

**The Asian EFL Journal Quarterly**  
**March 2008**  
**Volume 10, Issue 1**



**Senior Editors: Paul Robertson and Roger Nunn**



Published by the Asian EFL Journal Press

Asian EFL Journal Press  
A Division of Time Taylor International Ltd  
Time Taylor College  
Daen dong  
Busan, Korea

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

©Asian EFL Journal Press 2008

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

**No unauthorized photocopying**

All rights reserved. No part of this book may be reproduced, stored  
in a retrieval system or transmitted in any form or by any means,  
electronic, mechanical, photocopying or otherwise, without the prior  
written permission of the Asian EFL Journal.  
editor@asian-efl-journal.com

Publisher: Dr. Paul Robertson  
Senior Associate Editor: Dr. Roger Nunn  
Associate Editor, Journal Production Editor: Darren Lingley

ISSN 1738-1460



**Table of Contents:**

<b>Foreword by Roger Nunn.....</b>	<b>4-6</b>
<b>1. Kyoung Rang Lee and Rebecca Oxford .....</b>	<b>7-32</b>
- <i>Understanding EFL Learners' Strategy Use and Strategy Awareness</i>	
<b>2. Mohammad Rahimi.....</b>	<b>33-47</b>
- <i>Using Dictation to Improve Language Proficiency</i>	
<b>3. Souvannasy Bouangeune, Masashi Sakigawa and Yukiko Hirakawa..</b>	<b>48-64</b>
- <i>Determinants and Issues in Student Achievement in English at Lao Secondary Education Level</i>	
<b>4. Junko Otoshi and Neil Heffernan.....</b>	<b>65-78</b>
- <i>Factors Predicting Effective Oral Presentations in EFL Classrooms</i>	
<b>5. Chi Yen Chiu.....</b>	<b>79-110</b>
- <i>The Discourse of an English Teacher in a Cyber Writing Course: Roles and Autonomy</i>	
<b>6. Nguyen Thi Mai Hoa.....</b>	<b>111-132</b>
- <i>Mentoring beginning EFL teachers at tertiary level in Vietnam</i>	
<b>7. Zhu Xinhua.....</b>	<b>133-153</b>
- <i>Is Syntactic Maturity a Reliable Measurement to Investigate the Relationship between English Speaking and Writing?</i>	
<b>8. Mahmood Rouhani.....</b>	<b>154-180</b>
- <i>Another look at the C-Test: A Validation Study with Iranian EFL Learners</i>	
<b>9. Jeng-yih Tim Hsu and Chu-yao Chiu.....</b>	<b>181-204</b>
- <i>Lexical Collocations and their Relation to Speaking Proficiency</i>	
<b>10. Lu-Fang Lin.....</b>	<b>205-227</b>
- <i>The Study of English Learners' Synthesizing Process While Reading</i>	
<b>11. Nilton Hitotuzi.....</b>	<b>228-258</b>
- <i>An Economical Approach towards Interaction in the L2 Classroom: A Task-based Learning Experiment</i>	
<b>12. Daniela Nikolova.....</b>	<b>259-274</b>
- <i>English-teaching in Elementary Schools in Japan: A Review of a Current Government Survey</i>	

**Book Reviews**

<b>1. Assessing Language through Computer Technology</b> Carol A. Chapelle and Dan Douglas Reviewed by <b>Deepti Gupta.....</b>	<b>275-276</b>
<b>2. Literature and Stylistics for Language Learners: Theory and Practice</b> Greg Watson and Sonia Zyngier (Eds.) Reviewed by <b>Vander Viana.....</b>	<b>277-278</b>
<b>3. Practical English Language Teaching: Speaking</b> Kathleen M. Bailey Reviewed by <b>Aysha Viswamohan and Hannah M. Sanala.....</b>	<b>279-281</b>
<b>4. Culturally Contested Pedagogy: Battles of Literacy and Schooling between Mainstream Teachers and Asian Immigrant Parents</b> Guofang Li Reviewed by <b>Ozgur Yildirim.....</b>	<b>282-283</b>

<i>Asian EFL Journal</i> editorial information and guidelines.....	<b>284-290</b>
--	----------------



## Foreword

In this first issue of 2008 we are delighted to be able to present such a wide variety of studies conducted by Asian researchers. This issue marks the tenth anniversary of the Asian EFL Journal. We are both proud of our growth as an academic forum for teachers and researchers, and grateful for the hard work and dedication of our reviewers and editorial team.

Lee and Oxford, in *Understanding EFL Learners' Strategy Use and Strategy Awareness*, focus on the influence of strategy awareness, English-learning self-image and the importance of English in the Korean context. They found that students who valued English as important, evaluated their own proficiency as high and were already aware of a variety of language learning strategies employed learning strategies more frequently than those who did not. While gender and major were expected to be helpful indicators of successful learning, Lee and Oxford found that they did not affect strategy use and awareness unless combined with other variables, concluding that emphasizing strategy use based on gender or majors could be promoting a stereotype.

In *Using Dictation to Improve Language Proficiency*, Mohammad Rahimi revisits an old debate about the value of dictation as both a teaching and a testing tool. Rahimi focuses on the former. Rahimi concludes that dictation can be considered a good learning technique to improve learners' proficiency. In the study, improvement was recorded in grammar, vocabulary, reading, and listening comprehension, while the control group showed improvement only in vocabulary. As Rahimi points out, improvements in proficiency are difficult to pin down to a single cause and the period of incubation needs careful consideration. The fact that the post-test was done after around ten months leads to different possible interpretations.

In their study, Bouangeune, Sakigawa and Hirakawa look at the challenges facing an English education system which is in its infancy. The research for *Determinants and issues in student achievement in English at Lao Secondary Education Level* was located in Laos and investigates the proficiency level of Lao secondary school students. They identify a need to focus more on the learning of basic vocabulary, and to simplify the textbooks.

Otoshi and Heffernan, in *Factors Predicting Effective Oral Presentations in EFL Classrooms* discuss the criteria EFL learners consider to be important when making presentations. They provide a useful checklist as a basis for similar studies. This paper

underlines the importance of involving students in the learning process and, as it allows students' criteria to lead the process, raises important issues about the effectiveness of involving students in assessment.

Chi Yen Chiu, in *The Discourse of an English Teacher in a Cyber Writing Course: Roles and Autonomy*, investigates the relationships of teacher roles and learner autonomy in a course in which the written e-mail text was the only means of communication. Teaching roles were not found to provide opportunities for promoting learner autonomy but the counseling roles adopted were found to create the kind of supportive learning environment that helped develop autonomy in language learning.

AEJ is always happy to receive papers that not only identify problems but also suggest feasible solutions. Nguyen Thi Mai Hoa, in *Mentoring beginning EFL teachers at tertiary level in Vietnam*, argues that in contrast to some Western countries, mentoring is underemployed in beginning teacher education in Vietnam. Her study aims at raising awareness of the benefits of providing guidance to teachers who are starting up in the profession and, more importantly, this paper suggests a solution to the problem identified, providing a blueprint for implementing a mentoring programme.

Zhu Xinhua's study, *Is Syntactic Maturity a Reliable Measurement To Investigate The Relationship Between English Speaking And Writing?*, revealed that measures of syntactic maturity can differentiate between proficiency levels and is therefore a useful measurement tool for other researchers in this area. Zhu Xinhua does also point out the limitations of this method for distinguishing between the two modes.

Mahmood Rouhani, in *Another look at the C Test: A Validation Study with Iranian EFL Learners*, investigates the validity of the C-Test for the assessment of global language proficiency. This study, while focusing on Iranian learners, should be of interest to other researchers on the C-Test across national borders. His results indicated that the C-Test enjoyed high reliability and acceptable content relevance and fairly high criterion-related validity. On the other hand, Rouhani concluded that at least in this context, the C-Test texts did not behave consistently with examinees across proficiency levels, failing to discriminate well, for example, between participants of lower and upper intermediate levels.

In *Lexical Collocations and their Relation to Speaking Proficiency*, Jeng-yih (Tim) Hsu and Chu-yao Chiu explore the knowledge and use of English lexical collocations in relation to speaking proficiency. Their findings support a view that *knowledge* of such phenomena as collocation is a significant component of proficiency. While there was a significant correlation between these Taiwanese EFL learners' knowledge of lexical collocations and

their speaking proficiency, no significant correlation was found between the subjects' actual use of lexical collocations and their speaking proficiency.

Lu-Fang Lin, in *The Study of English Learners' Synthesizing Process While Reading*, examines the ability to retell synthetic information in familiar and unfamiliar topic passages. Her study confirms the view that macrostructure formation occurred as an integral part of comprehension. The impact of cross-cultural knowledge was also confirmed. However, the study did not find that non-native students had more difficulty in synthesizing information at intra- and inter-sentential levels.

Nilton Hitotuzi, in *An Economical Approach towards Interaction in the L2 Classroom: A Task-based Learning Experiment*, considers the advantages (as an alternative to task-repetition) of keeping learners engaged in meaningful interaction in the classroom for an extended period of time. Hitotuzi opts for a holistic model that considers different phases within micro-task frameworks which then feed important cumulative data into a macro-task cycle. This is another interesting example of the growing perception that different tasks, or tasks and exercises of different kinds and sizes, need to be considered within some kind of broad framework to respond to a broad range of SLA criteria for effective language learning.

Finally, Daniela Nikolova in *English-teaching in Elementary Schools in Japan: A Review of a Current Government Survey*, presents a critical overview of the current situation in English-Teaching in public elementary schools in Japan. This piece adopts what AEJ currently calls an "alternative" approach to article writing in that it emphasizes a first person view of a critical EFL situation. Nikolova briefly reviews the history of English language education in Japan and underlines the challenges that the most recent English Language programs face in current school curricula. Nikolova makes suggestions for radical changes in the English curriculum.

**Dr. Roger Nunn**  
**Senior Associate Editor**  
**Asian EFL Journal**



## **Understanding EFL Learners' Strategy Use and Strategy Awareness**

**Kyoung Rang Lee, Ph.D.**  
*Sejong University, Korea*

**Rebecca Oxford, Ph.D.**  
*University of Maryland, U.S.A.*

### **Bio Data:**

**Kyoung Rang Lee** is an Assistant Professor at Sejong University in Seoul, Korea. She is interested in individual differences in teaching and learning English, including learning strategies of both teachers and students. Currently, she is devoted to better understanding and promoting Koreans' English learning strategy awareness and use.

**Rebecca Oxford** is Professor and Distinguished Scholar-Teacher at the University of Maryland. She has authored or edited a number of books on learning strategies, motivation, and language education. She has also edited the Tapestry Program, a series of English textbooks for college students.

### **Abstract**

This study discusses the statistically significant impact of Korean students' (from middle school, high school, and university, N=1,110) strategy awareness, English-learning self-image, and Importance of English on language learning strategy use. Students who had certain characteristics – valuing English as important (Importance of English), evaluating their own proficiency as high (English-learning self-image), and being already aware of many language learning strategies – employed learning strategies more frequently than those who did not (all significant at  $p < .000$ ). As expected, strategy awareness and strategy use were related to the Korean cultural context. Cognitive and cultural interpretations of the results were presented. Implications for English language learners including Korean students were also addressed.

**Keywords:** English learning strategies, strategy use, strategy awareness, self-image, Importance of English

### **Introduction**

English is one of the most important subjects in many schools around the world, including Korea. These days many countries have introduced English classes to elementary schools, and many adults study English even after graduating from universities. For example, Korean students study English for at least 10 years on the average: three years in middle school, three years in high school, and four years at the university. In spite of this enthusiasm for studying English, the profile of Korean students' strategies for learning

English is not well researched, which is similar to that of English learners' strategies in the context of English as a foreign language (EFL). Likewise, "learning strategy" is still quite a vague concept to Korean EFL learners, although learning strategies could definitely help them learn English more efficiently if they knew and employed such strategies consciously.

Many education studies have investigated learning strategies since the 1980's, and this has also been a trend in second and foreign language education. Researchers have discovered that successful L2 learners, compared with their less successful classmates, used more strategic mental processes (learning strategies) and employed them more frequently; this strategy use was shown to occur before, during, and after L2 tasks (Oxford, 1994; Oxford, Cho, Leung, & Kim, 2004). Oxford (1990) pointed out how important learning strategies are, both in theory and in practice, for language learners. Rubin (1994) described learning strategies as behaviors that would contribute to developing learner's language system affecting learning directly. This information provides a background for understanding the L2 strategy research reviewed in this section.

### **Strategy Use**

#### ***Strategy Use and Gender***

Many empirical studies show women are different from men in language learning strategy use, with women generally using more strategies than men, but not in all cases (Dreyer & Oxford, 1996; Green & Oxford, 1995; Lan & Oxford, 2003; Lee & Oh, 2001; Oxford & Ehrman, 1995).

Oxford, Nyikos, and Ehrman (1988) summarized four studies concerning gender differences in language learning, confirming that females use a greater range of language learning strategies. Zimmerman and Martinez-Pons (1990) discovered that girls use metacognitive strategies, such as goal-setting, planning, keeping records, and monitoring, more than boys. According to Green and Oxford (1995), 15 out of 50 strategies on the *Strategy Inventory for Language Learning (SILL)*; Oxford, 1990) showed differences between women and men in terms of strategy use, with women using them more frequently, while only one strategy was used more often by men than women. Oxford and Ehrman's (1995) comprehensive study, with 520 language learners for an average of 20 weeks, also discovered that females' use of strategies was more frequent than males'. Lan and Oxford (2003) found that with the *Taiwanese Children's SILL*, significant differences in strategy use between girls and boys were present for 11 out of 50 strategies, with these

differences in favor of greater strategy use by girls.

In contrast to these significant gender differences, there are also studies showing a less clear distinction in strategy use between males and females (Dadour & Robbins, 1996; Oh, 1996; Park, 1999). Kaylani (1996) found out that girls are different from boys in terms of strategy use, not because of only gender, but because of gender in relation to proficiency.

### **Strategy Use and Major**

Similar to gender, academic majors generally affect students' use of learning strategies. Generally speaking, students majoring in humanities used more and a wider range of strategies than those majoring in science/engineering in several studies (e.g., Lee, 1994; Park, 1999). Dreyer and Oxford (1996) and Oxford and Nyikos (1989) also showed significant influences of university majors on students' strategy use.

### **Strategy Use and Age**

Many strategy studies have been conducted with college students or adults (Dadour & Robbins, 1996; Green & Oxford, 1995; Leki, 1995; Oxford & Ehrman, 1995; Phillips, 1991). Some studies have focused on younger students or have compared younger learners with college students (Dörnyei, 1995; Kaylani, 1996; Lan & Oxford, 2003; Lee, 2000; National Capital Language Resource Center [NCLRC], 1996 & 2000). Several studies showed that young learners tended to use social strategies more than other types of strategies, including discussing with and asking help from others (Lee, 2000; Wong Fillmore et al., 1985). In contrast, adult learners have shown high use of metacognitive strategies for planning, organizing, and evaluating their own L2 learning (Oh, 1992; Touba, 1992).

### ***Strategy Use and English Proficiency (English-Learning Self-Image)***

Since numerous studies about L2 learning strategies have been rooted in the distinction between good and poor learners, there are many studies about the relationship between strategy use and L2 proficiency. Some have used actual proficiency test scores (Dreyer & Oxford, 1995; Green & Oxford, 1995; Phillips, 1991), while others have used proficiency self-ratings (Wharton, 2000). Most researchers have agreed that more proficient learners employ a wider range of strategies more efficiently than less proficient learners (Green & Oxford, 1995; Kaylani, 1996; Lan & Oxford, 2003; Oxford, 1996; Oxford & Ehrman, 1995; Philips, 1991). Strategy use was significantly correlated with English proficiency

scores of university students learning English as a second language (ESL) in South Africa ( $r=.73$ ,  $p<.0001$ ) (Dreyer & Oxford, 1996). Research in Asia, such as in Thailand (Mullins, 1992), in Japan (Watanabe, 1990), and in Korea (Kim, 2000; Lee, 2000; Lee & Oh, 2001; J. Park, 2001; Y. Park, 1999; Yoon, Won, & Kang, 2001), also showed strong, positive correlations between strategy use and EFL proficiency.

### **Strategy Awareness**

Most investigators have agreed that awareness helps students learn a language and use strategies, at least in the earlier stages of learning (Chamot, 1998; Cohen, 1995; NCLRC, 1996; O'Malley & Chamot, 1990; Oxford, 1990; Oxford & Cohen, 1992). According to Cohen (1995), when students are no longer aware of their behaviors to learn a language, these behaviors are, by definition, no longer strategies but are instead processes; thus, he was emphasizing, through definitions, the importance of strategy awareness (a key element of consciousness) (see also, Carrell, Gajdusek, & Wise, 1998). Chamot (1998) stressed that learning strategies are teachable (see also Green & Oxford, 1995), i.e., that students can become more aware of strategies through strategy instruction. Chamot (1998) also stated that awareness of one's own strategies is closely related to metacognition, adding that more successful learners have better and more metacognitive awareness.

Despite the importance of awareness, few studies have been done concerning students' metacognitive awareness of L2 strategies. Carrell (1989, 1991) used a questionnaire, *Metacognitive Awareness Strategies Questionnaire (MASQ)*, (1989), to measure metacognitive strategies. Results underscored the effects of metacognitive awareness on reading and the importance of discerning learners' level of awareness of reading strategies. Though many researchers believe that learner awareness is a necessary feature of strategy use (e.g., Carrell, 1989; Cohen, 1995), we still can see situations like Baker and Brown (1984) pointed out: when a reader did not describe how to use a particular strategy but did in fact use it. This will become important in considering the results of the current study.

Based on the previous research, the current study aims to help to better understand Korean learners' English learning strategies in terms of strategy awareness, English-learning self-image, and importance of English, which have been known to be related with strategy use. In this study, Oxford's (1990) *SILL* is used to measure Korean English language learners' strategy use, and a column has been added next to each strategy on the *SILL* to ask whether the participants are aware of each strategy<sup>1</sup>. The *Background Information Questionnaire* was administered to measure other relevant

variables, such as gender, major, education level, English-learning self-image, and importance of English. At the same time, the study examined the culturally-related construct validity of the *SILL* for Korean EFL learners.

Based on this goal, this study asks the following research questions:

- (1) Are there any significant main effects or interaction effects of the following variables on strategy use: (a) gender, (b) major, (c) education level, (d) English-learning self-image, (e) importance of English, and (f) strategy awareness?
- (2) Which are the best predictors of strategy use among the given variables?
- (3) Are there any significant main effects or interaction effects of the given variables on strategy awareness: (a) gender, (b) major, (c) education level, (d) English-learning self-image, and (e) importance of English?
- (4) Which are the best predictors of strategy awareness among the given variables?
- (5) What strategies do Korean EFL students use that are not reflected in the *SILL*?
- (6) What strategies in the *SILL* do students perceive as culturally relevant in the Korean EFL situation?

### Methods

This study primarily used quantitative methods to gain broader perspectives on Korean EFL learners' strategy use. However, two open-ended questions were added to the main questionnaire, the *SILL*, to provide in-depth, qualitative data.

### Participants

More than 1,000 students, ranging from middle school to university students, participated in this study. Table 1 shows the characteristics of the participants. Data were collected from one middle school, three high schools, and two universities. Table 2 displays the locations of the schools and the socioeconomic status of the participants.

Table 1. *Total Participants and Number of Participants by Gender*

	Total	Male	Female
Middle School	379	188	191
High School	438	256	182
University	293	173	120
Total	1,110	617	493

Table 2. *Geographic Location of Each Educational Institution and General Socioeconomic Status of Students in Each Institution*

	Geographic Location	General Socioeconomic Status (SES) of Students
Middle School	-This middle school is in the eastern part of Seoul	Medium to high SES
High School	-One high school is in the suburbs of Seoul -Others are in the eastern part of Seoul	Medium to high SES Medium to high SES
University	-One university is in Seoul -One university is in the middle of Korea	Both have mixed SES (ranging from low SES to high SES). These are the best-known universities in their regions.

**Instrumentation**

Version 7.0 of the *SILL* (Oxford, 1990) was chosen for this study but was extended in certain new ways, as described below. This version includes 50 quantitative, close-ended items<sup>2</sup>, with each item measuring the frequency with which respondents use the particular strategy. Examples of items on Version 7.0 are: “I plan my schedule so I will have enough time to study English,” and “I use new English words in a sentence so I can remember them.” Students responded to each strategy item using a Likert scale of 1 through 5 to reveal the frequency with which they used the strategy, with 1 representing “Never or almost never true of me” (i.e., rarely or never used) and 5 representing “Always or almost always true of me” (i.e., always or almost always used).

The *SILL* contains six factor-analytically created strategy categories: memory-related (primarily for helping vocabulary enter long-term memory), cognitive (for processing and practicing language information), compensatory (for making up for missing knowledge, as in guessing from the context), metacognitive (for planning, organizing, and evaluating one’s own learning), affective (for managing emotion and motivation), and social strategies (for working with others or asking questions of others). The exploratory factor analysis on which the categories were initially based was conducted with 1,200 U.S. foreign language students studying a variety of languages, including French, Spanish, German, Russian, and Italian (Nyikos & Oxford, 1993; Oxford, 1996). Other studies have factor analyzed the *SILL* in EFL contexts (Hsiao & Oxford, 2002; Oxford, 1996; Oxford & Burry-Stock, 1995). Hsiao and Oxford (2002) reported a confirmatory factor analysis, which tested 15 potential strategy-use models with several strategy questionnaires. The *SILL* was the only strategy questionnaire that showed acceptable fit indices<sup>3</sup>.

The first author translated the *SILL* into Korean and then compared it to three other,

already published Korean translations of the *SILL*. We wanted to ensure that the translation used in this study was as faithful as possible to the English version. Since three versions of the translated *SILL* in Korean (Lee, 1994; Park, 1999; Park, Kim, & Park, 2003) were already available, the first author used an adapted Parallel Blind Technique (Behling & Law, 2000) to compare the four Korean translations, hers and the three published translations. In the Parallel Blind Technique, two or more translators independently translate the material. After the translations are complete, the works are compared for the purpose of identifying and resolving discrepancies. This study's translation of the *SILL* was further checked and approved by a Korean English teacher who holds a master's degree in English education and was in her thirteenth year of teaching English in Korea at the time of the study. This additional check was done to ensure that the wording of all items was appropriate for students from middle school through university. In this study the 50 Korean-translated strategy-use items had an overall reliability (Cronbach alpha) of .94,  $N=1,110$ . This is approximately the same level of reliability as typically found with the 50-item version of the *SILL* (Oxford, 1996).

