	Traditional Adaptive	S, L, R, W S, L, R, W
9 4/16 – 4/20	Group vocab review	Traditional	S, L, R

Questionnaire Responses

As mentioned previously, many discrepancies with the intended responses and the actual responses were found in the questionnaire. For example, in Questions 3 and 4, the word "expect" translates into

Chinese as *xi wang* (希望). When translating from Chinese to English *xi wang* not only means "to expect," but also "to hope for." Therefore, one interpretation of Question 4 might have been "On which section of the TEM4 do you *hope* to get the lowest score?" Question 7 had a similar problem with the word "regular." "Regular" was intended to mean "during the day," but this meaning was not explained to students. As a result, some students commented on classes taught by other teachers.

Although there were numerous technical issues with its design, the questionnaire allowed students to reflect on the CLTP course as a whole and provided helpful, albeit limited, insight into their perceptions and feelings. Students overwhelmingly felt that the CLTP course helped to improve their English. When asked to explain their answers, the most frequent student response was "dictation," presumably meaning that dictation helped improve English proficiency or ability. Despite the sense of improved English ability, the questionnaire revealed that students did not feel well prepared to take the TEM4. On a scale of one to ten, the average level of preparation was 5.34, 5.39, and 5.64 for the three sections, or an average 5.46. Interestingly, these numbers are very near the pre-determined passing score (60), a score that students are quite aware of. That the perceived level of preparation is very near the passing score possibly reflects social and cultural factors related to the level of difficulty of the test and are not necessarily indicative of the students' chances for success.

Other than aforementioned ambiguous words in English, the questionnaire had other limitations. First, no measures were taken to ensure the trustworthiness and consistency of the questionnaire and its items. Neither was the questionnaire piloted nor were colleagues asked to provide feedback on the questionnaire. Measuring the level of preparation students perceived prior to and after the CLTP course by means of a preliminary questionnaire would also have provided additional insight into the students' perceptions.

Test Scores

Based on the relatively low mean scores, it is reasonable to state that the test was too difficult for most students. However, looking past the difficulty level of the test, some differences stand out when comparing the scores for the two years. First of all, nearly every statistical category experienced positive increases. The average test score increased by approximately 2.5 points, the mean score increased by 3, the mode increased by 4, and the high score increased by 8 out of 100 possible points. The only negative change was in the low score, which decreased by 1 point. It is also worth noting that six students passed the test in 2012. While this number may seem low considering that 102 students

took the exam, it is a 50% increase from the previous year, and according to the English department, is the largest number of first-time test takers ever to pass the exam.

The nature of the research design makes it difficult to establish causal relationship between the CLTP and the increase in student test scores. The pre-experimental design left many variables unaccounted for. Nevertheless, the increases in nearly all statistics measured merits attention. While further research is required to confirm causality, the finding that focusing on communicative competence and language ability has a positive impact on both language ability and test scores is consistent with the findings by Read and Hayes (2003), mentioned above.

Limitations and Future Research

One limitation of this study related to cultural resources that could be drawn on in the creation of the CLTP course. Despite best efforts to construct a culturally appropriate innovation, it is possible that attempts to understand the micro and macro environments in which the course was developed and implemented were incomplete. It is also possible that certain elements in the course had the effect of Othering students in some way (Palfreyman, 2005). To address this problem, a diary kept for future studies should reflect on not only what happens inside the classroom, but also on macro-environmental factors including the relationship of teaching practice to institutional and societal norms, including expectations of passing the exam. To accomplish this broader scope, diary entries could include data from interviews with students to help delve deeper into feelings about the preparation course.

There were also limitations to the quantitative analysis. For example, the research was only conducted one time in one location. Increasing the sample size and number of sites would allow for more robust claims to be made about the effects of the course on test scores. Similarly, longitudinal studies could look at the impact of CLTP on students who did not pass the TEM4 the first time. It is possible that the impact of CLTP is more evident on test takers who take the TEM4 a second time. Thus, analyzing students' scores after their second attempt may prove useful. Additionally, the present study was primarily qualitative in nature and quantitative data were analyzed using an ex post facto design. While ex post facto design is not a limitation in and of itself, in the current study, the nature of the design did not allow for any controls to be implemented.