For the first time, a measurement of metacognitive strategy awareness was added to the *SILL*. Specifically, students were asked to respond to the following question with regard to each strategy-use item: "Did you know (think about) this strategy before now?" The intention was to discover whether the respondents were aware of a given strategy prior to taking the *SILL*. This was important because this basic awareness should precede "metacognitive knowledge about specific strategies [the value of procedures as well as when and where to use strategies]" (Pressley, Goodchild, Fleet, Zajchowski, & Evans, 1989, p.305), which plays a very important role in adequate transfer of strategy use. This move thus added to the *SILL* 50 dichotomous (yes-no) strategy-awareness items, each corresponding to a given strategy. The overall reliability of these 50 strategy-awareness items was calculated with a total of 1,090 Korean students responding to these items; the Cronbach alpha reliability index was .94.

New qualitative items were also added to the *SILL*. The qualitative items were as follows: (a) "Can you identify any strategies you have been using but that you can't find in this *SILL*? If so, please list the strategies below." (b) "If you found any strategy on the *SILL* that cannot easily be employed in Korea, please write down the strategy. If you can revise it for use in the Korean situation, please do so." These items were added for this study to discover any cultural beliefs and practices relating to strategy use, as well as to

examine the culturally-related validity of the *SILL* in an EFL situation where English native speakers are rare.

The *Background Information Questionnaire* was adapted from Oxford's (1990) *Background Questionnaire* (p. 282). It was originally written in English and then translated into Korean. Items concerned gender, major, current education level (middle school, high school, and university), importance of English, and English-learning self-image. The item assessing importance of English was, "How important is it for you to become proficient in English?" with the following response options: very important, important, somewhat important, not so important, and not important at all. Also, the item measuring English-learning self-image was, "How do rate your overall English proficiency as compared with the proficiency of *other students in your class*?" (emphasis in original). Response options in the current study included: excellent, very good, good, fair, and poor. We felt it was very important to obtain students' metacognitive self-assessment of proficiency vis-à-vis their peers because we agree with motivational theorists like Harter (1986). They believe that "humans have complex perceptions of themselves and their competences. . . All of these self-perceptions have motivational properties, with competence on a task more directly influenced by task-limited self-perceptions than global self-esteem" (Pressley et al., 1989, p. 307). We consider that this self-perception on his/her own English proficiency will play a very important role in actual performances. In addition, one of the most important themes of this study is awareness of one's own strategy use; therefore, we considered self-rated English proficiency to be more relevant to our study than standardized test scores<sup>4</sup>.

#### **Data Collection and Analysis Procedures**

The questionnaires were distributed to eight teachers, who administered them to students as in-class activities. This made it possible to obtain complete responses from almost all intended participants. A very small number of students (<1%) answered all the same answers to every item in an insincere manner, so their responses were excluded from this study. The first author provided guidelines to help teachers administer the questionnaires, and the teachers explained to their students that the results would not affect their grades in any way. The students answered the questionnaires voluntarily, and those who did not wish to participate did not have to do so.

Quantitative analyses were conducted using the *Statistical Package for the Social*

*Sciences* (SPSS 13.0). For various research questions, independent t-tests, and multiple regression analysis were employed. For the Likert-scaled strategy-use items on the *SILL*, the following key helped to interpret the means: 3.5 to 5.0, high use; 2.5 to 3.4, medium use; and 1.0 to 2.4, low use (Oxford, 1990). The qualitative results were summarized, categorized, and analyzed for frequency of mention. The answers to the two open-ended *SILL* questions were translated from Korean to English. To assess the quality of the translation, the Korean English-teacher in Korea and a Korean-English bilingual doctoral student in the U.S. were asked to compare the Korean responses to the two open-ended *SILL* items with the corresponding English translations. This process revealed that there were no areas of translation discrepancy.

### **Results**

In this section, these qualitative as well as quantitative results are presented in terms of strategy use, strategy awareness, and Korean EFL learners' English learning strategies.

#### **Strategy Use**

In contrast to preconceptions about Asians as constant memory-strategy users, compensatory strategies for making up for missing knowledge were used the most frequently (mean=3.10, SD=.83, medium use range) of the six strategy categories. The other strategy categories in order of frequency of use for the entire sample (all ages) were as follows: metacognitive (mean=2.83, SD=.81, medium use range), cognitive (mean=2.79, SD=.67, medium use range), memory (mean=2.61, SD=.71, medium use range), affective (mean=2.48, SD=.76, medium to low use range), and social strategies (mean=2.43, SD=.87, low use range).

Table 3 shows the different profile of strategy use in terms of the given variables. As shown by many previous studies, (a) women used strategies more often than men ( $t=3.47$ ,  $df=1108$ ,  $p<.01$ ), and (b) students majoring in humanities, more often than those majoring in science/engineering ( $t=1.96$ ,  $df=989$ ,  $p<.05$ ). In addition, (c) strategy use was higher for the following groups than for other groups: university students ( $F=29.74$ ,  $df=2$ ,  $p<.05$ ), students who rated their proficiency as excellent ( $F=54.45$ ,  $df=4$ ,  $p<.001$ ), students who considered English to be important ( $F=37.47$ ,  $df=4$ ,  $p<.001$ ), and students who were typically aware of strategies ( $t=15.69$ ,  $df=1088$ ,  $p<.001$ ).

Table 3. Strategy Use by Different Variables

Variable	Strategy Use Means (1 to 5)				
Awareness	Yes	No			
	2.94 (N=641)	2.41 (N=449)			
Gender	Female	Male			
	2.79 (N=493)	2.67 (N=617)			
Major	Humanities	Science/Engineering			
	2.77 (N=558)	2.69 (N=433)			
Education Level	Middle	High School	University		
	School	2.57 (N=438)	2.91		
	2.75 (N=379)	(N=293)			
English-learning Self-image	Poor	Fair	Good	Very	Excellent
	2.24	2.47 (N=159)	2.63	good	3.14
	(N=89)	(N=388)			(N=112)
	(N=358)				
Importance	Not	Not so important	Somewhat	Quite	Very
	important	2.16 (N=23)	important	important	important
	1.90	2.41			2.70
	(N=16)	(N=159)			(N=452)
	(N=454)				

Table 4 shows that the main effects of (a) strategy awareness, (b) education level, (c) English-learning self-image, and (d) importance of English on strategy use were very significant. Gender and major did not have significant main effects alone; however, gender showed significant interaction effects with other variables. Strategy awareness had a significant main effect on strategy use; the more that a student was aware of strategies, the more the student reported using strategies. Evidence from means also showed the significantly higher use of strategies by those who answered “yes” than “no” to the awareness questions. Education level also showed a significant main effect on strategy use. Specifically, university students used strategies the most frequently, followed by middle school students and high school students. Reasons for this possibly surprising finding are explored in the discussion section. Another significant main effect was found for self-image, that is, the students’ self-rating of their English proficiency. A fourth significant main effect appeared for importance of English (perceived importance of English).

Table 4. Main Effects and Interaction Effects on Strategy Use

Dependent variable: Strategy Use						
Source	Type III Sum of Squares	df	Mean Squares	F	Sig.	Eta Squared
Corrected Model	179.55	179	1.00	4.30	.000	.49
Intercept	289.81	1	289.81	1241.33	.000	.61
Awareness	28.85	1	28.85	123.56	.000	.14
Gender	.014	1	.014	.06	.808	.00
Education Level	1.71	2	.86	3.66	.026	.01
Major	.003	1	.003	.011	.916	.00
English-learning Self-image	5.77	4	1.44	4.18	.000	.03
Importance of English	7.36	4	1.84	7.88	.000	.04
Gender * Education level	1.74	2	.87	3.73	.024	.01
Gender * Major	1.15	1	1.15	4.96	.027	.01
Gender * Self-image	3.32	4	.83	3.55	.007	.02
Gender * Importance of English	2.31	3	.77	3.30	.020	.01
Gender * Self-image * Importance of English	3.75	8	.47	2.01	.043	.02

Five interaction effects were significant. Without a significant main effect, gender played an important role in all the interaction effects: (a) gender with education level, (b) gender with major, (c) gender with self-image, (d) gender with importance of English, and (e) gender with self-image and importance of English.

According to Cohen (1988), a large effect size is .40 or over, a medium effect size is .25 to .39, and a small effect size is .10 to .24. Applying Cohen's criteria for effect size, the overall model had a large effect size (.49). Effect sizes for the individual main effects and the interaction effects were not as impressive as the overall model's effect size. The significant main effect for strategy awareness had a small effect size (.14). Effect sizes for other significant main effects (i.e., for education level, self-image, and importance of English) and for significant interaction effects were all very small, although they contributed to the large overall model effect size.

To discover the best predictors of overall strategy use among the given variables, a stepwise multiple linear regression was conducted. The relationship between the predictors and strategy use was  $R=.60$ , and the predictors explained 36% of the variability in strategy use ( $R^2=.36$ , Adjusted  $R^2=.36$ ,  $p=.000$ ). This is a large amount of variance explained for an educational study such as this.

Table 5. Multiple Regression Results of Strategy Use

Model	R	R <sup>2</sup>	Adjusted R <sup>2</sup>	F	Sig.
1	.50 <sup>a</sup>	.25	.25	321.44	.000 <sup>a</sup>
2	.57 <sup>b</sup>	.32	.32	229.04	.000 <sup>b</sup>
3	.60 <sup>c</sup>	.36	.36	181.24	.000 <sup>c</sup>
4	.60 <sup>d</sup>	.36	.36	137.79	.000 <sup>d</sup>

- a. predictor: strategy awareness
- b. predictors: strategy awareness, English-learning self-image
- c. predictors: strategy awareness, English-learning self-image, importance of English
- d. predictors: strategy awareness, English-learning self-image, importance of English, education level

Strategy awareness was the best predictor (Beta=.37, p=.000) of strategy use, followed by English-learning self-image (self-rated proficiency) (Beta=.25, p=.000), importance of English (Beta=.19, p=.000), and education level (Beta=.07, p=.012). Gender and major were not significant predictors of overall strategy use. One quarter of the variance in strategy use was explained by strategy awareness alone (R<sup>2</sup>=.25, Adjusted R<sup>2</sup>=.25, p=.000).

**Strategy Awareness**

Similarly to the results of strategy use, Korean students were aware of compensatory strategies (mean=.62, SD=.30) best, followed by metacognitive (mean=.60, SD=.30), cognitive (mean=.56, SD=.26), memory (mean=.53, SD=.28), social (mean=.49, SD=.34), and affective strategies (mean=.46, SD=.30).

Table 6 shows the different profile of strategy awareness in terms of the given variables: (a) university students were the most aware of strategies of all (F=44.30, df=2, p<.001), and (b) students who rated their English proficiency as excellent (F=2.493, df=4, p<.001) and those who considered English to be very important (F=16.62, df=4, p<.001) were more aware of strategies than other students.

Table 6. Strategy Awareness by Different Variables

Variable	Strategy Awareness Means (Yes=1, No=0)		
Gender	Female	Male	
	.56 (N=486)	.54 (N=604)	
Major	Humanities	Science/Engineering	
	.56 (N=548)	.53 (N=425)	
Education Level	Middle School	High School	University
		.48	.64

	.56 (N=371)	(N=435)	(N=284)		
English-learning Self-image	Poor .48 (N=88)	Fair .46 (N=157)	Good .51 (N=382)	Very good .61 (N=352)	Excellent .65 (N=108)
Importance	Not important .42 (N=16)	Not so important .37 (N=23)	Somewhat important .45 (N=158)	Quite important .55 (N=439)	Very important .59 (N=449)

Table 7 shows that the main effects of (a) education level, (b) English-learning self-image, and (c) importance of English on strategy awareness were very significant. Gender and major did not have significant main effects. The only significant interaction effect on strategy awareness was found for the combination of these variables: major, English-learning self-image, and importance of English.

Table 7. *Main Effects and Interaction Effects on Strategy Awareness*

Dependent variable: Strategy Use						
Source	Type III Sum of Squares	df	Mean Squares	F	Sig.	Eta Squared
Corrected Model	18.17	178	.10	2.26	.000	.34
Intercept	33.20	1	33.20	735.66	.000	.48
Gender	.021	1	.021	.46	.496	.00
Education Level	.70	2	.35	7.73	.000	.02
Major	.01	1	.01	.20	.658	.00
English-learning Self-image	.68	4	.17	3.79	.005	.02
Importance of English	1.02	4	.26	5.65	.000	.03
Major * Self-image * Importance of English	.92	8	.12	2.54	.010	.03

The effect size of the overall strategy awareness model had a medium effect size (.34), a little lower than that of the overall strategy use model. However, the effect sizes of the variables were very small despite the significant main effects.

Predictors of strategy awareness were also tested by a stepwise multiple linear regression. Because strategy awareness was one of the best predictors of strategy use, finding its best predictors would ultimately help teachers teach students how to promote strategy use and how to use strategies properly. However, the overall relationship between strategy awareness and its predictors was not very high ( $R=.35$ ), so the proportion of the variability in strategy awareness explained by the predictors was small ( $R^2=.12$ , Adjusted

$R^2=.12$ ,  $p=.000$ ). Among the given variables, students' self-rated English proficiency (English-learning self-image) ( $Beta=.24$ ,  $p=.000$ ) was the best predictor of strategy awareness, followed by rating of importance of English ( $Beta=.17$ ,  $p=.000$ ) and education level ( $Beta=.13$ ,  $p=.000$ ). Gender and major were not significant predictors.

### ***Korean EFL Learners' Strategies***

The results of this question indicated that participants were using different kinds of memory strategies from those on the *SILL*. Most write-in answers to Research Question 5 were about the specific memory strategies favored by Koreans, such as repetition and dictation. Most students said, "I write a lot to memorize English words or expressions," and "I repeat a tape a lot so that I can get used to its expressions." Table 8 shows the top five write-in responses of this question. More specifically, Table 9 shows the different strategies used by different education levels.

Table 8. *Koreans' Strategies Not Reflected in the SILL*

Order	Strategy
1	I write a lot to memorize English words or expressions.
2	I repeat a tape a lot so that I can get used to its expressions.
3	I dictate while listening to English tapes several times.
4	I watch a movie / listen to a pop song until I memorize the lines.
5	I put new words or expressions on everywhere to memorize them any time.

Middle school students especially liked writing English words or expressions several times to memorize them, or having pen pals to write English letters regularly. They also tried to write English words or sentences and post them wherever they could see them often. In addition, they recorded their reading in English to correct wrong pronunciations, or they played games to learn English.

High school students also liked writing new words or expressions several times to learn them by heart. Some high school students even kept a journal of wrong answers not to make the same mistake again, which is a test-oriented strategy. Similar to the quantitative *SILL* results about frequency of strategy use, high school students listed fewer strategies than those in the middle school or universities. In particular, only high school students did not write in a new strategy related to seeing movies in English, because studying for the entrance examination had discouraged them from seeing movies. Instead, high school students tried to come up with an motivating reason why they must study English, or associate words or sentences with what they liked, which may be related to the fact that more than half of high school students did not enjoy English.

In contrast, university students did not have to memorize English words or sentences like high school and middle school students did. Thus, they wrote that they used entertaining materials, such as English tapes, movies, the Internet, newspaper, or magazines, to make themselves get used to sounds and expressions. They usually filled in blanks or dictated whole sentences while listening to English. In addition, they imitated English speaking actors or actresses, or tried to come up with their lines in advance while watching a movie. University students were more positive about opportunities for contacting native speakers and for traveling to English-speaking countries. They also used monolingual dictionaries instead of bilingual dictionaries.

In general, Korean students depended on mass media, such as English movies, popular songs, and books written in English, in order to compensate for the lack of native English speakers. This seemed a reason that made most students tended not to use social strategies, except for a few university students. Korean specific educational situation, where the entrance examination is the most important in students' life, high school students used test-oriented strategies mostly unlike the others who are not immediately facing it.

*Table 9. Different Strategies by Different Education Levels*

	Strategy
Middle School Students	<ul style="list-style-type: none"> <li>-write to memorize English a lot</li> <li>-make a pen pal to write English letter regularly</li> <li>-read books written in English</li> <li>-put new words on everywhere to read and memorize the words</li> <li>-memorize a whole sentence</li> <li>-repeat the tape a lot</li> <li>-record my own voice and listen to correct it</li> <li>-listen to English one hour a day</li> <li>-watch movies without scripts</li> <li>-learn English through games</li> </ul>
High School Students	<ul style="list-style-type: none"> <li>-write to memorize English a lot</li> <li>-prepare a notebook to gather wrong answers</li> <li>-write new expressions and read them very often</li> <li>-don't give up despite the bad scores</li> <li>-come up a motivation saying that I must study English</li> <li>-associate words or sentences with what I like</li> </ul>
University Students	<ul style="list-style-type: none"> <li>-repeat new words a lot to memorize</li> <li>-dictate while listening English tapes</li> <li>-read English novel or newspaper</li> <li>-read various English materials about something that I am interested in (book, internet...)</li> <li>-read books written in English</li> <li>-make English sentence for every expression</li> <li>-use monolingual dictionaries instead of bilingual dictionaries -try to come</li> </ul>

	up the next line when watching movies -try to find incorrect scripts while watching movies -imitate an actor or an actress in movies and radio/TV programs -travel to the English speaking country -make a friend with a native speaker and talk in English -try to think in English
--	---

Regarding the research question 6, most strategies that students considered not feasible in the Korean EFL situation were strategies related to native speakers: “I look for people I can talk to in English”; “I ask English speakers to correct me when I talk”; and “I ask for help from English speakers.” Also, some strategies relevant to the ESL situation were mentioned as difficult to use in the Korean EFL context, such as “I remember new English words or phrases by remembering their location on the page, on the board, or on a street sign,” and “I start conversations in English.” The most reasons given by the participants were “It’s difficult to find native English speakers,” “We can’t see native English speakers, so we can’t ask help from them,” and “I’ve never talked to native English speakers.” Many middle school students reported that they had never seen native speakers, so they could not even imagine a situation where English speakers could help them. University students interpreted native speakers as peers or English teachers, who were good at English, even without such guidelines.

Again, high school students answered differently from the others because of their test-oriented attitudes due to the specific Korean educational situation. Some of them explicitly mentioned the entrance examinations, saying that the examinations had prevented them from relating English to fun and emphasizing that they never studied English for fun.

### **Discussion**

Even though there were several strategies that Korean students considered to be impossible in the Korean EFL situation, the general profile of strategy use, drawn from the *SILL* in this study, was very consistent throughout the different education levels of the participants. Specifically, the order of strategy frequency was the same for all education levels: compensatory, metacognitive, cognitive, memory, social, and affective strategies. This section presents cognitive and cultural interpretations of the results in terms of (a) gender, (b) major, (c) education level, (d) English-learning self-image, and (e) importance of English. The relationship between strategy use and strategy awareness is also addressed.

***Effect of Gender Differences on Strategy Use and Strategy Awareness***

Different from the previous research and expectations, gender did not affect strategy use and awareness significantly. It did not have large effect sizes either. Gender played a significant role only with other variables, not by itself. More interestingly, except social strategies, females used five of the six categories significantly more often than males ( $p < .001$ ). This is very interesting, considering previous studies showing that social strategies are the best-known strategies employed more by women than by men (Oxford, Nyikos, & Ehrman, 1988).

Regardless of gender, Korean students are not typically encouraged to talk with classmates, so it stands to reason that social strategies might not show a significant gender effect. This may also be true in such countries where teachers are authoritative figures. Another possibility is that Korean students could not use the social strategies in the *SILL*; half of the social strategies in the *SILL* involve working with or communicating with native English speakers, and since there are few such speakers available in Korea, this might have suppressed social strategy use for both males and females. All the other EFL learners will experience such difficulties as Koreans.

Surprisingly, metacognitive strategies ( $t = 2.80$ ,  $p < .005$ ) showed significantly more awareness by males than by females. This is also very interesting because men were more aware of their metacognitive strategies than women, while using them less often than women. According to Pressley (2000), good readers use a strategy only when it is necessary to help their learning, thus, it is possible that men were better English learners than women. However, the English-learning self-image of men (mean = 3.19) and that of women (mean = 3.26) were almost the same (both considered their English proficiency to be good). Comparison between the self-ratings and the official scores is needed.

In sum, gender, though reaching statistical significance on some interaction effects, had effect sizes that were negligible. We think that it is still good for learners because we cannot change nor manipulate gender. The smaller the effect size of gender, the more possible it is for teachers or learners to improve learning.

***Effect of Major on Strategy Use and Strategy Awareness***

Students who were majoring—or want to major—in humanities had a higher overall mean (2.77, medium use range) than students majoring in science/engineering fields (2.69, medium use range), but it was not significant,  $p > .05$ . The former were also more aware of

their strategies than the latter ( $p > .05$ ). Metacognitive strategy use was the only category in which major had a significant difference, with humanities majors using the metacognitive category of strategies more frequently ( $p < .05$ ) and being more aware ( $p < .05$ ) than their science/engineering peers. This conforms to the general belief that humanities students are more reflective (metacognitive) than science/engineering students.

Except for this, major did not significantly influence strategy use or strategy awareness. Because middle school students chose as majors what they wanted to focus on in high school, they might not have known exactly what humanities and science/engineering fields were like. This might have related to non-significant influences of major on both strategy use and awareness. Based on this, it may be better to study the relationship between strategy and major of only university students who have clear ideas on their majors.

#### ***Effect of Education Level on Strategy Use and Strategy Awareness***

Education levels turned out to affect both strategy use and strategy awareness very significantly. As addressed in the results section, middle school and university students were more aware of as well as employed strategies significantly more often than high school students,  $p < .001$ . Korea's college entrance examinations seem to cause this difference between high school students and the others, probably because students in high school are discouraged from employing a wide range of strategies with great frequency, while middle school and university students are encouraged to do so. Lee (1994) found that Korean students used compensatory strategies as the most frequently used category, and this result appeared again in the present study. Since 1969, Korean students have taken multiple-choice entrance examinations, equivalent to the *Scholastic Aptitude Test* in the U.S., and we speculate that such examinations might promote compensatory strategies for guessing the right choice from the context, even if the details are not fully understood.

Compensatory strategies were among the two most frequently used categories for Anglophone children in Canada learning French (Gunning, Oxford, & Gatbonton, forthcoming) and for Taiwanese children learning English too (Lan & Oxford, 2003). It is possible that these children are also used to a similar type of a foreign language test—multiple choice.

#### ***Effect of English-learning Self-image on Strategy Use and Strategy Awareness***

The ranges of strategy use conformed to students' self-ratings of English proficiency

( $F=54.45$ ,  $p<.001$ ) and their awareness of strategies ( $F=24.93$ ,  $p<.001$ ), which means that the more highly they rated their English proficiency, the more they used strategies and the more they were aware of them. These results suggest that those who viewed themselves more positively than the others used strategies more often and were more of them than those who did not. Teachers should keep this result in mind so that they can help the learners promote their positive self-image, which will eventually help them learn English better.

***Effect of Importance of English on Strategy Use and Strategy Awareness***

The frequencies of strategy use conformed to students' ratings of the importance of English ( $F=37.47$ ,  $p<.001$ ). Strategy awareness was also positively influenced by students' attitudes toward the importance of English ( $F=16.62$ ,  $p<.001$ ). This is very encouraging because students' ratings of importance of English had a positive relationship with both strategy use and strategy awareness. Some students become more nervous when they consider something very important, so they perform worse than when they consider it unimportant. Teachers should properly emphasize how important English is, and the certain amount of anxiety will help students learn English more effectively.

***Effect of the Relationship between Strategy Use and Strategy Awareness***

As shown earlier, strategy awareness had a significant main effect on strategy use, although the effect size was small (.14). In a different analysis, strategy awareness, among the possible predictors included in the model, had the greatest ability to predict strategy use, but even so, it only explained 25% of the variability in strategy use. These relationships were significant, but other variables beyond strategy awareness (and the others in this study) might also influence strategy use. Baker and Brown (1984) noted that there is not a perfect connection between strategy use and strategy knowledge/awareness. Moreover, the role of awareness in strategy use is very complex. A strategy might fade from awareness by becoming automatic and habitual, at which time it is called a non-strategic "process"; but it can be brought back into consciousness (as a strategy) through direct instruction, reflection, or discussion. One thing is certain, however: As Pressley et al. (1989) noted, the learner can actively transfer a given strategy to a new learning situation *only* when the strategy is in awareness, i.e., when the learner has metacognitive knowledge of the strategy.

In particular, these four items showed non-significant impact of strategy awareness on

strategy use: “6: I use flashcards to remember new English words ( $p=.197$ )”; “14: I start conversations in English ( $p=.164$ )”; “27: I read English without looking up every new words ( $p=.147$ )”; and “42: I notice if I am tense or nervous when I am studying or using English ( $p=.054$ ).” The possible reason why several items showed the non-significant correlations between strategy use and strategy awareness is that, for example, students decided not to use flashcards even though they knew what the flashcards were, and they also decided to look up many words even though they knew it was not an effective strategy. Thus, it is necessary to practice strategies properly so that students can employ effective strategies and avoid ineffective behaviors.

Among different categories of strategies, the correlation between metacognitive strategy use and awareness was the highest ( $r=.43$ ,  $p<.001$ ). Metacognitive strategies involve awareness of cognitive processes, so this strongest relationship was not surprising at all. It suggests that teachers or learners should start from metacognitive strategies to promote their overall strategies. In addition, there are many ways for students to become aware of their strategies. For example, students might observe that they are using certain strategies. They might observe their friends’ or siblings’ strategy use. Their teachers could teach specific strategies. Uncovering the precise mechanism of strategy awareness might be somewhat difficult.

### ***Construct Validity of the SILL in Korea***

At first it might have been surprising that the *SILL* category of memory strategies was not more frequently employed by Korean students. We have to look into the strategy items in the *SILL* more carefully. The memory strategies in the *SILL* are mostly related to vocabulary. Also, it does not include any rote memory and repetition items, which are the basis of much of the successful memorizing by Asian students. The *SILL* memory category was structured to include a range of memory strategies based on visual, auditory, and kinesthetic modalities, and the Korean memory strategies might not have fitted that overall profile.

Regarding the strategies that the Korean participants considered difficult to employ in Korea, they were mostly the items related to native English speakers. It is not so easy to find native English speakers in Korea as in other ESL countries. However, those items can be easily adaptable when the items are revised to mean English teachers or better English learners, instead of native English speakers. Some students actually mentioned that they

asked help from English teachers because they can't from native English speakers.

Despite some memory strategies not reflected in the *SILL* and some other social strategies not easy to employ in Korea, the strategies in the *SILL* were usable and actually used in Korea. With minor revisions, the *SILL* can get the higher construct validity in the Korean EFL situation than as it is, resulting in helping both teachers and students in EFL countries like Korea.

### **Conclusions**

This study has focused on the influences of rarely-considered but very important variables, in addition to other variables, on EFL students' strategy use: (a) strategy awareness, (b) English-learning self-image, and (c) importance of English. Focusing on Korean EFL students, this study showed that except for major and gender, all the other variables had significant influences on strategy use and strategy awareness. Gender and major were expected to be helpful indicators of successful learning, but they turned out not to affect strategy use and awareness alone. However, when gender and major are combined with other variables, they interactively affected strategy use and awareness. Therefore, teachers should not emphasize stereotypical strategy use based on gender or majors.

This study also found that their strategy use was closely tied to strategy awareness and English-learning self-image. Both strategy awareness and English-learning self-images can be considered as metacognition about their English learning. Thus, the significant influences of strategy awareness and English-learning self-image on strategy use imply how to teach English effectively: teachers should promote positive self-image and strategy awareness when they teach learning strategies more effectively and efficiently.

In Korea, the entrance examinations' impact on students' learning proved to be very important, especially for high school students, according to the write-in data for the two qualitative questions and the quantitative findings. High school students used test-related strategies, but they did not use other strategies as often as students in middle school and universities. Moreover, high school students were less aware of these strategies than were students in the other education levels. In addition, the entrance examinations are multiple-choice standardized tests, which makes it unnecessary to understand the specific content correctly; this might have caused compensatory strategies to be the most frequently used strategies. Teachers in other countries where an entrance examination is seriously important, as well as in Korea, should encourage students to understand the details on standardized tests, not just the guessable outlines, so that students will not limit

themselves to using compensatory strategies the most frequently.

Lastly, to teach English in the EFL settings more effectively, several items in the *SILL* can be added or revised, according to Korean EFL students' suggestions in the open-ended questions. Teaching appropriate learning strategies is very important because learning strategies can enable students to become more independent, autonomous, lifelong learners (Allwright, 1990; Little, 1991).

### **Notes**

1. Strategy awareness is analyzed both as a dependent variable and as an independent variable.
2. *SILL* Version 7.0, designed for speakers of other languages learning English, was selected for this study. Another version of the *SILL* (Version 5.1) contains 80 items and was originally designed for native English speakers learning other languages, although it has also been extensively used in translation in a variety of settings.
3. For the current Korean translation of Version 7.0, we conducted a new exploratory factor analysis to determine whether the underlying factor structure was similar to that found for the *SILL* in earlier studies. We found that the factor structure was very close to the *SILL* factors reported elsewhere. For a summary of such studies, see Oxford (1996).
4. Just as in most countries, it was not possible to obtain a standardized external criterion measure of English achievement or proficiency for Korean students across all educational levels: middle school, high school, and university. Existing English assessments were not comparable across levels, and statistically creating standard scores for comparison would not begin to address that issue.

### **References**

- Allwright, D. (1990). *Autonomy in language pedagogy. CRILE Working Paper 6*. Centre for Research in Education, University of Lancaster, U.K.
- Baker, L., & Brown, A. L. (1984). Metacognitive skills and reading. In P.D. Pearson, R. Barr, M. Kamil, & P. Mosenthal (Eds.), *Handbook of reading research* (pp. 353-394). New York: Longman.
- Behling, O., & Law, K. S. (2000). *Translating questionnaires and other research instruments: Problems and solutions*. Thousand Oaks: Sage.
- Carrell, P. L. (1989). Metacognitive awareness and second language reading. *Modern*

- Language Journal*, 73, 121-133.
- Carrell, P. L. (1991). Second language reading: Reading ability or language proficiency? *Applied Linguistics*, 12, 159-179.
- Carrell, P. L., Gajdusek, L., & Wise, T. (1998). Metacognition and ESL/EFL reading. *Instructional Science*, 26, 97-112.
- Chamot, A. U. (1998). Teaching learning language strategies to language students. (Report No. EDO-FL-025-976). ERIC Clearinghouse on Language and Linguistics: Washington, D.C. (ERIC Document Reproduction Service No. ED433719).
- Chamot, A. U., & El-Dinary, P. B. (1999). Children's learning strategies in immersion classrooms. *Modern Language Journal*, 83(3), 319-341.
- Cohen, A. D. (1995). Second language learning and use strategies: Clarifying the issues. (Report No. EDO-FL-023-639). ERIC Clearinghouse on Language and Linguistics: Washington, D.C. (ERIC Document Reproduction Service No. ED393307).
- Cohen, A. D. (1998). *Strategies in learning and using a second language*. Essex, U.K.: Longman.
- Cohen, A. D., & Hosenfeld, C. (1981). Some uses of mentalistic data in second language research. *Language Learning*, 31, 285-313.
- Dadour, E. S., & Robbins, J. (1996). University-level studies using strategy instruction to improve speaking ability in Egypt and Japan. In Oxford, R. L. (Ed.), *Language learning strategies around the world: Cross-cultural perspectives* (pp.157-166). Manoa: University of Hawaii Press.
- Dörnyei, Z. (1995). On the teachability of communication strategies. *TESOL Quarterly*, 29(1), 55-85.
- Dreyer, C., & Oxford, R. L. (1996). Learning strategies and other predictors of ESL proficiency among Afrikaans-speakers in South Africa. In Oxford, R. L. (Ed.), *Language learning strategies around the world: Cross-cultural perspectives* (pp.61-74). Manoa: University of Hawaii Press.
- Green, J., & Oxford, R. L. (1995). A closer look at learner strategies, L2 proficiency, and gender. *TESOL Quarterly*, 29, 261-297.
- Hsiao, T-Y., & Oxford, R. L. (2002). Comparing theories of language learning strategies: A confirmatory factor analysis. *Modern Language Journal*, 86(3), 368-383.
- Kaylani, C. (1996). The influence of gender and motivation on EFL learning strategy use in Jordan. In Oxford, R. L. (Ed.), *Language learning strategies around the world: Cross-cultural perspectives* (pp.75-88). Manoa: University of Hawaii Press.

- Kim, H. (2000). Learner variables and their relationships to English achievement of Korean high school students. *Foreign Language Education*, 7(2), 233-258.
- Lan R., & Oxford, R. L. (2003). Language learning strategy profiles of elementary school students in Taiwan. *International Review of Applied Linguistics in Language Teaching*, 41, 339-379.
- Lee, H. (1994). Investigating the factors affecting the use of foreign language learning strategies and comparing the strategy use of EFL students. *English Teaching*, 48, 51-99.
- Lee, H. (2000). A relationship between English language learning strategies and cloze test. *Journal of English Language Teaching*, 12(1), 247-270.
- Lee, H., & Oh, J. (2001). The relationship between attitudes and proficiency in learning English. *English Teaching*, 55(4), 389-409.
- Leki, I. (1995). Coping strategies of ESL students in writing tasks across the curriculum. *TESOL Quarterly*, 29(2), 235-260.
- Little, D. (1991). *Learner autonomy 1: Definitions, issues, and problems*. Dublin: Authentik.
- Mullins, P. (1992). *Successful English language learning strategies of students enrolled in the Faculty of Arts, Chulalongkorn University, Bangkok, Thailand*. Unpublished doctoral dissertation, United States International University, San Diego, CA.
- National Capital Language Resource Center (1996). *Language learning strategies use by secondary foreign language students* (EDO-FL-026-393). Washington, D.C.: Department of Education.
- National Capital Language Resource Center (2000). *High school foreign language students' perception of language learning strategies use and self-efficacy* (EDO-FL-026-388). Washington, D.C.: Department of Education.
- Nyikos, M. & Oxford, R. L. (1993). A factor-analytic study of language-learning strategy use: Interpretations from information processing and social psychology. *Modern Language Journal*, 77, 11-22.
- Oh, J. (1992). Learning strategies used by university EFL students in Korea. *Language Teaching*, 1, 23-53.
- Oh, J. (1996). The effects of attitude and sex on use of EFL learner strategies. *English Teaching*, 51(2), 35-53.
- O'Malley, J. M., & Chamot, A. U. (1990). *Learning strategies in second language acquisition*. Cambridge: Cambridge University Press.
- Oxford, R. L. (1990). *Language learning strategies: What every teacher should know*.

Boston: Heinle & Heinle.

Oxford, R. L. (1994). Language learning strategies: An update. *ERIC Digest*. (Report No. EDO-FL-95-02). ERIC Clearinghouse on Language and Linguistics: Washington, D.C. (ERIC Document Reproduction Service No. ED376707).

Oxford, R. L. (1996). Employing a questionnaire to assess the use of language learning strategies. *Applied Language Learning*, 7, 25-45.

Oxford, R. L., & Burry-Stock, J. A. (1995). Assessing the use of language learning strategies worldwide with the ESL/EFL version of the strategy inventory for language learning (SILL). *System*, 23(1), 1-23.

Oxford, R. L., & Ehrman, M. E. (1995). Adults' language learning strategies in an intensive foreign language program in the United States. *System*, 23(3), 359-386.

Oxford, R. L., & Nyikos, M. (1989). Variables affecting choice of language learning strategies by university students. *Modern Language Journal*, 73, 291-300.

Oxford, R. L., Nyikos M., & Ehrman, M. E. (1988). Vive la difference? Reflections on sex differences in use of language learning strategies. *Foreign Language Annals*, 21(4), 321-328.

Oxford, R. L., Cho, Y., Leung, S., & Kim, H. (2004). Effect of the presence and difficulty of task on strategy use: An exploratory study. *International Review of Applied Linguistics*, 42, 1-47.

Park, J. (2001). Korean EFL learners' vocabulary learning strategies. *English Teaching*, 56(4), 3-30.

Park, Y. (1999). An analysis of interrelationship among language learning strategies, learning styles, and learner variables of university students. *English Teaching*, 54(4), 281-308.

Park, K., Kim, H., & Park, H. (2003). *Language learning strategies*. Seoul: Kyobo-Thompson.

Phillips, V. (1991). A look at learner strategy use and ESL proficiency. *CATESOL Journal*, 57-67.

Pressley, M. (2000). What should comprehension instruction be the instruction of? In M. L. Kamil, et al. (Eds.), *Handbook of reading research* (pp.545-561). Mahwah, NJ: Lawrence Erlbaum Associates.

Pressley, M., Goodchild, F., Fleet, J., Zajchowski, R., & Evans, E. D. (1989). The challenges of classroom strategy instruction. *Elementary School Journal*, 89(3), 301-342.

- Rubin, J. (1994). A review of second language listening comprehension research. *Modern Language Journal*, 78(2), 199-221.
- Touba, N. (1992). *Language learning strategies of Egyptian student teachers of English*. Paper presented at the Twelfth National Symposium on English Teaching in Egypt, Nasr City, Cairo, Egypt.
- Watanabe, Y. (1990). *External variables affecting language learning strategies of Japanese EFL learners: Effects of entrance examination, years spent at college/university, and staying overseas*. Unpublished master's thesis, Lancaster University, Lancaster, UK.
- Wharton, G. (2000). Language learning strategy use of bilingual foreign language learners in Singapore. *Language Learning*, 50(2), 203-243.
- Yoon, W., Won, M., & Kang, H. (2001). The effects of metacognitive strategy training for reading in EFL. *Journal of English Language Teaching*, 13(2), 203-221
- Zimmerman, B., & Martinez-Pons, M. (1990). Student differences in self-regulated learning: Relating to grade, sex, and giftedness to self-efficacy and strategy use. *Journal of Educational Psychology*, 82(1), 51-59.



## **Using Dictation to Improve Language Proficiency**

**Mohammad Rahimi**  
*Shiraz University*

### **Bio Data:**

Dr. Mohammad Rahimi (PhD in TEFL) is currently a full-time assistant professor at the Department of Foreign Languages and Linguistics of Shiraz University, Iran. He teaches both graduate and undergraduate courses in EFL teaching, language testing, research methods, and writing. His areas of interest are research in testing and teaching writing, listening, and speaking skills.

### **Abstract**

Dictation, though widely known as a testing device, can be considered a good learning technique to improve learners' proficiency. This study investigates the effect of dictation with 65 EFL learners to see if it helps them improve their language proficiency. To this end, an experimental group (EG) and a control group (CG) were selected. They both took a proficiency test as a pre-test. All the conditions, especially teaching materials, were the same for both groups except that the EG had regular practice with dictation for one semester. The post-test was administered in a semester after the experiment had terminated. The results showed that both groups showed improvement in their performance on the whole proficiency test. However, with respect to the components of the test, the EG showed improvement in grammar, vocabulary, reading, and listening comprehension, while the control group showed improvement just in the vocabulary section of the test.

### **I**

### **Introduction**

Dictation has been widely known as a testing device for measuring language proficiency. However, using dictation as a language testing technique has suffered huge oppositions. One of the strongest claims made against dictation has been offered by Lado (1961). He considers dictation as having very little value in testing the language. He believes that since everything is dictated to the learner, dictation cannot measure any aspect of the language. Harris (1965), too, rejects dictation as an "uneconomical" and "imprecise" technique of language testing. Somaratne (1957) goes even further and considers it as just a test of spelling. Finally, Stanfield (1985) rejects using dictation as a test of foreign language proficiency due to its strong association with the Grammar Translation Method.

On the other hand, the proponents of dictation advocate it as a good device for measuring language proficiency, due to the high correlation it has shown with other integrative language proficiency tests such as cloze and TOEFL (Oller 1971 and Irvine *et*

*al.*, 1974). Besides, Oller and Streiff (1975) believe that as dictation triggers the internalized expectancy grammar of the learner, it can give us information about his overall language proficiency.

With respect to its teaching value, a few people believe that using dictation in class and practicing with it has nothing to do with the improvement of the learners' proficiency. For instance, Cartledge (1968) believes that dictation is not a teaching device. He claims that it can merely help learners have some practice in oral comprehension. Although Valette (1964) considers dictation as a method of both testing and learning, she believes the fact that learners are proficient in dictation does not necessarily mean that they are proficient in other aspect of language learning.

Nonetheless, research shows that dictation can help learners improve different aspects of their proficiency. For instance, Sawyer and Silver (1961) believe that dictation is primarily a learning exercise and only secondarily a testing device. They assert;

It is true that such a dramatic change does not mean that the student's control of his new language is proportionately that great. He has, however, learned to listen, to concentrate, to write from dictation; he has become familiar with the teacher's particular voice quality; but these abilities are also part of learning a language. In acquiring them he has learned some part of his second language (p. 37).

Morris (1983), too, drawing on the mistakes made by EFL learners on three dictations, concludes that dictation is a technique which can be used both as a testing technique and, more importantly, as a learning activity which helps students develop their accuracy in listening and writing and reinforces structure and vocabulary.

Moreover, Whitaker (1976) considers dictation as a good teaching device. He says, "... wherever aural comprehension is prized, together with literacy and ability to read the FL, dictation may be found to be both profitable for teaching, and valid for testing" (p. 92). Similarly, Pappas (1977) considers dictation as a good means of developing learners' listening comprehension.

Valette (1964) asserts that practice with dictation can help students learn the language. She believes the teacher's concentration on different component of the language, including sound, sentence structure, etc, encourages the students to correct their papers. In fact, she believes dictation can stimulate the students' awareness of the written language.

Some other researchers believe that dictation can be utilized as a teaching technique due to the amount of input it provides for the learner. Stanfield (1985) mentions that since in dictation the learner employs more than one faculty, he is more successful in

internalizing the language. Furthermore, Neisser (1967) considers dictation as an effective technique since the language learner goes through two cognitive processes, i.e., synthesis and analysis. Likewise, Jafarpur and Yamini (1993) claim that, "Dictation allows the language learner to both comprehend and produce the language in the context of meaningful discourse" (p. 360). Accordingly, they call dictation "dual-access processing" because "in taking dictation the subject alters and harmonizes his perception, conception, and expression." Nevertheless, the results of their study with Iranian EFL learners show that practice with dictation cannot improve learners' proficiency. However, as they mention, since language learning is a process which improves over time and needs an incubation period before any learning can be seen in the performance of the learners, the results of their study may not hold for the impact of practice with dictation when longer period of time is concerned.

The incubation period was first pointed out by the proponents of the natural approach to language teaching in such methods and approaches as Total Physical Response, and Natural Approach. In fact, they believe in the priority of listening to speaking in language learning. That is, we should wait for the language to which the learners are exposed to, to be internalized and then show its effects in the learners' production. This incubation or silent period is generalized to other areas of language learning. According to Alwright (1984) and Lightbown (1983), any teaching or learning technique does not immediately show its effect, so we should wait for some time (the incubation period) to see the effect of teaching and learning in the learners' performance.

Therefore, there seems to be an urgent need for another study to avoid the problem of the short time length between practicing with dictation and the post test given to the students to investigate the effect of dictation on the improvement of EFL learners' proficiency. This study is an attempt to deal with this issue.

## **Method**

### *Subjects*

The subjects of the study consisted of 65 English majors enrolled in Conversation II classes. They took part in two separate classes: one taken as the experimental group (EG) consisted of 34 students and the other, the control group (CG), consisted of 31 students. They varied in age from 18 to 25 years with an average of 22.

### *Instrumentation*

The main instrumentation of the study was Shiraz University Placement Test. It consisted of three sections: grammar, vocabulary, and reading comprehension with 60, 40, and 40 items, respectively. In addition, since the test did not contain any listening comprehension items, 25 listening items selected from Sharpe (1995) were added to it. Altogether, the test included 165 multiple-choice items (the test format was like the TOEFL). It was used as both pre-and post-tests to measure the subjects' overall language proficiency.

In addition, the subjects took a listening comprehension multiple-choice test and a dictation consisting of 150 words as their final exam. In the former they listened to an unheard excerpt from Soars and Soars (2001) followed by some multiple-choice questions constructed by the researcher. The latter was excerpted from Hartley and Viney (1995) played for the examinees in the form of a systematic dictation. The scores of these tests, however, were not included in the analysis since they did not contribute any information to the study.

### **Procedure**

In the first session of the conversation class, all the subjects took the proficiency test as the pre-test. Except for the listening comprehension part, there was no time limit for the test so that all the students would finish the test and a complete evaluation of their proficiency would be made.

The conversation class met three one-and-a-half-hour sessions a week for a total of 17 weeks - 51 sessions on the whole. The objective of the course was to enable the students to produce and understand spoken English at an intermediate level. That is, according to the syllabus, they must be able to use English at a comprehensible level while interacting with the teacher and each other rather fluently in spite of the mistakes in grammar, pronunciation, and vocabulary choice. With respect to the listening ability, too, they must be able to have a thorough, though not detailed, understanding of the spoken English in ordinary situations and contexts. For instance, when they listen to a dialog or a short lecture, they must be able to understand the whole message and be able to answer some global and general questions on them.

Class activities were divided into two parts: each week, in the first and the second session, they practiced *New Interchange* and *New Headway*. In the third session, the students gave lectures on already-determined and agreed-upon topics. Part of their course grades were based on these lectures. Both sections followed the same syllabus.