Further research could help gain support for language-focused test preparation where high stakes language tests are non-communicative in nature. As explained in the introduction, testing plays a critical role in Chinese society. Reforming the language policy in China simply requires making a

decree from the highest level of the government. However, the perceived need to adhere to the national policy varies dramatically from one educational context to the next in China. Changing the Chinese system to a more communicative one may be attractive, especially to parties who have a financial or political interest in such a change, but not all students, teachers, or administrators see communicative language teaching as important or necessary.

One final suggestion for future research is to expand the focus of the effects of the CLTP course. The current study looked primarily at test results, but did not explore the effects of CLT on English proficiency. Creating opportunities to interact in the English does not guarantee improved language ability, and because communicative competence is a goal of both CLT and CLTP, a broader focus could illuminate advantages and disadvantages beyond language testing.

Conclusion

This study investigated the teaching diary kept during the spring semester of 2012. The diary proved to be a valuable tool in accomplishing the goal of preparing students for the TEM4. CLT was used in the design of the course, and while CLT may not be an ideal methodology in every social and cultural context, it does contain the underpinnings of a flexible, adaptable methodology. Despite the somewhat ethnocentric nature of the initial CLT implementation in China, the researcher feels that the current study may be applicable to many university classrooms throughout the county.

Participant test results revealed that the TEM4 is too difficult for many students. Although the number of passing scores was low, the number of passing scores represented a 50% increase from the previous year, as well as the largest number of students from the university ever to pass the exam. Also, when compared to test scores from the previous year, there was a significant increase of average test scores, indicating the likelihood that the CLTP course had a positive impact on the students' overall test preparation. The fact that all but one statistical category positively increased from one year to the next also makes it likely that the CLTP course had an effect on improving scores for students, meaning that focusing on communicative competence and language ability while at the same time preparing for exams has more positive results than focusing on test taking ability alone.

Other than gains in test scores, the CLTP course provided students experience with language tasks similar to those that they might encounter in dialogue outside of the classroom. Using CLT to prepare for the TEM4 provided opportunities for meaningful interaction, critical reflection, and noticing in the L2 that would not have been available using traditional practice tests or teacher centered methods.

For better or worse, high-stakes testing plays a vital role in the Chinese education system. Students prepare for these tests, including language tests, to the best of their ability, but memorization of information, an integral part of Confucianism, makes it unlikely that students will communicate with one another as part of the preparation process. However, the fact cannot be overlooked that language is indeed a means for interpersonal and other forms of communication, and finding innovative, communicative means of preparing students to succeed provides simultaneous support for both language and test-taking skills.

References

Alderson, J. C., & Hamp-Lyons, L. (1996). TOEFL preparation courses: A study of washback. *Language Testing, 13*, 280-297.

Bailey, K. M. (2001). What my EFL students taught me. *PAC Journal, 1*(1), 7-31.

Bailey, K. M., Curtis, A., & Nunan, D. (2001).*Pursuing professional development: The self as source*. Boston, MA: Heinle.

Bailey, K. M., & Ochsner, R. (1983). A methodological review of the diary studies: Windmill tilting or social science? In K. M. Bailey, M. H. Long, & J. Peck (Eds.), *Second language acquisition studies* (pp. 188-198). Rowley, MA: Newbury House.

Canale, M., & Swain, M. (1980). Theoretical bases of communicative approaches to second language teaching and testing. *Applied Linguistics, 1*, 1–47.

Cheng, L. (2008). The key to success: English language testing in China. *Language Testing, 28*(1), 15-37.

Deng, C., & Carless, D. R. (2010). Exam preparation or effective teaching: Conflicting priorities in the implementation of pedagogic innovation. *Language Assessment Quarterly, 7*(4), 285-302.

Holliday, A. (1994). *Appropriate methodology and social context*. Cambridge, England: Cambridge University Press.

Hu, G. (2005). 'CLT is best for China' – an untenable, absolutist claim. *ELT Journal, 59*(1), 65-68.