The EG, in addition, took a total of 50 dictations during the semester--each session, one dictation. The passages were selected from different sources, but mainly from books 2, 3, and 4 of *Lado English Series* (Lado, 1990). The reason was that the texts increase in length and difficulty levels as we move from one unit to another and from one book to another. The length of the passages ranged from 50 to 150 words. In each session the difficulty level and the length of the passages were increased. Thirty passages were dictated by the researcher; 20 others were read by other teachers, tape recorded, and were then played for the subjects. Each dictation exercise took about 10 minutes at the beginning of each session. It was read three times. First, it was read at a normal speed while the students just listened. Then, it was read in chunks with pauses between each chunk and the students wrote down what they heard. Finally, it was read at normal speed again. This time the students reviewed their writing to restore any missing parts.

The students' writings were collected and corrected by the researcher and were given back to the students in the next session. The missing words and misspelled ones which changed the meaning of the words were considered mistakes. For instance, writing *peace* for *piece* was considered as a mistake but not *peece* for *peace*. For each mistake one point was subtracted from the total number of the words the passage contained. In the next session, before starting a new dictation, the corrected ones were returned to and reviewed with the subjects. In the review, the words that had been missed by the majority of the students were discussed. This procedure was followed since dictation was used as a teaching device and the purpose was for the students to learn from the dictations they had written and the errors they had made.

The students were asked to guess what part of speech the missing word had been. Then they were asked to say what the word would be with respect to the meaning of the whole text and the sentence it appeared in. In some occasions, they were asked to use reading skills, such as recognizing cohesion and coherence to guess the missing words/phrases. For instance, in the sentence "*Then, he would blow the..... off,*" they were asked to guess what part of speech the missing word was. They used such clues as the existence of the article *the* before the missing word and its appearance after a verb to conclude that they needed a noun. Then, the researcher made them think what things are blown off. Finally, since the topic was on "*a birthday party,*" they managed to find that the missing word was "*candles.*"

If they had made a spelling mistake due to the wrong recognition of the word while listening, some pronunciation practice was also done. For example, if they had written

“sing” for “hing,” due to lack of the existence of the sound /θ/ in Persian, this sound and the words in which it appeared were presented and practiced in class.

In order to compensate for the time spent on the dictations, the EG class was held 10 minutes longer than the CG class. The subjects were not informed of the objectives of the study.

After a lapse of one semester, at the end of the next semester (after about 10 months), the same proficiency test--the post-test this time--was given to the same subjects. Since the test was relatively long and the time interval between its two administrations was long enough it was very unlikely that the students remembered the contents of the test from the first administration.

#### *Data analysis*

The results of the tests were subjected to the following statistical procedures. Two independent t-tests were run for the difference between the mean scores of the EG and the CG on the pre-test as well as the post-test. A paired t-test was run to find the difference between the means of the scores on the following tests: the pre- and post- proficiency tests for the CG as well as for the EG to see if there is any difference between the performance of the subjects on the pre- and post-tests; different components of the pre- and post-tests for the CG as well as for the EG to see if there is any significant difference between the performance of the subjects on different components of the pre-test and the post-test. In addition, an independent t-test was run to find the difference between the means of the scores of the subjects in the EG and the means of the scores of the subjects in the CG in the pre-test; the same test was run in the case of the post-test, too. Furthermore, scores on the whole pre- and post-tests and their components for both groups were subjected to Kuder-Richardson 21 to estimate their reliability coefficients. In addition, true mean gains were calculated for the pre- and post-proficiency tests and their components for both groups to see if there is any difference between the gains of the subjects in the CG and the EG. Finally, an ANOVA test was run for the mean gains of the subjects in the components of the post-test.

#### **Results and discussions**

The mean scores of CG and EG on the pre-test were compared to see if they were the same or different before the experiment started. Table 1 shows the result of t-test.

Table 1  
Independent t-test for EG and CG on pre-test

Mean Difference between EG and CG in Pre-test	CG	EG	df	t	Sig
	83.8	88.38	63	-1.01	0.31

As evident in the table, there was not significant difference between the two groups before the experiment started and the two group started with the same proficiency level.

Table 2 illustrates the results of t-test for the performance of the two groups on the post test.

Table 2  
Independent t-test for EG and CG on post-test

Mean Difference between EG and CG in post-test	CG	EG	Df	t	Sig
	88.38	114.32	63	-8.13	0.00

As the results indicate, EG significantly did better than CG in the post-test. In other words, the experimental group who had regular practice with dictation made more improvement in their general proficiency than the control group. Since the two groups of students had almost the same classes during the time interval between the pre-and post-test, one can claim that the difference in their performance on the proficiency test was due to the fact that the experimental group had practice with dictation.

In order to see how much improvement each group had from the pre-test to the post-test, first descriptive statistics and then a matched t-test were conducted. Table 3 shows descriptive statistics for the results of the pre- and post-tests for the CG.

Table 3  
Descriptive statistics for the results of the pre-and post-tests for the CG

Post-test statistics	Pre-test					Post-test				
	Whole test	Gr	Voc	R	L	Whole test	Gr	Voc	R	L
K	165	60	40	40	25	165	60	40	40	25
$\bar{X}$	83.8	38.7	17.15	16.3	11.65	89.8	39.2	21.25	18.55	12.15
SD	19.57	5.82	6.39	6.36	4.88	20.28	5.09	6.72	7.43	5
Range	60	19	24	22	16	67	19	24	25	18

	(58-118)	(30-51)	(4-28)	(8-30)	(4-20)	(65-132)	(30-51)	(8-32)	(8-33)	(5-23)
Skewness	0.43	0.27	-0.2	0.55	0.53	0.71	0.29	-0.18	0.33	0.69
Kurt.	-0.99	-0.45	-0.52	-0.16	-0.77	-0.63	0.2	-0.52	-0.7	-0.4
KR-21	0.9	0.6	0.77	0.77	0.75	0.9	0.49	0.79	0.83	0.78

The mean of the pre-test for the CG is 83.8, and that of the post-test is 89.8. The standard deviation for the pre-test is 19.57, and that of the post-test 20.28. Besides, the range of the scores of the pre-test is 60, whereas that of the post-test is 67. Furthermore, the distribution of the scores of both pre- and post-tests is positively skewed, though the latter is more positively skewed (0.43 and 0.71). In addition, the distribution of the scores obtained from both tests is flat (-0.99 and -0.63). Finally, the reliability coefficients of both tests (measured through KR-21) are the same, i.e., 0.90.

With respect to the components of the pre-test and the post-test, in all the cases the means of the components of the post-test show some increase as compared with the pre-test. In the case of grammar, for instance, the mean for the pre-test is 38.7 and that of the post-test 39.2. With respect to vocabulary, too, the mean of the post-test is larger than that of the pre-test (21.25 and 17.15, respectively). The mean scores of reading comprehension and listening comprehension also show the same pattern.

Regarding the standard deviation, one cannot see a remarkable difference between the components of the two tests. The same is true with regard to the range of the scores of different components of the two tests. Concerning the distribution of the scores on the components of the tests, except for the vocabulary part whose distribution is negatively skewed in both pre- and post-tests (-0.2 and -0.18, respectively), in all other cases one can see positive skewness. This shows that in both pre- and post-tests the vocabulary part has been the easiest. Finally, as the kurtosis values for different components of the tests show, except for the vocabulary part, a large difference can be seen in the case of the other parts. That is, with respect to grammar, as it is evident in the table, the distribution of the scores in the pre-test is flat and that of the post-test is rather peaked (-0.45 and 0.2, respectively). With regard to listening comprehension and reading comprehension, though in both pre- and post-tests the distribution of the scores is flat, the degree of peakedness largely differs. In the case of listening comprehension it is flatter in the pre-test (-0.77) than in the post-test (-0.4), whereas it is quite the opposite in the case of reading comprehension (-0.16 in the pre-test and -0.7 in the post-test).

Table 4  
Difference between the means of the pre- and post-tests and their components in the CG

	The whole tests	Gr	Voc	R	L	The whole pre-and post-test excluding voc.
Difference between the means	-2.12 (s)	-0.57 (ns)	-3.33 (s)	-1.80 (ns)	-0.73 (ns)	-0.62 (ns)

According to Table 4, there is a significant difference between the results of the pre- and post-tests for the CG. However, looking at the components of the two tests, one will see that, except for the vocabulary part, there is no significant difference between the results of the two tests. Thus, one may doubt whether it is because of the scores on the vocabulary part that there exists a significant difference between the two tests. The hypothesis was verified when a t-test was run between the two tests excluding the vocabulary part. As it can be seen in the table, there is no significant difference between the two tests in this case.

Nevertheless, the results are quite different in the case of the EG. The descriptive statistics for the pre- and post-tests for this group are presented in Table 5.

Table 5  
Descriptive statistics of the pre- and post-tests for the EG

Post-test statistic	Pre-test									
	Whole test	Gr	Voc	R	L	Whole test	Gr	Voc	R	L
K	165	60	40	40	25	165	60	40	40	25
$\bar{X}$	88.38	41.41	20.25	17.87	12	114.32	43.87	27.83	28.19	13.74
SD	21.58	5	5.62	7.64	4.1	13.71	4.31	5.2	6.36	3.52
Range	121 (11-132)	23 (26-49)	23 (10-33)	30 (4-34)	14 (6-20)	61 (82-143)	21 (33-54)	25 (13-38)	24 (13-37)	13 (9-22)
Skewness	-1.39	-1.25	0.31	-0.009	0.29	-0.4	-0.19	-0.42	-1.19	0.47
Kurt	4.95	1.98	0.29	-0.67	-1.17	0.37	0.86	1.26	0.86	-0.6
KR-21	0.92	0.5	0.70	0.84	0.65	0.82	0.37	0.70	0.80	0.52

According to this table, the mean for the pre-test is 88.38, whereas that of the post-test

is 114.32. There is a relatively high difference between the standard deviations of the two tests: 21.58 in the case of the pre-test and 13.71 for the post-test. This shows that the subjects have performed closer to homogeneously on the post-test. This claim can be verified referring to the large difference between the range of the scores on the two tests (121 for the pre-test and 61 for the post-test). Furthermore, the distribution of the scores on the two tests is negatively skewed (-1.39 and -0.19), though that of the post-test tends to be more normal. Besides, the distribution of the scores of the two tests is peaked (4.95 and 0.37). However, that of the post test seems to be much closer to a normal distribution.

Regarding the components of the two tests, one can see an increase in the means in all the cases. The highest increase, however, can be seen in the case of the reading part, i.e., from 17.87 to 27.19. Concerning the standard deviation, on the other hand, one can see a relative decrease in all the cases. With regard to the range of the scores, except for the vocabulary part which shows some increase from the pre-test to the post-test--23 to 25--, in all other cases the range has decreased, indicating the fact that the subjects have performed more homogeneously on the subtests. As for the skewness, rather large differences can be seen between the components of the pre- and post-tests. In the case of grammar, while the distribution of the scores of both tests is negatively skewed (-1.25 in the case of the pre-test and -0.19 in the case of the post-test), that of the post test is much closer to the normal distribution. With respect to vocabulary, the difference is more remarkable. That is, the distribution of the scores in the pre-test is positively skewed (0.31), whereas that of the post-test is negatively skewed (-0.45), indicating the fact that the latter has been easier. Similarly, with regard to the reading comprehension part, one can see that the distribution of the scores tends to be more negatively skewed in the post-test. As it is evident in the table, the distribution of the scores in the pre-test is very much close to the normal distribution (-0.01), whereas it is rather highly negatively skewed in the post-test (-1.19). Finally, the distribution of the scores on listening comprehension of both tests is positively skewed (0.29 and 0.47), implying that this part has been difficult both in the pre- and post-tests. Concerning the peakedness of the distribution of the scores, the only component of the test which shows a difference in the pre- and post-tests is reading comprehension. That is, the distribution of the scores in the pre-test is flat (-0.67), while it has become peaked (0.86) in the post-test. This shows that the scores on the post-test have been much closer to the mean. In all other cases, the peakedness has remained unchanged in spite of some changes in the degree.

Table 6 represents the difference between the means of the two tests and their different

components for the EG.

Table 6

Difference between the means of the pre- and post-tests and their components for the EG

	The whole tests	Gr	Voc	R	L
Difference between the means	-7.25 (s)	-4.13 (s)	-9.33 (s)	-7.45 (s)	-3.32 (s)

The results show that there is a significant difference between the means on the pre- and post-tests and all their components. Since the mean scores of the subjects in all the cases have increased from the pre-test to the post-test, it can be claimed that practice with dictation has had a positive effect on the improvement of the proficiency of the subjects, as compared with the results of the tests taken by the CG where one can see such an increase only in the vocabulary part. The reason for the increase observed in both cases can be the fact that since the subjects majored in English Translation and in most of their courses much emphasis was laid on developing their vocabulary, after two semesters, both groups' knowledge of vocabulary has improved rather equally.

The hypothesis that practice with dictation has probably been the major factor improving the proficiency level of the subjects in the EG can be verified by two pieces of evidence. Firstly, according to Table 5 (below), there is no significant difference between the scores of the subjects in the EG and the CG on the pre-test. In the post-test, however, a significant difference can be observed between the scores of the two groups. Since the mean of the scores for the EG is larger than that of the CG, one can truly claim that practice with dictation has helped students improve their proficiency in the English language.

Table 7

Difference between the means of the scores of the EG and the CG

Difference between the means in the pre-test	Difference between the means in the post-test
0.78 (ns)	4.75 (s)

Another piece of evidence supporting this claim is the difference between the gains of the two groups. Table 8 presents the gains of the subjects in the EG and the CG.

Table 8  
The mean gain of the subjects in the EG and the CG

EG							
CG							
Statistics	The whole test	Gr	Voc	R	L	The whole test	Voc
$\bar{X}$	25.93	2.45	7.55	10.28	1.67	4.06	4.05
SD	17.11	5.47	4.33	5.4	3.83	4.71	8.04
Range	103.32 [99.27-(-4.05)]	26.97 [16.8-(-10.17)]	20.68 [18.07-(-2.61)]	22.46 [21.01-(-0.55)]	14.06 [7.39-(-6.68)]	15.79 [11.75-(-4.04)]	35.1 [28.61-(-6.49)]

As presented in the table, the true mean gain obtained by the subjects in the EG is much larger than that of the CG (25.93 and 4.06, respectively). In other words, though both groups have had some gain from the pre-test to the post-test, that of the EG who have made practice with dictation is much larger. The same is true regarding the vocabulary part. That is, the gain in the EG is approximately twice as much as that of the CG. With regard to other components of the test, since the difference between the pre- and post-tests in the CG was not significant, it was not possible to measure any gain.

Finally, in order to see if there is any difference between the mean gain of the subjects in the EG on different subtests, an ANOVA test was run. The results represented in Table 9 show that there is a significant difference among the mean gains obtained on different parts of the test

Table 9  
ANOVA test for different components of the post-test for the EG

Source	D F	Sum of sq	Mean sq	F Ratio	F Prob
Between groups	3	1587.84	529.28	22.82	0.00
Within groups	120	2782.12	23.18		
Total	123	4369.96			

Table 10  
Scheffe test for the components of the post-test for the EG

Mean gains		L	Gr	Voc	R
1.64	L				
2.45	Gr				
7.55	Voc	*	*		
10.28	R	*	*		

According to Table 10 which presents the results of the scheffe, except for reading and vocabulary on the one hand and listening and grammar on the other, the means on all other components show significant differences. That is, there is a significant difference between the mean gain of the subjects on vocabulary and listening and grammar. The same holds true concerning reading comprehension and vocabulary and grammar. Thus, since the mean gain in vocabulary and reading comprehension has been larger than the grammar and listening comprehension, one can claim that practice with dictation has improved the subjects' vocabulary and reading comprehension more than their grammar and listening comprehension.

### **Conclusions**

This study investigated the effect of practice with dictation on the learners' proficiency. There were an experimental and a control group. The experimental group took a dictation in every session of their class for a period of one semester. Each session, the researcher practiced the corrected dictation with them and elaborated on grammatical, phonological, and lexical points in the dictated passage. The results of the study showed that the experimental group made some improvement in their language proficiency. The control group, however, showed no improvement on the proficiency test and its components except for the vocabulary subtest.

The results of the study, to some extent, support those of Valette (1964) and Morris (1983), since they claimed that practice with dictation helps learners improve their language proficiency. However, they mostly emphasized the improvement in learners' listening and writing skills, whereas in the present study listening comprehension and grammar turned out to be the language components least improved and reading and vocabulary the ones most improved.

The results of the study, on the other hand, contradict those of Jafarpur and Yamini (1993) in that they saw no improvement in the proficiency of the experimental group as compared with the control group. Nevertheless, since they administered the post-test at the end of the semester during which the experiment was conducted (after about 4 months), the 'incubation hypothesis' might have been partly disregarded (as they themselves stated). The results of the present study lend support to the above hypothesis which implies that a teaching technique may show its effects in the long run.

This study, nonetheless, suffers from a couple of shortcomings. Firstly, since the post-test was given a long time after the pre-test, i.e., after about 10 months, some of the

subjects might have improved their proficiency via extracurricular activities, e.g., attending language institutes and reading short stories. The low number of the subjects participating in the study would amplify the effect of such activities on the results of the study.

Secondly, the pre- and post-tests measured just the learners' recognition rather than production. Including tests of writing and speaking in the pre- and post-tests could have shown if practice with dictation has any effect on the improvement of such skills or not.

Finally, as mentioned above, the class for EG was held about 10 minutes longer than that of CG. Though it was inevitable, as the researcher had to ensure that everything else was the same for the two classes except for the dictation, this longer time for EG might have contaminated the results. In fact, the better performance of EG on the post-test might be partly due to having more language practice as compared to CG.

## **References**

- Allwright, R.L. (1984). The importance of interaction in classroom language learning. *Applied Linguistics*, 5, 156-171.
- Cartledge, H. A. (1968). A defence of dictation. *ELT Journal*, 22, 226-231
- Harris, D.P. (1965). *Testing English as a second language*. New York: McGraw-Hill.
- Hartley, B., & Viney, P. (1995). *New American streamline: Connections*. Oxford: Oxford University Press.
- Irvine, P., Atai, P., & Oller, G.W. (1974). Cloze, dictation, and the test of English as a foreign language. *Language Learning*, 245-254.
- Jafarpur, A. & Yamini, M. (1993). Does practice with dictation improve language skills? *System*, 21, 359-369.
- Lado, R. (1961). *Language testing*. New York: McGraw-Hill.
- Lado, R. (1990). *Lado English series*. New Jersey: Prentice-Hall.
- Lightbown, P.M. (1983). Exploring relationships between developmental and instructional sequences in L2 acquisition. In H. W. Selinger & M. Long, (Eds.), *Classroom-oriented research in second language acquisition*, pp. 217-245. Rowley, MA: Newbery House.
- Lord, E.M. (1963). Elementary models for measuring change. In C. W. Harris (Ed.), *Problems in measuring change*. Madison: University of Wisconsin.
- Morris, S. (1983). Dictation-a technique in need of reappraisal. *ELT Journal*, 37, 121-126.

- Neisser, U. (1967). *Cognitive psychology*. New York: Appleton-Century Crofts.
- Oller, J. W. (1971). Dictation as a device for testing foreign-language proficiency. *ELT Journal*, 25, 254-259.
- Oller, J. W. & Streiff, V. (1975). Dictation: A test of grammar-based expectancies. *ELT Journal*, 30, 25-36.
- Pappas, G.S. (1977). You mean you still give dictations? *Language Arts*, 54, 936-939.
- Richards, J.C., Hull, J., & Proctor, S. (1997). *New interchange English for international communication*. New York: Cambridge University Press.
- Sawyer, J. & Silver, S. (1961). Dictation in language learning. *Language Learning*, 11, 33-42.
- Sharpe, P.J. (1995). *Barron's how to prepare for the TOEFL test of English as a foreign language*. Woodbury, New York: Barron's Educational Series.
- Soars, J. & Soars, L. (2001). *New headway intermediate*. Oxford: Oxford University Press .
- Somaratne, W. (1957). *Aids and Tests in the Teaching of English*. Oxford: Oxford University Press.
- Standfield, C. W. (1985). A history of dictation in foreign language teaching and testing. *Modern Language Journal*, 69, 121-128.
- Valette, R.M. (1964). The use of the dictee in the French language classroom. *Modern Language Journal*, 48, 431-434.
- Whitaker, S.F. (1976). What is the status of dictation? *Audio-Visual Language Journal*, 14, 87-93.



## **Determinants and issues in student achievement in English at the Lao Secondary Education Level**

**Souvannasy Bouangeune**  
*National University of Laos*  
**Masashi Sakigawa**  
*Hiroshima University, Japan*  
**Yukiko Hirakawa**  
*Hiroshima University, Japan*

### **Bio Data:**

**Souvannasy Bouangeune**, an instructor of English, has been teaching English at the Department of English, Faculty of Letters, National University of Laos for ten years. His major is English Language Teaching. He holds an M.A. in English Language Education from Hiroshima University. Currently, he is a doctoral candidate in the Department of Education and Culture, Graduate School for International Development and Cooperation at Hiroshima University in Japan.

**Masashi Sakigawa** is a doctoral candidate in the Department of Education and Culture, Graduate School for International Development and Cooperation at Hiroshima University in Japan. His major is educational development in Asian countries, especially Vietnam.

**Yukiko Hirakawa** is an Associate Professor in the Department of Education and Culture, Graduate School for International Development and Cooperation at Hiroshima University in Japan. Her fields of interest are educational planning, process of educational improvement and developing higher order cognitive skills. She is now working to improve English education in the Department of English, National University of Laos.

### **Abstract**

English language education was introduced into Lao secondary schools in 1986 under the “Chintanakan Mai” (new way of thinking) policy. Since then, the demand for it has been steadily increasing. This demand, however, has not been matched by a corresponding quality improvement. Although it is widely recognized that Lao secondary school students’ achievement in English is not very impressive, there is no actual data to support this claim. In order to improve Lao secondary school students’ proficiency in English, this research aimed to ascertain the students’ level of English and to identify the factors which might impact on their achievement. For this purpose, an English test and questionnaire was administered to 1170 grade 5 students in secondary schools. This accounts for 3.1% of all grade 5 students in Laos. The results suggest that 1) English language education in Lao secondary schools should focus more on the learning of basic vocabulary, and 2) the content of the textbooks should be more basic.

**Keywords:** English achievement, Influential factors, Secondary education, Laos

## **1. Introduction**

In 1986, Laos adopted the “Chintanakan Mai” (new way of thinking) policy, representing a change from a centrally planned economy and an Eastern bloc-oriented foreign policy toward a market economy and an omni-directional foreign policy. Since the implementation of this policy, an inflow of people and capital from the West has raised the importance of and the demand for English language learning. For instance, our survey found that 81.2% of Lao secondary school students “strongly agree” and another 12.0% “agree” that English is important; when English is compared with other subjects, 37.2% answered that they “work on English harder than any other subject” while 49.1% “work on English as hard as other subjects” and only 13.6% “work on English less hard than any other subject”<sup>(1)</sup>. Currently, many Lao students are anxious to be employed by foreign-affiliated companies, which usually offer salaries more than five times that paid by local companies or public agencies, and where English language skills are often required as a condition of employment. The demand for English is increasing within the government as well, as it seeks to recruit officers with English skills in order to attract more foreign capital for economic development. As a result, English language education was introduced into the Lao secondary school curriculum in 1986. English, together with French and Russian, became one of the few elective foreign languages in the national curriculum, prescribed in 1994 by the National Research Institute for Educational Sciences (NRIES, 1994).

## **2. English language education in Lao secondary schools**

Improvement in the quality of English language education has not kept up with the rapidly increasing demands for English in Lao secondary schools. There is a lack of textbooks, qualified English teachers and also a well-organized curriculum (Goh & Vonchith, 2003). As a result, it is widely believed that student achievement at secondary school level in English is unsatisfactory. However, as a large-scale study of student achievement at secondary school has never been undertaken, a conclusive statement cannot be made at this point.

In the world’s under-developed countries, such as Laos, theories of educational improvement introduced by developed countries are often accepted at face value. Such countries are often required to follow the assistance policies of the donor countries, which may not necessarily reflect their actual situation or needs. These practices impede the determination of the real causes of educational problems and one way of addressing them.

In Laos, Teaching English to Speakers of Other Languages (TESOL) has already been introduced in English classes at some secondary schools. But, according to this study, the employment of TESOL, as one of teaching methods for English, appears to be too difficult for the level of English proficiency achieved by Lao secondary school students.

The relevant literature regarding English education in Lao secondary schools is severely limited, especially compared to other Asian countries. If there is any, the work is not based on findings from field surveys and only describes the history or general situation of English language education in Lao secondary schools (e.g., Appleby et al. 2002; Goh & Vonchith, 2003; Sithirajvongsa & Goh, 2004).

Thus, in order to improve English educational policy in Laos, it is important to initially clarify the actual level of English proficiency of Lao students, and identify the factors which affect student achievement in English with research data.

### **3. Research method**

#### **3.1. Schedule and implementation**

Both the test and the questionnaire survey were conducted in September 2003. Under the supervision of Souvannasy Bouangeune and his four assistants, the tests and the questionnaires were implemented at the sample schools. The assistants were given instructions before arriving in the field; they followed a written manual and used checklists in implementing the survey.

#### **3.2 Sampling areas, schools and students**

In this study, we administered a test at the beginning of the school year to Grade 5 students, who had been learning English for at least four years — from Grade 1 of secondary school to Grade 4, the first year of upper secondary school. The expectation was that, after learning English for four years, students would have reached a certain level of proficiency.

The total numbers of schools in Laos consist of 847 lower-level secondary schools and 270 upper-level secondary schools, with gross enrolment rates in 2002 of 51.9% and 30.0% respectively (Ministry of Education, 2003). Among the 270 upper-level schools, 32 were independent upper secondary schools with students from Grade 4 to 6, and 238 were integrated secondary schools with students from Grade 1 to 6. Both types of schools accept students based on their scores in the provincial achievement test at the end of Grade 3, but the integrated secondary schools also allow their Grade 3 students to advance to

Grade 4 based on school tests.

For reasons of accessibility, the researchers chose Vientiane Municipality and Vientiane Province from the 18 provinces in Laos as survey areas. In Vientiane Municipality, four urban districts out of nine were excluded because the socioeconomic situation and conditions in the schools were very different from other areas in Laos. Even so, the two sample areas were considered to be privileged compared to other provinces, because schools in these areas are assumed to have, in general, more quality teachers. In mountainous provinces, many secondary schools cannot provide students with English as an elective subject due to a lack of qualified teachers.

To control for student aptitude, the researchers excluded the three best schools of the 31 upper secondary schools in the five target districts of Vientiane Municipality (Vientiane Municipality Education Office, 2003) and the two best schools of the 26 schools in Vientiane Province (Vientiane Province Education Office, 2003). These schools attract high-achieving students from large catchment areas and were identified by several officials in the provincial/municipal offices of education. The researchers also excluded those schools that had changed English teachers during the last school year (Grade 4). Schools that had tracking classes were also excluded, because it was difficult to do a random sampling of students in these schools.

As a result, twenty-five schools were selected: 15 in Vientiane Municipality and 10 in Vientiane Province. Twenty-one of the schools were integrated secondary schools and four were independent upper secondary schools. In each school, the researchers selected one Grade 5 class using random sampling. Thus, the sample size in terms of classes was 25. The total number of students was 1,170.

The twenty-five schools comprise 9.3% of all 270 upper secondary schools in Laos, and the 1,170 students make up 3.1% of all Grade 5 students (37,665 in total).

### **3.3. English Test**

The researchers developed a test to measure student achievement and the teachers' knowledge of English. At first, we tried to use a standardized test consisting of grammar, vocabulary and communication (Tanaka & Lauer, 2003a; Tanaka & Lauer, 2003b), with small modifications based on Lao textbooks. However, the results of a pilot study revealed that this test was too difficult for the Lao students. Consequently, the researchers were obliged to almost totally rewrite the test.

The rewritten test had three parts: grammar and sentence patterns (20 items),

vocabulary (20 items), and reading comprehension (5 items). Due to limitations in terms of time and equipment, the test did not include pronunciation. The content of the test was based mainly on Textbook 1 and 2, which were expected to be completed during Grade 1 to 3.

A multiple-choice format offering four answer choices for each question was used. Grammar and sentence patterns included the usage of the object pronoun; *wh*-questions; prepositions; comparative and superlative adjectives; tenses (present simple; present continuous; present perfect; past simple and future with *going to* + infinitive verb); quantifiers, plural nouns; requests with *could*; *yes/no* questions; and the passive voice and verb agreement. Twenty items accounted for one point each. Sixteen of these items appeared in Textbook 1 (NRIES, 1996) and four items in Textbook 2 (NRIES, 1997b).

The vocabulary section consisted of 20 items corresponding to 20 basic vocabulary words. Twelve of the words were selected from amongst the 1,000 most frequently used words of the General Service Word List (West, 1953), two were selected from words 1,001 to 2,000 of the list, and another two from words 2,001 to 2,284. According to Schmitt (2000), a vocabulary of 2,000 words is considered to be necessary for successful daily communication. The four remaining words (“three”, “slim”, “wedding” and “get up”) were not taken from the list. Fifteen of the words appeared in Textbook 1 (NRIES, 1996) and three were from Textbook 2 (NRIES, 1997b). The remaining two words appeared in the earlier chapters of Textbook 3 (NRIES, 1998b).

The reading comprehension section consisted of one simple passage of text (120 words) extracted from Textbook 2 (NRIES, 1997b, p. 115). Five items accounted for two points each.

### **3.4. English test scores**

The analysis treated English test scores as the dependent variable. Table 1 shows the descriptive statistics of the student scores on the test. The mean score was 16.4 out of a full score of 50 (32.8%). Considering the fact that the test was in a multiple-choice format with four answer choices, and that the rate of “no answer” was only 3.6%, the student scores were considered to be low.

Many Lao students have serious problems with basic vocabulary; a mean score of 7.5 out of a full score of 20 (37.5%) reflects this. It is striking that many students, after studying English for four years, still do not know the meaning of basic words like “know” and “water”. Vocabulary is the part of the test that accounted for the largest difference in

student scores, with a standard deviation of 4.0.

Many Lao students also have serious problems with regard to sentence patterns/grammar and reading comprehension tasks. These problems were reflected in a mean score of 6.7 out of a full score of 20 (33.5%) and 2.2 out of a full score of 10 (22%) respectively.

**Table 1: Descriptive Statistics of English Test Scores**

Student score	<u>M</u>	<u>SD</u>	<u>n</u>	Min	Max	Full score
	16.4	7.4	1,170	3	49	50
Sentence pattern/grammar	6.7	2.9	1,170	0	20	20
Vocabulary	7.5	4.0	1,170	0	20	20
Reading comprehension	2.2	2.3	1,170	0	10	10
Mean score of classes	16.0	3.5	25	13.0	24.9	50

*Note:* M (mean). SD (standard deviation). n (number in sample).

### **3.5. Questionnaire and variables**

Educational quantitative researchers look at different factors related to student achievement, depending on their own interests. To reveal the factors related to student achievement from broader viewpoints, combining different kinds of research approaches, rather than focusing on a specific factor, such as parental attitude toward education (a family factor) or learner-centered teaching methods (an educational factor), would be most effective. With this approach, however, the research would fail to provide a detailed description of how a specific factor works in a certain situation. So, to reveal how a specific factor works in a certain situation, a detailed description would be more useful; however, this approach may not provide an overall image of how each factor interrelates with student achievement from different kinds of factors' viewpoint. To make up for this, the results obtained by each kind of research method can be aggregated.

In Laos, where there is a limited amount of extant educational research on English language education, and where the students' main purpose for studying is to obtain high grades in order to gain entrance into institutes of higher education, a broad approach to English achievement which examines the relationship between various factors directly related to student achievement would be most instructive. Therefore, this research examines the related factors by combining three different kinds of factors: family factors (family background), motivation factors (motivation for English study) and educational factors (educational inputs). Previous research on the cause and effect relationship with regard to student achievement has shown that these kinds of factors are important in

assessing student achievement<sup>(2)</sup>.

Questionnaires were distributed to students who took the English test. The questionnaire consisted of 20 questions (Table 2). Nine of the questions related to family background; from the answers, four variables were created. Four of the questions related to motivation for English study; from the answers, four variables were created. The seven remaining questions related to educational inputs; from the answers, five variables were created. The variables created are often cited as basic variables in research on family background, motivation for English study and educational production. An evaluation scale for all 13 variables except for “average score in a class”, was employed for the answered questions<sup>(3)</sup>.

**Table 2: Descriptions, Means, and Standard Deviations for Variables**

Variable	Description	Metric	Mean	SD
<i>Family background</i>				
Family economic indicator	How many items out of the following six does your family have? (1) Car, (2) Motorbike, (3) Color TV, (4) VCR, (5) CD player or cassette player, (6) Radio	0=None of these items 6=All of these six items	2.53	1.57
Father's academic background	What is your father's [highest] academic qualification?	1=Below lower secondary school 6=University	3.70	1.32
Mother's academic background	What is your mother's [highest] academic qualification?	1=Below lower secondary school 6=University	3.00	1.35
Parental interest in study	Do your parents encourage you in your studies?	1=Never 5=Always	2.75	1.13
<i>Motivation for English study</i>				
English learning at home	How often do you study English at home?	1=Almost never 5=Almost everyday	3.40	1.33
Attitude toward English learning	What do you think about your study of English, compared to your other school subjects?	1=Work on English less hard than any other subject 3=Work harder on English than any other subject	1.76	0.67
Enjoyment of English learning	Do you enjoy learning English?	1=Strongly disagree 5=Strongly agree	4.48	0.76
Importance of English learning	Do you think that English is important?	1=Strongly disagree 5=Strongly agree	4.81	0.51
<i>Educational inputs</i>				
Possession of textbooks	How many kinds of the following three school English textbooks do you have? (1) textbook 1 (2) textbook 2 (3) textbook 3	0=None of them 3=All three kinds of textbooks	0.39	0.65
Possession of an English-Lao dictionary	Do you have an English-Lao dictionary at home?	0=No 1=Yes	0.21	0.40
Possession of English books	Do you have English books at home?	0=No 1=Yes	0.40	0.49
TV or radio English teaching program	Have you ever learned English through a TV or radio teaching program?	1=Never 5=Many times	2.33	0.98

Extra English classes or private tuition	Have you ever taken extra English classes or private tuition?	1=Never 4=More than 13 months	1.56	0.79
Average score in a class		13.05=Lowest score 24.88=Highest score	16.44	3.64

#### 4. Data analysis

As mentioned above, based on the answers to the 20 questions in the questionnaire, a total of 16 variables were created. The average score in each class was then added as an independent variable, and English test scores were added as a dependent variable. The data obtained was analyzed statistically, using correlations and a multiple linear regression equation.

##### 4.1. Correlations between the variables

The correlations between each variable, which are deemed to be important for student academic achievement, were determined by means of bivariate correlations. Although bivariate correlations do not necessarily indicate a relationship of cause and effect, they are expected to give an overall indication of which variables are related with which. Table 3 shows that many of the variables, especially the relationship between English test scores and other variables, were positively correlated with each other at a significance level of 5%. It can thus be inferred that these variables — family background, motivation for English study and educational inputs — have some influence on the English test scores achieved<sup>(4)</sup>.

**Table 3: Correlations between the Variables**

Variables		1	2	3	4	5	6	7	8
1. English test score	Pearson correlation	1	-	-	-	-	-	-	-
	Sig.(2-tailed)	.							
	N	1170							
2. Family economic indicator	Pearson correlation	.166**	1	-	-	-	-	-	-
	Sig.(2-tailed)	.000	.						
	N	1170	1170						
3. Father's academic background	Pearson correlation	.091**	.252**	1	-	-	-	-	-
	Sig.(2-tailed)	.002	.000	.					
	N	1133	1133	1133					
4. Mother's academic background	Pearson correlation	.052	.227**	.525**	1	-	-	-	-
	Sig.(2-tailed)	.082	.000	.000	.				
	N	1125	1125	1106	1125				
5. Parental interest in study	Pearson correlation	.095**	.167**	.155**	.159**	1	-	-	-
	Sig.(2-tailed)	.001	.000	.000	.000	.			
	N	1133	1133	1103	1094	1133			
6. English learning at home	Pearson correlation	.145**	.091**	.047	.093**	.265**	1	-	-
	Sig.(2-tailed)	.000	.002	.116	.002	.000	.		
	N	1131	1131	1103	1098	1103	1131		
7. Attitude toward English learning	Pearson correlation	.157**	.056	.078*	.081**	.230**	.291**	1	-
	Sig.(2-tailed)	.000	.062	.011	.008	.000	.000	.	
	N	1106	1106	1076	1072	1079	1082	1106	
8. Enjoyment of English learning	Pearson correlation	.082**	-.008	-.032	-.034	.076*	.149**	.127**	1
	Sig.(2-tailed)	.007	.789	.295	.265	.013	.000	.000	.
	N	1098	1098	1064	1056	1070	1063	1039	1098
9. Importance of English learning	Pearson correlation	.047	.022	-.096**	-.083**	.093**	.117**	.060*	.206**
	Sig.(2-tailed)	.118	.460	.002	.006	.002	.000	.049	.000
	N	1119	1119	1084	1077	1090	1084	1060	1095
10. Possession of	Pearson correlation	.043	.121**	.070*	.057	.108**	.073*	.023	.011

English textbooks	Sig.(2-tailed)	.142	.000	.019	.056	.000	.014	.449	.705
	N	1170	1170	1133	1125	1133	1131	1106	1098
11. Possession of an English-Lao dictionary	Pearson correlation	.189**	.195**	.100**	.114**	.100**	.113**	.115**	.021
	Sig.(2-tailed)	.000	.000	.001	.000	.001	.000	.000	.493
	N	1170	1170	1133	1125	1133	1131	1106	1098
12. Possession of English books	Pearson correlation	.154**	.128**	.073**	.027	.090**	.155**	.078**	-.016
	Sig.(2-tailed)	.000	.000	.014	.372	.002	.000	.010	.587
	N	1170	1170	1133	1125	1133	1131	1106	1098
13. TV or radio English teaching programs	Pearson correlation	.165**	.202**	.153**	.174**	.162**	.248**	.190**	.028
	Sig.(2-tailed)	.000	.000	.000	.000	.000	.000	.000	.373
	N	1106	1106	1077	1067	1078	1077	1053	1039
14. Extra English classes or private tuition	Pearson correlation	.389**	.190**	.152**	.085**	.242**	.286**	.253**	.110**
	Sig.(2-tailed)	.000	.000	.000	.005	.000	.000	.000	.000
	N	1110	1100	1076	1072	1078	1079	1056	1043
15. Average score in a class	Pearson correlation	.492**	.097**	.079**	.054	.017	.069*	.170**	.034
	Sig.(2-tailed)	.000	.001	.008	.071	.557	.020	.000	.264
	N	1170	1170	1133	1125	1133	1131	1106	1098

\*\* Correlation is significant at the 0.01 level (2-tailed).

\* Correlation is significant at the 0.05 level (2-tailed).

**Table 3: Correlations between the Variables (Continued)**

Variables		9	10	11	12	13	14	15
9. Importance of English learning	Pearson correlation	1	-	-	-	-	-	-
	Sig.(2-tailed)	.						
	N	1119						
10. Possession of English textbooks	Pearson correlation	.022	1	-	-	-	-	-
	Sig.(2-tailed)	.463	.					
	N	1119	1170					
11. Possession of an English-Lao dictionary	Pearson correlation	.068*	.012	1	-	-	-	-
	Sig.(2-tailed)	.023	.670	.				
	N	1119	1170	1170				
12. Possession of English books	Pearson correlation	-.023	-.112**	.063*	1	-	-	-
	Sig.(2-tailed)	.448	.000	.031	.			
	N	1119	1170	1170	1170			
13. TV or radio English teaching programs	Pearson correlation	.054	.040	.151**	.181**	1	-	-
	Sig.(2-tailed)	.080	.182	.000	.000	.		
	N	1059	1106	1106	1106	1106		
14. Extra English classes or private tuition	Pearson correlation	.057	.110**	.257**	.160**	.273**	1	-
	Sig.(2-tailed)	.065	.000	.000	.000	.000	.	
	N	1064	1110	1110	1110	1061	1110	
15. Average score in a class	Pearson correlation	-.056	.098**	.087**	.055	.167**	.301**	1
	Sig.(2-tailed)	.063	.001	.003	.060	.000	.000	.
	N	1119	1170	1170	1170	1106	1110	1170

#### 4.2.1 Standardized regression coefficients from multiple regressions of English test scores on variables

To determine which variables strongly affected English test scores, a multiple linear regression equation was employed. English test scores were regressed on variables (Table 4). Six of the seven significant variables related to educational inputs and it appears that educational inputs have a stronger effect on students' English test scores than the other two kinds of variables. The average score in a class, especially, seems to have the strongest effect on English test scores<sup>(5)</sup>.

Extra English classes or private tuition is also likely to be one factor with a strong effect on English test scores. Extra classes are usually held after school, for one period (one hour) a day, five days a week during the semester. Six months of extra classes are expected to provide about 85 periods of teaching. The English teachers at the schools usually teach the extra class. Students pay a fee for the extra class, which goes to the

teacher. In most cases, all students, regardless of their grade level or their previous experience in attending extra classes, are taught together in the same class. Students in Grade 5 have studied English at school for more than 396 periods, as they have had three periods a week for 33 weeks a year over four years. It is difficult to believe that by adding 85 periods to the 396 periods, student scores could improve dramatically. The results suggest that the effect of extra classes is not only the result of additional teaching time, but also of the quality of the teaching. It is widely recognized that in the extra classes, many teachers use imported textbooks that are better-structured and more systematic in teaching sentence patterns/grammar.

**Table 4: Standardized Regression Coefficients from Multiple Regressions of Total English Test Scores**

	Total English test score
<i>Family background</i>	
Family economic indicator	.064*
Father's academic background	.014
Mother's academic background	-.026
Parental interest in study	.013
<i>Motivation for English study</i>	
English learning at home	.022
Attitude toward English learning	-.004
Enjoyment of English learning	.018
Importance of English learning	.041
<i>Educational inputs</i>	
Possession of English textbooks	-.020
Possession of an English-Lao dictionary	.059*
Possession of English books	.093***
TV or radio English teaching program	-.017
Extra English classes or private tuition	.220***
Average score in a class	.426***
$R^2$	.337

\* $p < .05$ , \*\* $p < .01$ , \*\*\* $p < .001$

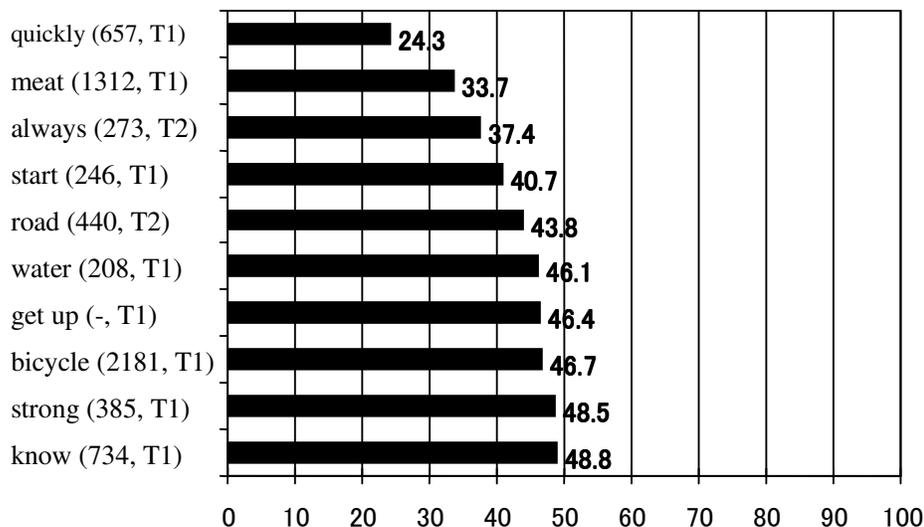
Possession of English textbooks also indicates a significant effect on English test scores. Of course, it is natural that reading English books improves English ability, especially reading ability. However, judging by the test scores of the survey, it seems that even a simple English book would be too difficult for these students. Therefore, more importantly, possession of English books might mean that families with English books have a more positive attitude toward English, and that some other family members are also familiar with English.

On one hand, some educational inputs are very likely to have an effect on English test

scores but, on the other hand, some others are not. It is important to note that problems exist in the educational inputs with little or no effect on English test scores; why aren't these inputs working effectively? Of all these factors, in this research, possession of textbooks gives a result particularly contrary to previous research, and is worthy of further study in a Lao context.

Many previous studies have revealed that possession of textbooks is an important input for improving student achievement in developing countries (e.g., Fuller & Clarke, 1994; Harber & Davies, 1997; Scheerens, 1999). However, the content of the English textbooks used in Laos seems too difficult for Grade 5 secondary school students. Figure 1 shows the results from the vocabulary section of the English test given to Grade 5 students in the study. The test results demonstrate that many students are not familiar with basic vocabulary items; for instance, 66.3 % of students did not know the meaning of “meat”; 53.3% did not know the meaning of “bicycle”; 51.5% of students did not know the meaning of “strong”.

A comparison of the test results and the content of the textbook reveal a mismatch between the students’ English level and the content of the textbook. For example, Picture 1 shows pages 2 and 3 of Textbook 3, used by Grade 5 students in Laos<sup>(6)</sup>. In spite of the fact that many Grade 5 students are not familiar with some basic vocabulary items, the readability of the textbook expects them to understand long sentences with more difficult vocabulary items. Furthermore, most of instruction and explanation are written in English<sup>(7)</sup>.