Kitao, S., & Kitao, K. (1996).Testing communicative competence. *The Internet TESL Journal, 2*(5). Retrieved from http://iteslj.org/Articles/Kitao-Testing.html

Liao, X. (2000). Communicative language teaching innovation in China: Difficulties and solutions. (ERIC Document Reproduction Service No. 443 294).

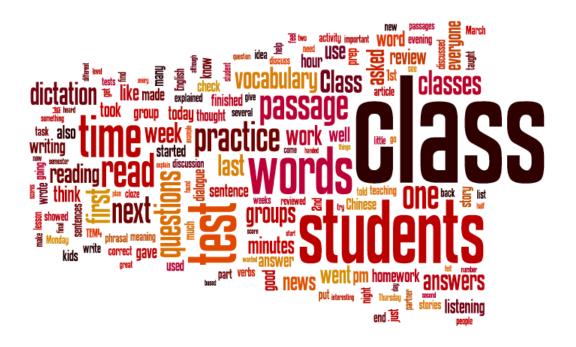Liao, X. (2004).The need for communicative language teaching in China. *ELT Journal, 58*(3), 270-273.

Lieblich, A., Tuval-Maschiach, R., & Zilber, T. (1998). *Narrative research: Reading, analysis, and interpretation*. Thousand Oaks, CA: Sage.

Markee, N. (1997). *Managing curricular innovation*. Cambridge, England: Cambridge.

McNamara, T. (2001). Language assessment as social practice: Challenges for research. *Language Testing, 18*(4), 333-349.

National Council on Measurement in Education. (2012). *Test preparation*. Retrieved from http://ncme.org/resource-center/glossary/t/test-preparation/

Nevo, D., & Shohamy, E. (1986). Evaluation standards for the assessment of alternative testing methods: An application. *Studies in Education Evaluation, 12*(1), 149-158.

Nunan, D., & Bailey, K. M. (2009). *Exploring second language classroom research: A comprehensive guide*. Boston, MA: Heinle.

Ouyang, H. (2012). Dominant pedagogical approaches and diverse teaching conditions: Integrating CLT in a Chinese university as a *danwei* community of practices. *Applied Linguistics Review, 3*(2), 273-293.

Palfreyman, D. (2005). Othering in an English language program. *TESOL Quarterly, 39*(2), 211-233.

Qiang, N., Wolff, M., Hai, T., & Gregory, A-M. (2004). Can you get a first class education at a third tier college in China? *Asian EFL Journal, 6*(1). Retrieved from http://www.asian-efl-journal.com/04_nq_mw.php

Raimes, A. (1990). The TOEFL test of written English: Causes for concern. *TESOL Quarterly, 24*(3), 427-442.

Read, J., & Hayes, B. (2004). The impact of IELTS on preparation for academic study in New Zealand. *IELTS Research Reports, 4*, 153-191.

Richards, J. C., & Rogers, T. S. (1986). *Approaches and methods in language teaching: A description and analysis*. Cambridge, England: Cambridge University Press.

Richards, J. C., & Schmidt, R. (2010). *Longman dictionary of language teaching & applied linguistics* (4th ed.). Harlow, England: Pearson.

Said, E. W. (1978). *Orientalism*. New York, NY: Pantheon Books.

Scouller, K. (1998). The influence of assessment method on students' learning approaches: Multiple choice question examination versus assignment essay. *Higher Education, 35*(4), 453-472.

Shamim, F. (1996). Learner resistance to innovation in classroom methodology. In H. Coleman (Ed.), *Society and the language classroom.* (pp. 105-121). Cambridge, England: Cambridge University Press.

Shaw, P. A. (1982). "Ad Hoc needs analysis." *Modern English Teacher, 10*(1), 12-14.

Shohamy, E. (2001). The social responsibility of the language tester. In R.L Cooper, E. Shohamy, & J. Walters (Eds.), *New perspectives and issues in educational language policy* (pp. 113-130). Amsterdam, Netherlands. John Benjamins.

Smith, C. D. (2010). Pair-dictation activities for teaching English to university students. *The Internet TESL Journal, 16*(2). Retrieved from http://iteslj.org/Techniques/Smith-Dictation.html

State Education Development Commission of China (1992). *Syllabus of teaching English for compulsory education in junior school of nine-year system.* Beijing, China: People's Education Press.