**Figure 1: Rate of Correct Answers in Basic Vocabulary\***

\*The number in parenthesis after each word is the order in which it appears on the General Service List of English Words (West, 1953). T1, T2 and T3 indicate the textbook “English

for Lao Secondary Schools 1, 2 or 3” (NRIES, 1996; NRIES, 1997a; NRIES, 1997b; NRIES, 1998a; NRIES, 1998b) in which the vocabulary word appears for the first time. The symbol “-” means that the vocabulary item is not on the General Service List.

Unit 1

**B How big is your family?**

**1 Pre-listening**  
Look up these words in the back of the book:  
*washing-up laundry*

**2 Listen and write**  
Kham's friend, Thongda, is asking Kham about his family.  
Listen to the dialogue. Write a list of all the family words you hear. For example:  
parents  
brothers

**3 Complete**  
Complete these sentences.  
1 My father's sister is my  *aunt*.  
2 My mother's brother is my ...  
3 My brother is my parents' ...  
4 My aunt's son is my ...  
5 My mother and father are my ...  
6 They have three sons and one ...  
7 My father's mother is my ...  
8 My brother's son is my ...  
9 My sister's daughter is my ...  
10 My mother's father is my ...

**4 Choose**  
Sometimes 's means *is*; sometimes it means *has*.  
Read the sentences and choose: *is* or *has*.  
1 Thong's quite fat. is  
2 Deng's got big ears. —  
3 My father's got short hair. —  
4 Chansy's quiet and very shy. —  
5 My mother's got two brothers. —  
6 Uncle La's only fifteen. —  
7 My father's gone to work. —  
8 My grandmother's not very old. —  
9 She's got problems with her eyes. —

**5 Complete**  
Complete the sentences, using *is, are, has* or *have*.  
1 There — a lot of people in Kham's family.  
2 Kham — got two brothers and two sisters.  
3 I don't — any sisters.  
4 There — nobody at home.  
5 My father — not in the office. He — gone home.  
6 My mother — very busy. She — a lot of work to do.  
7 My brother — handsome, so he — a lot of girlfriends.

**6 Write**  
Write two paragraphs about your family. In the first paragraph, give information about who is in your family, and give the names of the people. For example:  
*I've got two brothers, Bounthong and Soya...*  
In the second paragraph, describe people and say who looks like who. For example:  
*Bounthong is quite tall and slim. He looks like my mother...*



Family life

**C The traditional Lao family**

**1 Pre-reading**  
What do you know about traditional families in Laos? What kinds of work do different people do?  
Look at the words in the box on the right.

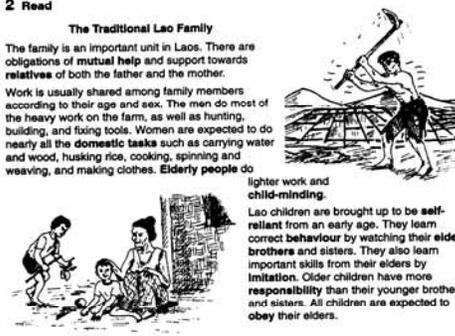
**2 Read**  
**The Traditional Lao Family**  
The family is an important unit in Laos. There are obligations of **mutual help** and support towards **relatives** of both the father and the mother. Work is usually shared among family members according to their age and sex. The men do most of the heavy work on the farm, as well as hunting, building, and fixing tools. Women are expected to do nearly all the **domestic tasks** such as carrying water and wood, husking rice, cooking, spinning and weaving, and making clothes. **Elderly people** do lighter work and **child-minding**. Lao children are brought up to be **self-reliant** from an early age. They learn correct **behaviour** by watching their **elder brothers** and sisters. They also learn important skills from their elders by **imitation**. Older children have more **responsibility** than their younger brothers and sisters. All children are expected to **obey** their elders.

**3 Match**  
Write these English words and phrases in your notebook. For each word, find a Lao word in the box with the same meaning. For example:  
1 mutual help – ຄວບຄູ່ກັນແລະກັນ

1 mutual help 7 obey  
2 relatives 8 behaviour  
3 domestic tasks 9 elder brothers  
4 elderly people 10 imitation  
5 child-minding 11 responsibility  
6 self-reliant

**Vocabulary –**  
obligation ຜົນຜົນ  
support ການຊ່ວຍເຫຼືອ  
hunting ການລ່າສັດ  
spinning ການຜັດໂກຍ  
bring up ຜິດຜອນລູກ

ມີສອງ ຄວາມໝາຍ ຄວາມໝາຍດຽວ  
ຜິດຜອນລູກ ຜິດຜອນ ຜິດຜອນ  
ສິ່ງສາມາດເຮັດໄດ້ ສິ່ງ ການສູນແສນ  
ຄວບຄູ່ກັນແລະກັນ ລູກເຮືອນ



Picture 1: English Textbook 3

### 4.3.1 Standardized regression coefficients from multiple regressions of English test scores of each section on variables

Next, English test scores were classified into the three sections - vocabulary scores, grammar scores and reading scores to compare the differences in the effect of the same variables on each section (Table 5). Table 5 is the result of a multiple linear regression equation (Enter method) in which vocabulary scores, grammar scores and reading scores became dependent variables. Comparing the results of vocabulary scores as a dependent variable with the results of grammar scores and reading scores as dependent variables, for vocabulary scores, the family economic indicator is not significant, but possession of an English-Lao dictionary is significant, while possession of English-Lao dictionary for grammar scores and for reading scores (Table 5: Step 1) are not significant. This result partly shows that, unlike the results for textbooks, possession of an English-Lao dictionary

contributes to mastering vocabulary. In general, textbooks need to target a specific level of learner, while many English dictionaries, including advanced dictionaries, can be used by different levels of learners<sup>(8)</sup>. For this reason, a Lao-English dictionary might work effectively.

Stanovich (1991) pointed out that poor reading comprehension could occur even with a large vocabulary in terms of size, but a poor vocabulary never leads to good reading comprehension. Reading scores (Table 5: Step 2) are the result of a multiple linear regression equation in which vocabulary scores and grammar scores are also added to the independent variables (Table 5: Step1). Three variables emerge at the significance level of 1%. Of the three, vocabulary scores show the highest standardized coefficients. The strongest effect on reading scores is shown by vocabulary scores as an independent variable. Considering the poor vocabulary of Lao students, the necessity of vocabulary learning should be emphasized for reading comprehension.

While possession of English books is significant for vocabulary scores and grammar scores, it is not significant for reading scores (both Steps 1 and 2). This result confirms the aforementioned inference that possession of English books does not lead to an improvement in students' reading ability. There may be another reason that possession of English books is significant only in vocabulary scores and grammar scores.

**Table 5: Standardized Regression Coefficients from Multiple Regressions of Vocabulary Scores, Grammar Scores and Reading Scores on Each Variable**

	Vocabulary scores	Grammar scores	Reading scores (Step 1)	Reading scores (Step 2)
<i>Family background</i>				
Family economic indicator	.026	.067*	.075*	.062
Father's academic background	.021	.001	.007	.002
Mother's academic background	-.002	-.046	-.022	-.017
Parental interest in study	.015	.027	-.018	-.025
<i>Motivation for English study</i>				
English learning at home	.004	.041	.013	.008
Attitude toward English learning	.003	-.027	.017	.020
Enjoyment of English learning	.018	.018	.003	-.003
Importance of English learning	.054	.016	.018	.004
<i>Educational inputs</i>				
Possession of English textbooks	-.003	-.051	.008	.014
Possession of an English-Lao dictionary	.068*	.019	.047	.029
Possession of English books	.096***	.099***	.008	-.025
TV or radio English teaching program	.002	-.005	-.054	-.054
Extra English classes or private tuition	.219***	.179***	.099**	.031
Average score in a class	.378***	.358***	.257***	.133***
<i>Test score</i>				
Vocabulary scores				.221***
Grammar scores				.115**
<i>R</i> <sup>2</sup>	.292	.232	.097	.156

\* $p < .05$ , \*\* $p < .01$ , \*\*\* $p < .001$

## **5. Summary**

This study has revealed several thought-provoking findings, with ramifications for English language education in Laos. First, Lao secondary students have a problem with basic vocabulary. Moreover, as vocabulary skills seem to have a significant effect on reading comprehension, basic vocabulary learning should be emphasized before reading comprehension is practiced. In order to do this, an English-Lao dictionary could be an important tool. In addition, the content of the textbook should be revised to focus more on basic learning, and it would be better to use Lao, rather than English, for instructions and explanations. As grammar scores are also low and are likely to affect reading comprehension, students should also focus more on basic grammar before tackling reading comprehension.

## **Endnotes:**

- (1) Because the percentages are rounded up to the second decimal place, total percentages do not equal 100%.
- (2) For example, regarding family factors, refer to Kellaghan (1994) or Buchmann & Hannum (2001). Regarding motivational factors, refer to Joseph et al. (1997) or Lens (1994). Regarding educational factors, refer to Fuller & Clarke (1994) or Scheerens (1999).
- (3) In the research, the variable “average score in a class” is considered to be an integrated indicator of the effects of a variety of school factors such as class facilities, class climate, leadership of the principal, teaching quality, and so on. These were not examined in the research. Further detailed investigation is needed in order to clarify which of the school factors contributed to the results.
- (4) Results of the t-test show that gender differences in English test scores were not observed at a significance level of 5%. The average scores of male and female students were 16.23 and 16.73 respectively.
- (5) This might not be surprising, since the average score in a class comprises an individual student's English test score. Nevertheless, as the relationship only accounts for one of approximately 46.8 students (the average number of students per class in the sampling schools), the effect of an individual student's English test score on the whole relationship is not large.
- (6) The total number of pages in the textbook is 162. Secondary students are expected to finish the entire textbook by the end of the first semester of Grade 5. On the whole,

with each increasing page number the content of the textbook becomes more difficult.

- (7) In fact, it is doubtful that the teachers themselves fully understand the textbooks. Twenty-four English teachers, who taught the students who participated in the study, were asked to take the same test as the students. They only answered 82% of the questions correctly. For the reading comprehension section, they only answered 64% of the questions correctly. It can be thus inferred that the lack of ability of teachers poses limitations on their ability to successfully teach students.
- (8) For instance, students learning English do not usually change English dictionaries every year, although they change English textbooks for higher level ones every year.

## **References**

- Appleby, R., Copley, K., Sithirajvongsa, S. & Pennycook, A. (2002). Language in development constrained: Three contexts. *TESOL Quarterly*, 36(3), 323-346.
- Buchmann, C. & Hannum, E. (2001). Education and stratification in developed and developing countries: A review of theory and research. *Annual Review of Sociology*, 27(1), 77-102.
- Fuller, B. & Clarke, P. (1994). Raising school effects while ignoring culture? Local conditions and the influence of classroom tools, rules, and pedagogy. *Review of Educational Research*, 64(1), 119-157.
- Goh, E. & Vonchith, T. (2003). English language teaching in Laos. In H. W. Kam & R. Y. L. Wong (Eds.), *English language teaching in East Asia today: Changing policies and practices* (pp. 257-268). Singapore: Times Academic Press.
- Harber, C. & Davies, L. (1997). *School management and effectiveness in developing countries*. London: Cassell.
- Joseph, O. J., Ruth, T. J. & Christianah, A. U. (1997). Effects of achievement motivation and study habits on Nigerian secondary school students' academic performance. *Journal of Psychology*, 131(5), 523-529.
- Kellaghan, T. (1994). Family and schooling. In *The international encyclopedia of education* (2nd Ed., Vols. 4, pp. 2250-2258). Oxford: Pergamon Press.
- Lens, W. (1994). Motivation and learning. In *The international encyclopedia of education* (2nd Ed., Vols. 7, pp. 3936-3945). Oxford: Pergamon Press.
- Ministry of Education. (2003). *Bothsalub Kanjathtang Patibath Phanphatthana Kanhsuksa 2002-2003 Lae Phanphatthana Kanhsuksa 2003-2004 [Summary of the*

*educational implementation plan for the academic year 2002-2003 and educational plan for the academic year 2003-2004*]. Vientiane.

National Research Institute for Education Sciences, Ministry of Education. (1994).

Laksuth Sanmatthayon Samansuka [*National curriculum for secondary school*]. Vientiane.

National Research Institute for Educational Sciences, Ministry of Education. (1996).

*English for Lao secondary schools 1*, Vientiane: Printing Press of the Ministry of Education.

National Research Institute for Educational Sciences, Ministry of Education. (1997a).

*Laksuth Phasa Angkith Sanmatthayomton [English curriculum for lower secondary schools]*, Vientiane: Ministry of Education.

National Research Institute for Educational Sciences, Ministry of Education. (1997b).

*English for Lao secondary schools 2*, Vientiane: Printing Press of the Ministry of Education.

National Research Institute for Educational Sciences, Ministry of Education. (1998a).

*Laksuth Phasa Angkith Sanmatthayon Tonpay [English curriculum for upper secondary schools]*. Vientiane: Printing Press of the Ministry of Education.

National Research Institute for Educational Sciences, Ministry of Education. (1998b).

*English for Lao Secondary Schools*, Vientiane: Printing Press of the Ministry of Education.

Reena, S. & Peter, B. (2001). Reading in a second language: Factors associated with progress in young children, *Educational Psychology*, 21(2), 189-202.

Scheerens, J. (1999). School effectiveness in developed and developing countries; A review of the research evidence. *World Bank Paper*, Retrieved on February 4, 2006 from <http://www1.worldbank.org/education/est/resources/topic%20papers/jaap699.doc>

Schmitt, N. (2000). *Vocabulary in language teaching*. Cambridge: Cambridge University Press.

Sithirajvongsa, S. & Goh, E. (2004). Lao PDR. In H. W. Kam & R. Y. L. Wong (Eds.), *Language policies and language education: The impact in East Asian countries in the next decade* (2nd Ed., pp.196-206). Singapore: Eastern Universities Press.

Stanovich, K. (1991). Word recognition: Changing perspectives. In R. Barr, M. L. Kamil, P. Mosenthal & P. D. Pearson (Eds.), *Handbook of reading research* (pp. 418-452). New York: Longman.

Tanaka, M. & Lauer, J. (2003a). *English qualification test for Japanese university students*.

Hiroshima: Hiroshima University.

Tanaka, M. & Lauer, J. (2003b). *Vocabulary levels test for Japanese learners of English.*

Hiroshima: Hiroshima University.

Vientiane Municipality Education Office. (2003). *Bothsalub Kanjathtang Patibath Phanphatthana Kanhsuksa 2002-2003 Lae Phanphatthana Kanhsuksa 2003-2004 [Summary of Educational Implementation Plan for the Academic Year 2002-2003 and Educational Plan for the Academic Year 2003-2004]*. Vientiane.

Vientiane Province Education Office. (2003). *Bothsalub Kanjathtang Patibath Phanphatthana Kanhsuksa 2002-2003 Lae Phanphatthana Kanhsuksa 2003-2004 [Summary of Educational Implementation Plan for the Academic Year 2002-2003 and Educational Plan for the Academic Year 2003-2004]*. Vientiane Province.

West, M. (1953). A general service list of English words. *London: Longman*, Retrieved on February 4, 2006 from <http://jbauman.com/gsl.html>



## **Factors Predicting Effective Oral Presentations in EFL Classrooms**

**Junko Otoshi**

*Gunma Prefectural Women's University, Japan*

**Neil Heffernan**

*Ehime University, Japan*

Junko Otoshi currently teaches at Gunma Prefectural Women's University in the Department of International Communication. Her research interests include second language writing studies and testing/evaluation.

Neil Heffernan is an Associate Professor at Ehime University. His research interests include testing/evaluation and CALL.

### **Abstract**

This paper outlines a study that explores what factors EFL learners consider to be important when making presentations. 304 undergraduate students were surveyed at a large private university in Japan. A 30-item questionnaire was used to discover what components are considered to be important to learners in doing effective English presentations. Descriptive statistics and principal component analysis were used to analyze the results of the data. The results indicate that the participants consider the following three factors as the major criteria for effective English oral presentations: clarity of speech and voice quality; correctness of language; and interaction with the audience. Based on the outcome of the study, this paper further delineates the pedagogical implications for oral presentations in EFL.

**Keywords:** Oral presentations; peer evaluation; peer-assessment; feedback; autonomy.

### **Introduction**

As an alternative form of assessment, peer assessment has been utilized in oral presentation activities in various educational contexts (Boud, Cohen & Sampson, 1999; Patri, 2002). Echoing with the concept of a student-centered approach to instruction, students can take an active role in their own language learning through the use of peer assessment activities. Opposed to teacher-only assessment, oral performances can also be evaluated from the views of learners' peers. Obtaining feedback from peers is vital to communicative language learning situations like oral presentations because of the notion of interaction between learners (Earl, 1986; Rust, Price & O'Donovan, 2003). Furthermore, students can achieve a higher level of learning through interaction with their peers and instructors (Earl, 1986; Rust et al., 2003).

Thus, peer assessment is considered an important activity to develop students' learning

and to facilitate autonomy among learners. In most classrooms, however, the assessment criteria are already established by the classroom teacher, with students' ideas not incorporated into these pre-existing rubrics. This lack of student input might lead to a low reliability of peer assessment. Because students are not well aware of the description of each evaluation criterion, some students might not be able to assess their peers' performances properly. As a result, their views might be different from those of their instructor. Additionally, pre-existing criteria might deprive students of the aforementioned benefit of increased autonomy from peer assessment activities because they are not involved with establishing the criteria. Therefore, it is useful to explore students' views of which aspects of oral presentations are most effective which, as a result, will become the criteria of peer assessment activities. In doing so, instructors will find out what aspects of oral presentation students consider important. Defining and creating the evaluation rubric together with their instructor, learners will gain more responsibility for their learning as well as to improve the reliability of the peer assessment activities themselves.

The present study explores exactly which aspects of oral presentations Japanese university students view to be most effective. Incorporating students' ideas while establishing the criteria for presentations is an important factor when considering a learner-centered approach in EFL classes. Therefore, this paper will suggest specific ideas as to how to carry out oral presentation activities using a student-established evaluation criteria. A description of Japanese university students' views on effective presentations will yield valuable data both for oral presentation activities and their assessment. A secondary aim of the research is to suggest a set of specific recommendations to improve peer assessment activities for oral presentations in ES/FL classrooms.

To summarize the above, the primary focus of this research is as follows: to make learners more cognizant of the importance of presentation skills in English; to get learners more involved in the evaluation process; to urge learners to think about the criteria that form an effective presentation; to have learners involved in the formulation of the evaluation criteria; and have learners receive evaluation from their peers as well as reflect critically on this method of measurement and how it effects their own oral presentations. It should be noted that this paper should not be considered exclusive to the Japanese context. In fact, the fundamental concepts and results of this paper can be applied to most Asian contexts and beyond.

## **Review of the Literature**

There has been a marked increase in the popularity of peer-evaluation activities in EFL classes (Cheng & Warren, 2005; Finch, 2003; Hadley, 2001; Johnson & Johnson, 1999; King, 2002) due to the increasing number of teachers emphasizing the importance of communication skills in the classroom (Hadley, 2001; King, 2002; Nunan, 2002; Richards & Rodgers, 2001). Thus, teachers need to become more cognizant of the principles and methods of peer evaluation. EFL students, especially in business fields, are well aware of the impact of competent presentation skills in English. In fact, students who are preparing for the working world tend to take presentation classes as a part of their regular degree programs (Nakamura, 2002). However, to date, most of the work in peer evaluation has been done in ESL and EFL writing classes (see Duke & Sanchez, 1994; Hedgcock & Leftkowitz, 1992; Pica, 1984; Lalande, 1982; Zamel, 1985), so a more thorough look at peer assessment for EFL oral presentations is warranted. Advocates of peer assessment highlight learner autonomy as one of the benefits of conducting this type of evaluation (Cheng & Warren, 2005; Johnson & Johnson, 1999), as each student takes responsibility for his or her own active learning (Cheng & Warren, 2005; Patri, 2002). Thus, learners can become more cognizant of exactly what they are learning through peer assessment activities. In addition, there is a general consensus that involving students in their own correction is helpful in developing skills useful in academic life (Cheng & Warren, 2005) and enhancing interpersonal relationships among learners (Earl, 1986). A combination of teacher-, peer-, and self-evaluation seems to yield the most successful results (Bachman & Palmer, 1989; Hadley, 2001; Hedgcock & Lefkowitz, 1982; Lalande, 1982). However, in spite of the usefulness and reliability of peer evaluation (Freeman, 1995; Hughes & Large, 1993; Patri, 2002), it remains an area of EFL that has not been well researched, as seen in the dearth of theoretical and empirical studies published in this emerging field of study (Patri, 2002).

Zamel (1985) maintains that teachers' feedback on ESL compositions is at times inconsistent. Cohen (1985) suggests that teachers' comments on student work are confusing and vague, while Truscott (1996) came up with similar results stating that teachers should abandon correcting student errors altogether. Hedgcock and Leftkowitz (1982) state that more attention is paid to content and organization in peer-evaluated work; two very important elements in the evaluation of oral work. Indeed, it seems teachers tend to focus on form and accuracy when marking, so getting learners to conduct peer-evaluation has the potential of circumventing this problem. The problems of

traditional feedback methods found by these researchers clearly have implications for the classroom, and points to the need for a more useful evaluation method. Peer evaluation offers a remedy for these problems.

Although teachers will always feel the need to correct their learners' work, a balance needs to be found between teacher-assessed work and work that is peer-evaluated. In a traditional teacher-centered classroom, students tend to adopt a passive attitude toward the assessment of their oral performances because they are being rated by a single person: the teacher. Further, the evaluation criteria are usually set by the teacher, so learners have no input into the grades they receive. One method around this is to have learners involved in the creation of the assessment rubrics to be used in class (Patri, 2002).

This evidence points to the fact that it is the type of feedback given to learners that is of ultimate importance. Peer-evaluation gives the benefit of learners learning from their peers, while being actively involved in their classmates' work (Birdsong & Sharplin, 1986; Cheng & Warren, 2005). Indeed, peer-evaluation tends to have a double-pronged effect: it helps the evaluator notice errors, and also assists the presenter in learning where he or she can improve.

Another related factor is the possibility that students not trained in the art of peer evaluation may demonstrate problems that have been outlined in the literature. For example, previous studies have indicated certain biases – such as friendship bias – may exist when learners evaluate their peers (Dochy, Segers & Sluijsmans, 1999; Falchikov, 1995; Morahan-Martin, 1996). If learners are explicitly taught the form and practice of peer evaluations, this problem can be remedied quite easily (Finch, 2003; Jacobs, 1987; Newkirk, 1984). Therefore, giving our learners the correct tools to undertake peer-evaluation activities has the potential to greatly enhance the reliability and usefulness of such activities.