Xie, Q., & Andrews, S. (2013). Do test design and uses influence test preparation? Testing a model of washback with Structural Equation Modeling. *Language Testing, 30*(1), 49-70.

Zhan, Y. (2008). Washback on Chinese learners: An impact study of the College English Test Band 4. Retrieved from http://www.iaea.info/papers.aspx?id=71

Appendix A

Teaching diary word salience

Representation by Wordle.com

Appendix B

Course Evaluation

1. Do you think your English has improved as a result of our test preparation? Explain
2. On a scale from 1 to 10, how well prepared do you feel for the TEM4?

very unprepared   1-----2-----3-----4----5-----6-----7-----8-----9-----10   very well prepared

3. How many complete practice tests have you taken?
4. On which part of the TEM4 do you expect the highest score? Why?
5. On which part of the TEM4 do you expect the lowest score? Why?
6. About how many times did you practice each section? Do you wish you had done more?
   Dictation –
   Listening Comprehension –
   Cloze –
   Grammar and Vocabulary –
   Reading Comprehension –
   Writing –
7. Do you feel our regular classes helped you prepare for the exam? Explain
8. Do you feel our night classes helped you prepare for the exam? Explain
9. If you were the teacher, what would you have done differently during our classes?

Language Curriculum Design and Socialisation.

Peter Mickan. Bristol: Multilingual Matters, 2013. Pp. xvi + 133.

Reviewed by Yuka Akiyama, Georgetown University, Washington DC.

With globalization and the expansion of social interaction across borders, educational policy makers as well as language educators have noticed an urgent need for the transformation of a traditional curriculum into a curriculum that enables language users to socialize with community members who share a specific discourse. In *Language Curriculum Design and Socialisation* (2013), Peter Mickan proposes a new way to design a curriculum that views language as "a meaning-making resource in human activity rather than presenting language as grammatical and lexical objects" (p. xiii). The book is written for curriculum developers, language teachers, teacher trainers, and student teachers who want to go beyond Communicative Language Teaching (CLT), currently the dominant approach whose syllabus design, despite its communicative interests, tends to separate language from its focus on communication (Mickan, Chapter 2). Mickan's fundamental argument, which is based on Halliday's Systemic Functional Linguistics is that texts should be used as the basis of curriculum design because texts (1) encompass daily language use, (2) contextualize language used for social purposes, and (3) are familiar to learners and readily available to teachers in all contexts and situations. The book consists of nine chapters, each of which ends with a summary, notes, extra suggested readings, and a set of tasks, making this a potentially useful textbook for second language teacher education courses.

Chapter 1: "Texts in the Fabric of Life" sets out the notion that humans live with a myriad of texts. The chapter then delves into the rationale behind why texts should be the foundation of curriculum design. The author argues that the patterned nature of language, as realized in the "ubiquity, propinquity, utility, and significance" (p. 4) of texts, makes texts familiar units for the design of curricula and useful organizers for teaching activities.

Chapter 2: "Change and Renewal in Curriculum Design" starts with a historical overview of language teaching approaches that have influenced curriculum design

Address correspondence:
Department of Linguistics, Georgetown University, Box 571051, Poulton Hall 240, 1421 37th Street., N.W.,Washington D.C. 20057-1051

for the last 50 years. Mickan lists problems with each of these approaches, especially focusing on the challenges in integrating decontextualized grammar into a communication-based curriculum that inconsistently includes speech acts, functions and notions, situational dialogues, communicative activities, tasks, and genres. He claims that the use of texts, which are already cohesive and contextualized by nature, avoids the separation of language from social practices and allows language to be analyzed as useful and authentic units of meaning.

In Chapter 3: "Learning the Language of Social Practices," Mickan first defines language learning as a process of socialization and regards language development as a shift in membership from peripheral to fuller participation. The chapter then briefly talks about how to organize a language curriculum based on social practices and texts. For instance, the author places importance on engaging in authentic texts such as "shared reading, reading for pleasure, solving problems, debating ideas, composing arguments and creating objects collaboratively" (p. 37-38) for socialisation to take place in language instruction.