Oral presentation assessment criteria consist of multi-faceted factors: language, content, delivery, and visual aids. It is incumbent among teachers to clearly delineate these factors to their learners before undertaking any peer assessment activities in class. In a classroom setting, the criteria for assessment are usually established by the teacher, who also has the responsibility of judging students' presentation performances. However, this single rater practice raises reliability issues (Bachman & Palmer, 1989). As such, there is a good possibility that an instructor's assessment is not objective or consistent enough to meet the needs of his or her learners. The consensus among researchers is that if learners are explicitly aware of the criteria used to assess them they will be better able to apply the

method of evaluation to their own learning styles, and thereby to their own benefit as learners (Rust et al., 2003).

Further, students are able to comprehend and critically assess marking criteria that they have a hand in making, and have more difficulty comprehending marking criteria that are imposed on them by teachers (Orsmond, Merry & Reiling, 2000). The development of the criteria is just as important as the assessment itself, so great care must be taken by learners when considering just how to mark their peers (Erwin & Knight, 1995). However, the criteria must be presented in an operational manner that is understandable to students; a process the students themselves can have a hand in creating (Dochy et al., 1999).

Lastly, it has been found that learners demonstrate extremely positive attitudes towards peer-evaluation activities (Birdsong & Sharplin, 1986; Cheng & Warren, 2005; Duke & Sanchez, 1994), and that such activities can foster intrinsic motivation and confidence in the language learner (Finch, 2003). Accordingly, learners actually enjoy evaluating each other's work, and have much to gain from performing this activity. Indeed, Topping (1998) maintains that learners evaluate their peers in a comparable manner to their teachers, which suggests it is worthwhile for teachers to seriously consider using peer-evaluation activities in class. However, learners must be given the correct tools to do so, and need to be trained accordingly before undertaking any serious peer-evaluation activities (Cheng & Warren, 2005).

### **Participants**

This study was conducted in the Autumn of 2005 at a large private university in Japan. Some 304 freshman undergraduate students voluntarily participated in the study. The students were studying in the Economics, Business Administration, and Science and Engineering faculties. All students were taking an oral presentation class, which was a required subject in their English course as indicated by the questionnaire.

### **Research Methods**

The aim of the study is to explore what kinds of oral presentations Japanese university students consider to be effective. In order to answer the research question, an attitude survey was conducted. Descriptive statistics and principal components analysis using SPSS13.0 were employed to interpret the results.

The questionnaire was developed based on the literature review and informal conversations with university students and EFL specialists who have taught oral

presentations in both EFL and ESL contexts. The questionnaire consisted of thirty items across the following evaluation criteria of oral presentations: eye-contact; voice; English; originality of content; clarity; Powerpoint; body language; and time management (see the evaluation rubrics in Appendix 1). The criteria and the description of the rubrics were modified for this study by identifying common rubrics from related studies (e.g., Cheng & Warren, 2005; King, 2002; Nakamura, 2002; Patri, 2002).

In order to spread the responses from the participants out (Likert, 1932), item statements were categorized into positively and negatively phrased statements. Each item was measured using a 5-point Likert-type scale (for positive statements, from 1= strongly disagree to 5= strongly agree; and for negative statements the coding was reversed). Chronbach's alpha was checked to establish the reliability of the instrument developed for gathering the data for the study. This figure was determined to be .843, and thus reliable.

### **Results of the Descriptive Analyses**

Table 1 presents the descriptive statistics of 30 items that dealt with university students' attitudes towards effective oral presentations in English. Item 23 (A presentation should be given in a clear voice) demonstrates the highest score and the participating students most agree with this statement. On the other hand, Item 28 (I don't mind grammatical mistakes in a presentation as long as the message is clearly delivered to the audience) has the lowest score among the thirty statements. This item was phrased negatively so the scale was reversed to properly measure students' attitudes towards effective oral presentations in English. It was noted that the participant students tended to agree with Item 23, demonstrating that they accepted grammatical errors by their peers.

**Table 1: Results of the Descriptive Statistics (N= 304: Max.=5 Min.= 1)**

Item Statement	Mean	SD
1. Presentation topics should be interesting to the audience.	4.18	.85
2. Smiling is <u>not good</u> while giving a presentation	4.07	.92
3. A presentation should be given in an organized way.	4.45	.75
4. Glancing at a transcript is <u>not good</u> while giving a presentation.	3.31	1.05
5. Using signal words such as "First" and "Second" are important when giving a presentation.	3.62	.85
6. Using PowerPoint is <u>not necessary</u> when giving a presentation.	4.08	.92
7. Speakers should avoid using difficult terms when giving a presentation.	3.90	.98
8. Speakers should just speak about whatever they want even if the audience does not understand it.	4.41	.83
9. Good presentations include detailed examples and reasons.	4.28	.71
10. I <u>don't mind</u> if I find grammatical errors in a PowerPoint presentation.	3.55	1.09
11. Speakers should pay attention to the audience's response while they	3.79	.88

12. A PowerPoint presentation <u>does not have to</u> include statistical data when speakers mention numerical information.	3.97	.88
13. Speakers should argue their own ideas or possible solutions in their talk.	3.91	.87
14. A good PowerPoint presentation includes pictures and photographs.	4.05	.83
15. Speakers <u>don't have to</u> speak fluent English.	2.70	1.00
16. Oral presentations should be given in informal language (as opposed to a formal, written style of language).	3.05	.97
17. Speakers should stick to the objectives of the presentation without confusing the audience.	4.25	.67
18. Speakers <u>don't have to</u> finish the presentation within an allotted time.	3.75	1.06
19. Speakers <u>don't have to</u> act cheerfully when speaking.	3.91	1.20
20. Speakers should make eye-contact with the audience.	3.81	1.00
21. Speakers <u>don't have to</u> outline the presentation objectives to the audience.	4.26	.86
22. The size of the letters in a PowerPoint presentation should be easy to read.	4.29	.77
<b>23.</b> A presentation should be given in a clear voice.	<b>4.47</b>	.71
24. The speaker should use some body language while speaking.	3.71	.88
25. A presentation should be given in a clear voice.	4.03	.88
26. Speakers should pay attention to the speed of the speech.	4.15	.82
27. Speakers should speak with confidence.	4.28	.80
<b>28.</b> I <u>don't mind</u> grammatical mistakes in a presentation as long as the message is clearly delivered to the audience.	<b>2.63</b>	1.11
29. A presentation should be delivered with correct pronunciation.	3.35	.93
30. Speakers <u>don't have to</u> speak loudly.	4.10	.85

### Results of the Principal Components Analysis

In order to ascertain what kinds of oral presentations the participants consider to be effective, a principal component analysis with a varimax rotation was carried out using SPSS13.0. Regarding the salient value for a component loading, we examined Tabachnick and Fidell (1996) who proposed a rule of thumb that loadings of .30 (absolute) or above be used to specify variables that load on each component, and Stevens's (1996) recommendation that using loadings which are approximately .40 (absolute) or higher for identifying variables that load on a particular principal component. Also, in a sample of 300 (Stevens, 1996), factor loading of .298 (.149 X 2) at alpha =.01 and above are required for practical significance. In a practical sense, selecting factor loadings of .40 meets the criteria. We decided to choose the value of .40 in order to select the components. As a result, three components were retained and interpreted as follows:

1. Factor 1 has the largest number of loadings from 9 variables (items 3, 8, 9, 17, 21, 23, 26, 27, and 30), accounting for 22.8% of the total variance. Items 3, 8, 9, 17, 21, and 22 dealt with comprehension of speech. Items 23, 26, 27, and 30 are all related to the speaker's voice quality. Therefore, Factor 1 was labeled as *Clarity of Speech and Voice*

Quality.

2. Next, the variance of Factor 2 drops sharply on the scree plot, and accounts for 6.84 % of the total. It received loadings from 4 variables (items 10, 15, 28, and 29). All of these items are concerned with the English language itself, so this group was labeled

*Correctness of Language.*

3. Factor 3 received loadings from three variables (items 11, 20, and 24). These items dealt with the speaker's attention to the audience. Therefore, it was labeled *Interaction with the Audience.*

**Table 2: Results of the Principal Component Analysis**

Factor	Loading	Items
Factor 1 (Clarity of speech and voice quality) $\alpha = .855$ contribution ratio: 22.8%	.709 .451 .468 .572 .675 .799 .527 .685 .540	3. A presentation should be given in an organized way. 8. Speakers should just speak about whatever they want even if the audience does not understand it. 9. Good presentations include detailed examples and reasons. 17. Speakers should stick to the objectives of the speech without confusing the audience. 21. Speakers don't have to outline the presentation objectives to the audience. 23. A presentation should be given in a clear voice. 26. Speakers should pay attention to the speed of the presentation. 27. Speakers should speak with confidence 30. Speakers don't have to speak loudly.
Factor 2 (Correctness of language ) $\alpha = .605$ contribution ratio: 6.84%	.660 .602 .700 .617	10. I don't care if I find grammatical errors in a PowerPoint presentation. 15. Speakers don't have to speak fluent English. 28. I <u>don't mind</u> grammatical mistakes in a presentation as long as the message is clearly delivered to the audience. 29. A presentation should be delivered with correct pronunciation.
Factor 3 (Interaction with the audience) $\alpha = .668$ contribution ratio: 5.22%	.763 .608 .451	11. Speakers should pay attention to the audience's response while they speak 20. Speakers should make eye-contact with the audience. 24. The speaker should use some body language while speaking.

### Discussion

From the results of the present study, three components were identified to influence Japanese university students' views on effective oral presentations: Clarity of Speech and Voice Quality; Correctness of Language; and Interaction with the Audience.

Regarding Factor 1, Clarity of Speech and Voice Quality, it was noted that participants judge an effective presentation on the combination of clarity of the presenter's presentation and his or her voice. Thus, in the process of making evaluation rubrics, teachers should pay attention to this factor by incorporating examples of good presentations and by demonstrating a clear and understandable voice. It is suspected that some students, especially those in the lower levels of proficiency, may confuse the clarity of speech component with that of a person who simply just speaks with a loud voice. Therefore, teachers must explain the component items of Factor 1, and how clarity of speech and voice quality can positively impact an oral presentation. This will serve to act as guidance for learners striving to become more effective presenters.

With respect to Factor 2, Correctness of Language, from the results of the descriptive statistics, participants in this study tended to be tolerant of grammatical errors during oral presentations. Although the students might not put serious emphasis on the correctness of language, the fact that it is included as a component here carries some weight. Component items of this factor involve the accuracy of speech and pronunciation. This factor is considered important for the preparation stages of oral presentations, and can be dealt with by teachers in class by clearly explaining the implications of these criteria to their learners, so they can factor this into their evaluations. Further, when preparing to assess their peers, learners should be made aware of the significance of the correctness of language criteria. Thus, learners can prepare transcripts of their presentations, which can then be checked by their peers. The presenter's peers can then focus on grammatical accuracy and vocabulary use. This will ensure that a complete understanding is obtained by both presenters and their peers.

Factor 3, or Interaction with the Audience, refers to the physical attention the presenter gives to the audience. As outlined in Table 2, this physical attention involves eye-contact, body language and paying attention to the audience's response to the presentation. It can be a difficult task to judge exactly how much eye-contact and body language is sufficient in oral presentations in the Japanese EFL context. This is partly due to these factors not being strongly rooted in Japanese culture (Lusting & Koester, 2003). However, with practice, teachers can assist learners in overcoming potential problems in these areas by reiterating that in real world situations presenters are expected to pay attention to the audience because the presentation should be aimed at the target audience (Lucas, 1997). Teachers can demonstrate and stress the importance of presenter-audience interaction, and learners can prepare for evaluating their peers' interaction with the audience with adequate

training. This can be done by teachers and students in class: by giving both good and bad examples of presenter-audience interaction.

### **Conclusion and pedagogical implications**

Based on the results of this study, we can draw some conclusions about recommendations on how to best prepare our learners for oral presentations. As teachers, we can prepare our learners to be aware of the above criteria (Table 1 and Table 2) and the effect they have on the efficacy of oral presentations in EFL classrooms.

In light of the work done on peer assessment in writing classes and elsewhere, a more open approach to our learners' evaluation is needed (Rust et al., 2003). One suggestion is to use an integrated method of feedback: one in which learners write an outline for their presentations first, evaluate each others work, and then make their presentations based on the corrections given by their peers. Learners then do their presentations in-class while the teacher and the other students evaluate that presentation using rubrics they have had a hand in creating (Cheng & Warren 2005; Orsmond, Merry & Reiling, 2000). Consequently, learners from the present study stated that these methods were exceptionally useful in assisting them in becoming more proficient at making oral presentations. Specifically, comments from learners suggested that if teachers take time to demonstrate how and when to use eye contact, how to organize a presentation, connect with an audience, use body language and manage time, and how to construct an effective PowerPoint presentation, the learners can build upon their existing knowledge to use in future presentations. Methods that seem commonplace to most teachers are not so to learners in our classes, thus we must ensure we expend the correct amount of time and energy into giving our learners the tools they need to become effective presenters in the target language. Students wanting to improve their competence in this area of language study will surely benefit from such training.

Along these lines, the first step is for teachers and learners to establish a marking criteria. Involving learners in this process will be of great benefit to them in preparing for their own presentations (Orsmond, Merry & Reiling, 2000). Secondly, learners must clearly understand the criteria set out for evaluating presentations. Teachers can carry out practice sessions with their learners by having them watch and evaluate past presentations on video. Teachers need to monitor that learners are using the criteria correctly and applying the tenets of the rubric in an accurate manner. Third, after the appropriate training has been completed (which may take two to three class periods), students prepare

their presentations (individually or in groups) with the criteria they helped create firmly in mind. Next, learners evaluate their peer's presentations with the rubric they helped create. Finally, after the presentations are completed, learners receive their evaluation from their peers and reflect on their performance and how to improve on it. All of this can serve to be extremely useful not only to the learners in our classes, but also to teachers using these techniques in EFL presentation classes.

The question remains as to exactly what constitutes good spoken performance in the EFL classroom? For the purposes of EFL classes and in general, the authors deem it vital for learners to be able to communicate their message to the audience in a convincing manner. Thus, grammatical correctness is not of the utmost importance (as noticed by its absence in the rubric in Appendix 1). What are most important are the objectives, organization and presentation of the speech, as well as the communication of both the message and of the presenter with his or her audience. These are factors which can literally make or break an oral presentation in the EFL classroom. As teachers, we must inform our learners of their importance, practice them repeatedly, and then reiterate exactly how they can affect the overall message in a presentation.

Peer evaluation can be a valuable method in assisting EFL learners in how to properly structure English oral presentations. By taking part in peer evaluation activities, learners gain a firm knowledge of the form and process of what makes an effective oral presentation. By involving them in the process of actually creating the rubrics to be used in evaluating their peers, we are giving our learners an opportunity gain independence while learning more about exactly what makes a successful presentation.

Future research should focus on surveys that target instructors' ideas of what constitutes a good oral presentation, since there is bound to be some differences in attitude between students and instructors over what constitutes a good oral presentation. For instance, native instructors of English might have different criteria from those of their students due to cultural differences. Discovering the differences between the two groups will become a stepping stone that will lead to more effective instruction in EFL oral presentation classes. In accordance with the notion of a student-centered approach, both teachers and students should negotiate with each other over their views on oral presentations.

As argued earlier in this paper, performing an oral presentation is a complicated task that involves multi-faced language characteristics. To enhance the reliability of oral presentation activities, they should be evaluated by both teachers and learners. They are

useful to both groups interested in utilizing classroom peer evaluation activities as a method of improving learners' knowledge of exactly what makes an effective oral presentation in English. Therefore, interaction with both sides is necessary so that peer evaluation can be considered a useful and effective communicative activity in EFL language learning contexts.

### **References**

- Bachman, L.F., & Palmer, A.S. (1989). The construct validation of self ratings of communicative language ability. *Language Testing*, 6, 14-29.
- Birdsong, T., & Sharplin, W. (1986). Peer evaluation enhances students' critical judgment. *Highway One*, 9, 23-28.
- Boud, D., Cohen, R., & Sampson, J. (1999). Peer learning and assessment. *Assessment and Evaluation in Higher Education*, 24(4), 413-26.
- Cheng, W. & Warren. (2005). Peer assessment of language proficiency. *Language Testing*, 22(1), 93-119
- Cohen, A. (1987). Student processing of feedback of their compositions. In A. Wenden and J. Rubin (Eds.), *Learner strategies in language learning*. Englewood Cliffs, N.J.: Prentice Hall.
- Dochy, F., Segers, M., & Sluijsmans, D. (1999). The use of self-, peer-, and co-assessment in higher education: A review. *Studies in Higher Education*, 24(3), 331-35.
- Duke, C., & Sanchez, R. (1994). Giving students control over writing assessment. *English Journal*, 83, 47-53.
- Earl, S.E. (1986). Staff and peer assessment: measuring an individual's contribution group performance. *Assessment and Evaluation in Higher Education*, 11, 60-69.
- Erwin, T.D. & Knight, P. (1995). A transatlantic view of assessment and quality in higher education. *Quality in Higher Education*, 1(2), 179-188.
- Falchikov, N. (1995). Peer feedback marking: developing peer assessment. *Innovations in Higher Education*, 11(2), 146-65.
- Finch, A. (2003). Reflective instruments for self-assessment in Korean EFL classrooms. *The Korea TESOL Journal*, 6(1), 63-86.
- Freeman, M. (1995). Peer assessment by groups of group work. *Assessment and Evaluation in Higher Education*, 20, 289-300.
- Hadley, A. O. (2001). *Teaching language in context* (3rd ed.). Boston, MA: Heinle & Heinle.

- Hedgcock, J., & Leftkowitz, N. (1992). Collaborative oral/oral revision in FL writing instruction. *Journal of second language writing*, 1(3), 255-76.
- Hughes, I.E., & Large, B.J. (1993). Staff and peer-group assessment of oral communication skills. *Studies in Higher Education*, 18(3), 379-85.
- Jacobs, G. (1987). First experiences with peer feedback on compositions: student and teacher reactions. *System*, 15, 325-33.
- Johnson, D., & Johnson, R. (1999). Cooperative learning and assessment. *JALT Applied Materials Cooperative Learning*, 164-178
- King, J. (2002) Preparing EFL learners for oral presentations. *The Internet TESOL Journal*, 8(3). Retrieved November 1, 2007 from <http://iteslj.org/Lessons/King-PublicSpeaking.html>
- Lalande, J. (1982). Reducing composition errors: An experiment. *The Modern Language Journal*, 66, 140-49.
- Likert, R.A. (1932). A technique for the measurement of attitudes. *Archives of Psychology*, 140, 5-53.
- Lucas, S.E. (1997). *The art of public speaking* (6<sup>th</sup> ed.). New York: McGraw Hill.
- Lusting, M.W. & Koester, J. (2003). *Intercultural competence interpersonal communication across cultures* (4<sup>th</sup> ed.). Boston, MA: Allyn&Bacon
- Morahan-Martin, J. (1996). Should peers' evaluations be used in class project?: Questions regarding reliability, leniency, and acceptance. *Psychological Reports*, 78, 1243-50.
- Nakamura, Y. (2002). Teacher assessment and peer assessment in practice. *Education Studies*, 44, 203-215
- Newkirk, T. (1984). Direction and misdirection in peer response. *College Composition and Communication*, 35, 301-11.
- Nunan, D. (2002). Listening in language learning. In J.C. Richards & W.A. Renandya (Eds.), *Methodology in language teaching* (pp. 238-241). Cambridge: Cambridge University Press.
- Orsmond, P., Merry, S., & Reiling, K. (2000). The use of student derived marking criteria in peer and self-assessment. *Assessment and Evaluation in Higher Education*, 25(1), 23-38.
- Patri, M. (2002). The influence of peer feedback on self-and peer assessment of oral skills. *Language Testing*, 19(2), 109-31
- Pica, T. (1984). Second language acquisition theory in the teaching of writing.

TESOL Newsletter, 18(2), 5-6.

Richards, J. C., & Rodgers, T. (2001). *Approaches and methods in language teaching*.

Cambridge: Cambridge University Press.

Rust, C., Price, M., & O'Donovan, B. (2003). Improving students' learning by developing their understanding of assessment criteria and processes. *Assessment and Evaluation in Higher Education*, 28(2), 147-164.

Stevens, J. (1996). *Applied multivariate statistics for the social sciences* (3<sup>rd</sup> ed.). Mahwah, NJ: Lawrence Erlbaum Associates.

Tabachnick, B.G. & Fidell, L.S. (1996). *Using multivariate statistics* (3<sup>rd</sup> ed.). New York: Harper Collins College Publishers.

Topping, K. (1998). Peer assessment between students in colleges and universities. *Review of Educational Research*, 68, 249-76.

Truscott, J. (1996). Review article: The case against grammar correction in L2 writing classes. *Language Learning*, 46(2), 327-69.

Zamel, V. (1985). Responding to student writing. *TESOL Quarterly*, 19, 79-101.

#### Appendix 1: The rubric for Judging Speech Quality

Eye-contact	rapport with, and sense of the audience
Voice	show confidence; appropriateness of pace; volume; and articulation
English	clear pronunciation ; appropriate use of vocabulary good structures/register; conciseness; and clarity of expressions
Originality of Content	good choice of topic; clear objective/ purpose statement; and unique ideas about the chosen topic
Clarity	good connection of ideas; appropriate use of signal words; well structured; clear conclusions
PowerPoint	attractive; appropriate use of paragraphing, headings, numbering, spacing, and illustrations.
Body Language	appropriate gestures to keep audience's attention
Time Management	Management of the allocated time of oral presentation



## **The Discourse of an English Teacher in a Cyber Writing Course: Roles and Autonomy**

**Chi Yen Chiu**

*National Formosa (Huwei) University*

### **Bio Data:**

Chi-Yen Chiu is an assistant professor in the department of Applied Foreign Languages at the National Formosa (Huwei) University in Taiwan. He gained his Doctorate of Applied Linguistics from the Pennsylvania State University. He currently teaches English pronunciation, Phonetics, TESL, and writing courses. His research interests include uses of technology, practices of feedback, and teacher roles for learner autonomy in teaching English writing.

### **Abstract**

The case study reported in this paper investigates the relationships of teacher roles and learner autonomy in the cyber teaching of two EFL adult writers. The discourse of the teacher and learners was constructed in the written e-mail text because e-mail was their only means of communication. A total of 362 e-mails were generated in the twenty-month period of the cyber writing course. The teacher as investigator used NVivo 1.1-3, a qualitative data analysis software to conduct a content analysis that identified his own discourse in terms of *teaching* and *counseling* roles in a sample of 90 teacher e-mails, spread equally among the beginning, middle and end phases of the instructional period. Linked to the content analysis, a follow-up discourse analysis further examined the ensuing learner-teacher interactions to see how the two learners reacted to the teaching and counseling roles of the teacher. The results suggested that teaching roles did not provide opportunities for promoting learner autonomy whereas counseling roles created a supportive learning environment for the development of autonomy in language learning. The results of the data analysis provided additional evidence in support of the connection between computer-mediated communication and autonomy in language learning. Implications were drawn to call into question the universality of established categories of teacher roles in autonomous language learning, suggesting that cultural context and experience need to be taken into consideration.