The author then moves on to talk about curriculum design and planning. Chapter 4 explains how to situate five core components of a curriculum (i.e., a syllabus, a lesson plans, resources, teaching activities, and assessment) in social practices while Chapter 5 expands on the previous chapter by describing a seven-step procedure for planning a socially-oriented curriculum.

In Chapter 6: "Teaching Practices," Mickan moves on to the pedagogical aspect of the curriculum design and outlines eight proposals for text-based teaching practices such as the appreciation of prior knowledge, explanation of lexico-grammatical selections and meanings, and assessment of social participation. He claims that his proposed teaching approach enables learners to become aware of the function and social practices of the text.

Chapter 7: "Curriculum Applications" perhaps is the most practical chapter of the volume that exemplifies how to apply social theory to language curriculum design in various language pedagogy contexts. The chapter depicts text-based curricula in Australia such as content-based instruction for ESL students as well as an extensive reading program in Korea. These examples will likely allow readers to gain practical insights and apply them to their own teaching.

Chapter 8: "Curriculum Design in Higher Education" expands on the previous chapters by discussing how to apply social theory beyond K-12 education. The author proposes that language instruction make explicit discipline-specific discourse, so students can see the defining features of academic literacies and develop disciplinary knowledge.

The final chapter of the volume "Language Planning, Curriculum Renewal and the Teacher as Researcher" again makes the case for text-based curriculum and emphasizes the importance of conducting action research to maintain socially-oriented curricula. The chapter discusses language planning from three perspectives: language policy and curriculum design, program structure, and teacher education and professional development, providing an overall picture of the inextricable relationship between language policies and curriculum design.

Overall, although this book alone may not be as sufficient as to enable readers to design a text-based curriculum, this volume provides sound theoretical insights into how social theory informs language curriculum design and can be credited as one of few recently published books that have shed light on both the theoretical and practical aspects of text-based curriculum design (see Crombie, 1985; Unsworth, 2001 for other examples). More readers would appreciate this book if more practical examples were included throughout the volume; however, language teachers can certainly make use of this collection as the point of departure for understanding a text-based language curriculum.

References

Crombie, W. (1985). *Discourse and language learning: A relational approach to syllabus design*. Oxford: Oxford University Press.

Unsworth, L. (2001). *Teaching multiliteracies across the curriculum: Changing contexts of text and image in classroom practice.* Philadelphia, PA: Open University Press.

The submission has not been previously published or is not being considered for publication elsewhere.

Yuka Akiyama is a Ph.D. student in Linguistics at Georgetown University. She holds a Master of Education in TESOL from Boston University. Her research interests include social interaction, corrective feedback, task-based language teaching, and teacher training in the context of telecollaboration/virtual language exchange. Her previous publications include studies on the effect of dynamic written corrective feedback and the effect of lexical categories on task-based telecollaborative interaction, a position paper on the use of telecollaboration for faculty development, and an assessment tool for Japanese language educators.

Exploring language pedagogy through second language acquisition research

Rod Ellis and Natsuko Shintani, Routledge: London/New York, 2014. Pp. ix+388.

Reviewed by Yingli Yang, University of International Business and Economics, Beijing, P.R. China

Second language acquisition (SLA) has developed into a multidisciplinary field in recent years. Exploring language pedagogy through second language research is a timely work. It is a must-have for educators and researchers who will benefit from its up-to-date and comprehensive references on research methods, theoretical framework, and pedagogical guidance.

This book comprises 12 chapters, which are divided into five parts. The purpose of the book is to establish a link between 'pedagogical discourse' and 'research-based discourse' (p. 2) and inform language practitioners of the pedagogical choices they can make in consideration of SLA theories and research results. Chapter 2 examines the 'method construct' and related SLA theories. Postmodernist pedagogy argues that teachers and learners are autonomous entities and therefore should dispense with the method construct. The authors conclude that the method construct still has its value in guiding classroom practice, but more rigorous observational or experimental studies are needed in order to evaluate the effects of one method over another.

Chapter 3 introduces three types of linguistic syllabuses. As the name suggests, the grammatical syllabus places grammar at the center of a language curriculum. However, this leads to the caveat that language meaning, which is also critical to the process of communication, is excluded. The lexical syllabus is introduced with the development of corpus linguistics. Yet such a kind of syllabus cannot ensure the mastery of enough lexical items, nor can this syllabus evolve grammatical rules and patterns that formulate correct sentences. Finally, although the notional syllabus is more advanced than the lexical syllabus, as the authors argue, it still fails to incorporate rule-based learning to achieve "the full development of competence in an L2" (p.79).

Address correspondence:
University of International Business and
Economics, Room 1312, Chengxin
Building, No. 10, Huixin Dongjie,
Chaoyang District, Beijing, China,
100029

While chapters 2 and 3 deal with the 'design' of a language course, chapters 4 and 5 address the implementation of the syllabus. Contrary to the non-interface position proposed by Krashen (1981), the empirical evidence presented in chapter 4 supports that explicit instruction assists learners' development of explicit knowledge. Additionally, together with ample opportunities for practice, explicit instruction can also facilitate the development of implicit knowledge.

Chapter 6 reviews theories and issues in task-based language teaching, including the definition of tasks, task types, task-based syllabuses, and cognitive theories related to task-based teaching, as well as the implementation of tasks and empirical studies. Finally, the authors suggest that in order to implement task-based teaching successfully, performance-based testing must be introduced accordingly.

Part III discusses the internal perspective of SLA by viewing the process of SLA as a process of communication or 'interaction'. Chapters 9 and 10 examine two arguments that are prominent in the SLA field, i.e., issues regarding the value of the use of L1 and the effectiveness of feedback in SLA. In Chapter 9, by reviewing literature on code-switching, the authors conclude that L1 is facilitative of L2 learning since it enables learners to "overcome communication problems" (p.245) and to "scaffold their production of L2" (p.243); however, due to the paucity of research that measures the direct effect of the use of L1 on L2 learning, it is also advisable for teachers to balance their L2 input with the use of L1. In Chapter 10, on the basis of empirical research conducted during the past three decades, the authors revisit the five central questions raised by Hendrickson (1978) regarding corrective feedback. Building on theories and empirical studies on corrective feedback, it was suggested that both oral and written feedback are facilitative in second language acquisition, particularly in improving learners' accuracy.

While Part II and Part III take a holistic approach towards second language teaching and discuss the general situations in second language teaching contexts, Part IV zooms in to look at individual differences of the learners during their learning process. By reviewing important components of individual differences in SLA literature such as age, language aptitude, motivation, and learning strategies, the authors offer practical suggestions for language teachers. There are two ways to accommodate learners with different kinds of language aptitude. One is through matching the learners' aptitude with the appropriate teaching method, the other is by offering "an eclectic selection of activities", which may meet diverse needs of all the learners in the classroom (p. 312). In terms of motivation, the authors suggested that teaching practitioners should actively engage students to maintain a high level of motivation. Part V reiterates the importance of linking SLA research with pedagogy and summarizes the contents presented in previous chapters.

The highlight of this book is that it combines a research perspective with well-grounded pedagogical advice. The authors' abounding teaching experience in various contexts offers a unique and balanced view of second language acquisition. Moreover, researchers and teaching practitioners will benefit from the extensive literature, research design, and multi-faceted perspectives introduced in the book. The only part that might challenge novice teachers is the terminology occurring in chapters 4 and 5. Fortunately, Table 5.1 deftly summarizes these otherwise confusing concepts. Overall, this book is a remarkable contribution to the field of SLA. With its well-organized structure and accessible language, it will be a valuable resource for researchers and teaching practitioners alike.

References

Hendrickson, J. (1978). Error correction in foreign language teaching: recent theory, research, and practice. *Modern Language Journal, 62*(8), 387-398.

Krashen, S. (1981). *Second language acquisition and second language learning*. Oxford: Pergamon.

The submission of this book review has not been previously published or is not being considered for publication elsewhere.

Dr. Yingli Yang is an associate professor at the school of international studies, University of International Business and Economics, China. She holds a Ph.D Degree in Second Language Education from McGill University, Canada. Her publications appear in Studies in Second Language Acquisition, Journal of Pragmatics, Journal of Asia TEFL, and Asia ESP journal. Her research interests include second language acquisition, corrective feedback, and business English.