**Keywords:** teacher roles, learner autonomy, computer-mediated communication, e-mail pal exchanges, EFL adult writers

### **1. Introduction**

For anyone who has access, the computer is now a key component of English language learning. As Graddol (1997) explains, the majority of Internet websites are based in English, and people who normally speak in other languages oftentimes find themselves communicating with each other in English. Not surprisingly, more and more websites are created to provide resources and materials for English teaching and learning. Some of

them even offer English lessons and courses for free. *English for Internet* (EFI), the pedagogical context of this study, is one such example. Described as *Your Free English School on the Net*, EFI offers a variety of classes to language learners from all over the world.

Advances in computer technologies such as the Internet and E-mail are often associated with developing autonomy in language learning. The Internet has become an important medium for the promotion of learner autonomy. On the one hand, the use of computer learning networks brings more learner-initiated interactions, more learner-centered discussion and a shift in authority from the teacher to the learner in language learning (Balester, Halasek & Peterson, 1992; Barker & Kemp, 1990; Faigley, 1990; Warschauer, Turbee & Roberts, 1996). On the other hand, technologies of the Internet provide opportunities for interactions among language learners, between L2 learners and L1 users, and between language learners and teachers that would be otherwise difficult to achieve in a L2 classroom or other language learning contexts (Benson 2001).

E-mail tandem language learning is a particularly useful tool of computer-mediated communication to facilitate autonomous language learners. The term *tandem* refers to organized language exchanges between two L2 learners who both wish to improve their proficiency in the other's L1 (Appel & Mullen, 2000). In e-mail tandem learning, two language learners are paired up to correspond with each other via e-mail. Both of them should be responsible for their learning process in which they will determine their learning goals and methods (Schwienhorst, 1997). Little and Brammerts (1996) point out that a fundamental principle of e-mail tandem language learning is autonomy. Learners in e-mail tandem must take control of their learning and have a mutual responsibility to make their partnership as beneficial to each other as possible.

Over the last twenty years, increasing attention has been drawn to the concept of autonomy in language learning. Autonomy has in fact become a buzzword (Little, 1991) and a central theme in language learning and teaching (Camilleri, 1999). The decision to promote autonomy usually comes from the teacher (Hill, 1994). Teachers of autonomous language learners are portrayed as helper, facilitator, resource, consultant, counselor, coordinator, and adviser (Voller, 1997). Nonetheless, there is a dearth of research to investigate the reactions of language learners in response to teacher roles said to promote autonomy.

The research reported in this paper has for its focus the investigation of teacher roles and learner autonomy in a cyber pedagogical context, a context wherein the teacher as

well as the learners are L2 users of English with diverse linguistic and cultural backgrounds and experiences. E-mail was the only means of communication, the medium used by the teacher to receive compositions, provide feedback and pair up the students for e-mail exchanges. The asynchronous nature of e-mail had extended time and space for the three EFL writers to engage in literacy practice of English anytime and anywhere at their own convenience. Because the teacher never interacted with the students via other medium, his roles were constructed solely through the written e-mail text. Therefore, an appropriate way to identify the teacher roles in this context is by investigating the teacher discourse in the written e-mail text.

## **2. Literature Review**

Computer-assisted language learning or CALL is a process in which learners use computers and as a result improve their language proficiency (Beatty, 2003). According to this rather broad definition, the applications of CALL can include word processing, computer games, corpus linguistics, computer-mediated communication, World Wide Web, and Personal Digital Assistants (PDA), etc. (ibid). Levy (1997) reviews the development of major CALL projects from a historical perspective by dividing the history of CALL into three phases: the 1960s to 1970s, the 1980s, and the 1990s. Kern and Warschauer (2000) further offer a linguistic perspective for the history of CALL: *structural approaches*, *cognitive approaches* and *sociocognitive approaches* (see Table 2.1).

**Table 2.1 - The History of CALL: A Linguistic Perspective**

Approaches	Representative Scholars	The Role of CALL	Major Program	Time Period
Structural	Bloomfield Fries Lado	Providing repetitive drill, corrective feedback, and mechanical exercise	PLATO	1960s 1970s
Cognitive	Chomsky Krashen	Offering language input and inferential tasks	Hyper-Card	1980s
Socio-cognitive	Hymes Halliday Savignon	Providing alternative contexts for social interaction; facilitating access to existing discourse communities as well as new ones	The Inter-national Email Tandem Network	1990s –

CALL programs in the 1960s and 1970s were basically designed to provide immediate feedback on grammatical accuracy of learner response (Kern & Warschauer, 2000). Initiated at the University of Illinois, the PLATO (Programmed Logic for Automatic

Teaching Operations) was a well-known project that offered language materials for repetitious drill and mechanical exercise (Beatty, 2003; Levy, 1997). This emphasis on drill-and-practice is consistent with *structural approaches* to language teaching that stress the importance of repetition to language learning (Kern & Warschauer, 2000). Influenced by behavioral psychologists such as John Watson and B. F. Skinner, language learning based on structural approaches was perceived as habit formation, a process in which pattern drills helped condition the learner to produce correct response.

CALL software applications from the 1980s tended to shift the locus of control from the computer to the learner (Kern & Warschauer, 2000). This generation of CALL saw the computer as *a tool* controlled by the learner rather than *an expert* that controlled the learner in the earlier generation. Situated within simulated environments, learners made use of computers to create their *microworlds* (Papert, 1980) in which learners were enabled to solve problems, test hypotheses and construct new concepts on the basis of their existing knowledge. This phase of CALL is in line with *cognitive approaches* to language learning which reject behaviorism's analogy of mind as a blank slate but instead assume that learners enter the classroom with "a rich set of ideas and experiences (Kern & Warschauer, 2000, p. 93).

The applications of electronic networks from the 1990s to the present stem from two important technological innovations: computer-mediated communication (CMC) and the World Wide Web (Kern & Warschauer, 2000). CMC provides language learners access to interaction with others via either asynchronous networks such as electronic mail or synchronous networks like Internet Relay Chat. The World Wide Web is a revolutionary medium for an abundance of language learning resources. The uses of electronic learning networks since the 1990s seemed to accord with the principles of *sociocognitive approaches* to language learning which stressed meaningful social interaction in authentic discourse communities (Kern & Warschauer, 2000). With the advent of computer networks, the dynamic of CALL was shifted from "learners' interaction with computers" in structural and cognitive approaches to "interaction with other humans via the computer" in sociocognitive approaches (ibid, p. 11).

Technology has been associated with learner autonomy especially when technology is taken in the broadest sense and autonomy as a super-ordinate term (Motteram, 1997). In Benson's (2001) *technology-based approaches* to the development of autonomy, he includes a good number of projects that incorporate e-mail language advising (Makin, 1994), student-produced video (Gardener, 1994), e-mail tandem learning (Lewis, Woodin

& St. John, 1996), informational CD-ROMs (Guillot, 1996), and electronic writing environments (Milton, 1997), etc. According to Benson, some of these projects provide the technology that facilitates interactions difficult or impossible in the classroom. For others, it is the interaction with technology that benefits the development of learner autonomy.

In addition, Benson (2002) sees the development of autonomy as closely tied to that of a communicative orientation towards language learning. At the University of Hong Kong, Benson and Nunan (2003) interview thirty-one freshmen on their English experiences from the earliest to the most recent learning stage (cited in Benson, 2002). They conclude that the intention of these students to take control of their language learning almost always arises out of their experiences of communication although expressing intention of taking control is not equivalent to gaining autonomy in language learning.

For Gardner and Miller (1999), the introduction of autonomy in language learning requires changes in the roles of both teachers and learners. Because the idea to incorporate autonomy in language learning typically comes from the teacher, the promotion of autonomy depends to a great extent on the teacher's redefinition of his or her own roles (Hill, 1994). Crabbe (1999) similarly suggests that a re-examination of teacher roles is essential if the learning mode of the students is to become more autonomous.

Barnes (1992) represents the potential roles of the teacher on a continuum from transmission at one end to interpretation at the other. A *transmission teacher* is one who transmits knowledge to learners. This kind of teacher is always ready to evaluate and correct the performance of learners to make sure the knowledge is successfully transmitted. On the contrary, *interpretation teachers* consider it important to help learners interpret knowledge by themselves. Their roles are to set up dialogues with learners and to help them re-organize the knowledge.

Adapting Barnes' distinction, Voller (1997) argues that teacher roles for the promotion of learner autonomy should always fall within the interpretation end of the continuum. He suggests the teacher roles in interpretation teaching to include helper (Tough, 1971), facilitator (Knowles, 1975), knower (Curran, 1976), resource (Breen & Candlin, 1980), consultant (Gremmo & Abe, 1985), counselor (Knowles, 1986), coordinator (Hammond & Collins, 1991), and adviser (Sturtridge, 1992). Following a thorough review, he classifies the teacher roles above into three categories: teacher as facilitator, teacher as counselor and teacher as resource.

According to Voller (1997), *teacher as facilitator* is a term widely used in the

literature of autonomous language learning, self-access language learning, self-instruction language learning, and communicative language learning. A facilitator of learning is generally represented as a helper who makes it easier for learning to occur. In his attempt to characterize the role of a facilitator, Voller uses Holec's (1985) two complementary roles, a provider of *technical support* and a provider of *psycho-social support*. The technical support provided by a facilitator includes 1) helping learners to plan and carry out their independent language learning by means of needs analysis, organizing interactions, etc., 2) helping learners evaluate themselves, and 3) helping learners to acquire the skills and knowledge needed to implement the above. The psycho-social support includes 1) being caring, supportive, patient, empathic, open and non-judgemental, 2) motivating learners and being prepared to enter into a dialogue with learners, and 3) raising learners' awareness of independent learning.

*Teacher as counselor* is another common role of interpretation teachers discussed in autonomous language learning although "little research has been done to determine exactly how counselors counsel" (Voller, 1997, p. 104). A counselor is a person who provides advice to those who need it. Teachers as counselors often refer to those who work in more individualized learning contexts such as the staff in self-access language learning centers. However, the term counselor has also been used in other contexts such as communicative language learning (Richards & Rodgers, 1986) and community language learning (Curran, 1976).

Regent (1993) compares the discourse of a traditional teacher in a French conversation class with that of a counselor in a self-directed learning center. She uses one extract from the teacher and two extracts from the counselor to investigate how discourse influences the development of learner autonomy. After a survey of teacher and counselor discourse, Regent comes to a conclusion by making a list of role categories that distinguish teaching from counseling. In the list, there are 19 categories in teaching and 20 categories in counseling, and as a result some of the categories do not pair up well (see Table 2.2).

**Table 2.2 - Teaching Vs. Counseling (Regent, 1993, pp. 37-8)**

TEACHING	COUNSELING
Setting objectives	Eliciting information about aims, needs & wishes
Determining course content	Why, what for, how, how long
Selecting materials	Giving information, clarifying
Determining time, place, pace	Suggesting materials
Determining learning tasks	Suggesting methodology
Determining use of L1/L2	Suggesting other sources
Managing class interaction,	Listening, responding
Initiating	Helping self monitoring
Monitoring learning situation	Interpreting information
Keeping records	Giving feedback, reformulating
Allocating homework	Suggesting organization procedures
Presenting vocabulary and grammar	Presenting materials
Explaining	Analyzing techniques
Answering questions	Offering alternative procedures
	Answering queries
Marking, grading	Suggesting self-assessment tools
Testing	Giving feedback on self-assessment
Motivating	Being positive
Rewarding, punishing	Supporting
Counseling	Putting into perspective

Borrowing Regent's (1993) categorization, Riley (1997) further analyzes the speech acts and functions that realize the roles in teaching and counseling. He re-organizes Regent's list into 15 sets of role categories and presents the roles as pairs through combining and deleting some of the categories in Regent's (see Table 2.3).

**Table 2.3 - Roles in Teaching and Counseling (Riley, 1997, p. 122)**

TEACHING	COUNSELING
1. Setting objectives	1. Eliciting information about aims, needs and wishes
2. Determining course content	2. Why, what for, how, how long: giving information, clarifying
3. Selecting materials	3. Suggesting materials, suggesting other sources
4. Deciding on time, place and pace	4. Suggesting organization procedures
5. Deciding on learning tasks	5. Suggesting methodology
6. Managing classroom interaction, initiating	6. Listening, responding
7. Monitoring the learning situation	7. Interpreting information
8. Keeping records, setting homework	8. Suggesting record-keeping and planning procedures
9. Presenting vocabulary and grammar	9. Presenting materials
10. Explaining	10. Analyzing techniques

---

11. Answering questions	11. Offering alternative procedures
12. Marking, grading	12. Suggesting self-assessment tools and techniques
13. Testing	13. Giving feedback on self-assessment
14. Motivating	14. Being positive
15. Rewarding, punishing	15. Supporting

---

Compared to those of facilitator and counselor, the role of *teacher as resource* is much less explored in the literature. Voller (1997) sees this role as comparable to that of teacher as knower, with a similar emphasis on the expertise of the teacher. Whereas teacher as knower often refers to classroom settings, teacher as resource is generally used in self-access learning contexts. However, he argues that defining a teacher role in terms of expertise in knowledge can imply an unequal power relationship between teachers and learners.

### **Research Questions**

Inasmuch as there was almost no institutional constraint from the EFI website, the teacher as investigator had the autonomy to make choices in conducting the cyber course of grammar and writing. In this pedagogical context, the only means of communication was e-mail, and therefore, the teacher roles as well as their influences on the learners could only be identified in the discourse of e-mail. To examine the relationships of teacher roles and learner autonomy in the cyber context, the study was guided by the following three research questions:

- 1) What teacher roles are constructed in the cyber course of grammar and writing?
- 2) How are the teacher roles constructed in the discourse of the written e-mail text?
- 3) How do the roles of the teacher influence the development of learner autonomy in the cyber pedagogical context?

### **3. Method**

#### **3.1. Context**

*English for Internet* (EFI), a free English school on the Internet, is a language learning website initiated and overseen by David Winet. Winet also teaches in the English Language Program at the University of California at Berkeley (extension). His concept is to recruit volunteer teachers to provide courses on EFI and then offer them free to English

learners from all over the world by matching them up with available online teachers (Stevens, 2003). During the year of 2000, EFI provided cyber courses in listening, speaking, reading, grammar and writing, as well as specialized instruction in ESP, TOEFL preparation, voice message board and online chat.

After a few e-mail exchanges with EFI coordinators, the investigator became a volunteer teacher for a cyber course of grammar and writing from February 12, 2001 to September 27, 2002. His contact with course participants was made possible by the coordinators who sent him the names and e-mail addresses of English learners from the basic to intermediate level. Since the EFI coordinators did not establish any requirements on how the course should be structured, the investigator was free to manage his own class. The contents of the assignments in this cyber course started with weekly grammar lessons and multiple-draft compositions. The grammar lessons were usually provided according to the frequent types of errors found in the learners' compositions. An e-mail pal activity was included beginning in the third month of the course when the two students in this study were paired up to exchange e-mails with each other.

In this case study, the investigator first retrospectively examined his own roles as the teacher in the cyber pedagogical context to see whether they were similar to or different from those that had been established in the literature of autonomous language learning. Two experienced English teachers were recruited to be the co-coders of the teacher roles in order to enhance the validity and inter-rater reliability of the study. Second, he examined the reactions of the two EFL adult learners in response to the teacher roles as well as the ensuing learner-teacher interaction. His main interest was to look at the locus of control between the teacher and the learners to see how teacher roles were associated with the development of learner autonomy in the cyber pedagogical context.

### **3.2. Participants**

The teacher as investigator or George (nickname) was a M.A. student at a state university in the United States. George majored in TESL and therefore became interested to have a teaching experience to apply what he learned from school. When he saw David Winet's announcement of online teaching positions, George immediately e-mailed professor Winet and volunteered to teach the cyber course of grammar and writing on EFI. Because of his full-time student status, George was only able to take two students from EFI at a time.

The students, Mick and Jing (pseudonyms) were the only two EFL adult learners who participated regularly and continuously in George's class of grammar and writing. Mick

was a Hungarian working as a production manager at a Hungarian-German joint venture company. He was the first student in this course. His participation began on February 12, 2001 and continued to June 20, 2002. Jing was a Thai homemaker raising two preschool children. She participated in the cyber course from March 6, 2001 to September 27, 2002. According to their proficiency in a grammar-focused writing test given by the EFI website, Mick was at the basic level, and Jing, the intermediate level.

### **3.3. Data Collection**

The teacher as investigator routinely kept personal e-mail messages and sorted them into different folders. Therefore, his Outlook Express stored all the e-mail correspondence and written assignments generated during the 20-month period of the instructional period. The complete data in this study included course schedules, grammar lessons, multiple-draft written assignments, online grammar assessment, online needs assessment, online course evaluation and 362 e-mail messages from the teacher and the two course participants. This study focused on the e-mail entries collected from the beginning to the end of the cyber class. The sample of 90 teacher e-mails used for NVivo content analyses spread equally between the early (Phase 1), middle (Phase 2) and end phase (Phase 3) of the 362 e-mail entries.

### **3.4. Methods of Data Analysis**

The data analysis of this study was divided into two parts: content analysis and discourse analysis. In the content analysis, a qualitative data analysis software known as NVivo 1.1-3 was applied to search for patterns of teacher discourse in terms of roles in teaching and counseling. In this pedagogical context, Nvivo allowed the investigator to manually assign units of text to categories of teacher roles in teaching and counseling. Based on the findings of the content analysis, a follow-up discourse analysis further investigated the reactions of the two learners associated with the frequent teacher roles identified in the content analysis.

Three steps were taken for the content analysis. First, the investigator did a preliminary analysis in which a set of coding schemes was built to generalize teacher discourse in a sample of 90 teacher e-mails. In coding the data, he found Riley's (1997) contrast of teacher roles in teaching and counseling to be helpful in building a coding scheme. Therefore, the coding scheme for the preliminary analysis was both derived from the data and connected to the conceptual ideas of teaching and counseling roles in the

literature.

In step 2, two experienced English teachers were recruited to be co-coders for the study. The co-coders included an Egyptian male, a Fulbright visiting scholar at the Pennsylvania State University, who had been a teacher trainer in English Education and teacher in Linguistics for more than twelve years, and a Taiwanese female, a second-year M.A. – PhD student at the Pennsylvania State University with six years' experiences teaching English as a foreign language. The two co-coders were trained individually to use NVivo 1.1-3 and code five e-mails outside the sample of 90 teacher e-mails. They both were asked to provide feedback to the coding scheme from the preliminary analysis. After the training sessions, the coding scheme underwent minor refinement based on the feedback of the two co-coders (see Table 3.1).

Finally, the three coders (including the investigator) started to use the final coding scheme to code the sample of 90 teacher e-mail messages. The results of their analyses were compared. Codings that matched for two or all three coders were recorded as the outcome of the inter-coder treatment. In other words, recorded data reflect a consensus of i) coders 1 and 2, ii) coders 1 and 3 (iii) coders 2 and 3, and (iv) coders 1, 2 and 3.

**Table 3.1 - The Final Coding Scheme**

Indicators for Teaching Roles	Indicators for Counseling Roles
1. Setting objectives	1. Eliciting and fulfilling learner aims, needs and wishes.
2. Determining course content	2. Clarifying and giving information for course content
3. Selecting materials	3. Suggesting materials and other sources
4. Deciding on time, place, and pace	4. Suggesting time line and organization procedures
5. Deciding on learning tasks and allocating homework	5. Suggesting learning tasks
6. Dominating class interaction	6. Being willing to listen and respond
7. Monitoring learning situation	7. Interpreting learning situation
8. Presenting grammar and vocabulary	8. Presenting learning materials
9. Using teaching technique	9. Analyzing learning technique
10. Providing right answers or right ways of doing things	10. Offering alternative procedures and choices
11. Marking and grading	11. Suggesting self-assessment techniques and tools
12. Rewarding and punishing	12. Supporting and being considerate
13. Keeping records	13. Giving positive or negative feedback

### 3.5. Limitations of the Study

This study involved a small population of learners in a cyber pedagogical context which allowed one-on-one instructions. The results of this study might not be generalized to the teaching contexts with larger population of students. For the sampling technique in the content analysis, the e-mail data were not sorted randomly but by time periods. The advantage of selecting samples by time is to preserve the rich context of the CMC data. On the other hand, the inter-coder treatment was taken as an alternative approach to establishing the inter-rater reliability in the content analysis. Two advantages account for the adoption of the alternative approach. First, the approach considers the perspectives of teachers or coders from different cultural backgrounds to ensure the qualitative nature of the analysis. Second, the inter-coder treatment realizes the purpose of guarding against subjective biases of individual coders and helps to enhance the inter-rater reliability in the content analysis.

## 4. Results & Discussion

### 4.1. Content Analysis

Table 4.1 reports the identification of teacher roles in *teaching* for each of the three phases in the 20-month instructional period. In a calculation of the total codings for each category of teaching roles in the three phases, the following three categories were coded more frequently than other categories of teaching roles: 1) *deciding on learning tasks and allocating homework* (21 total codings), 2) *presenting grammar and vocabulary* (3 codings), and 3) *deciding on time, place and pace* (2 codings).

**Table 4.1 - Number of Coding in Teaching Roles**

Number of Coding in	Phase 1	Phase 2	Phase 3
1. Setting objectives	0	0	0
2. Determining course content	1	0	0
3. Selecting materials	0	0	0
4. Deciding on time, place, and pace	2	0	0
5. Deciding on learning tasks and allocating homework	17	4	0
6. Dominating classroom interaction	0	0	0
7. Monitoring learning situation	0	0	0
8. Presenting grammar and vocabulary	0	2	1
9. Using teaching technique	0	0	0
10. Providing right answers or ways of doing things	0	1	0
11. Marking and grading	0	0	0
12. Rewarding and punishing	0	0	0
13. Keeping records	0	0	0
Total	20	7	1

Notes: The unit of coding is by sentence.

Phase 1 = Email 027 – 080; Phase 2 = Email 149 – 201; Phase 3 = Email 263 – 328

In contrast with teaching roles, the roles in *counseling* were identified with more frequency and in a variety of categories (see Table 4.2). The six more frequent counseling roles were as follows: 1) *giving positive or negative feedback* (33 total codings), 2) *supporting and being considerate* (26 codings), 3) *being willing to listen and respond* (18 codings), 4) *eliciting and fulfilling aims, needs and wishes* (12 codings), 5) *offering alternative procedures and choices* (11 codings), and 6) *suggesting learning tasks* (9 codings).

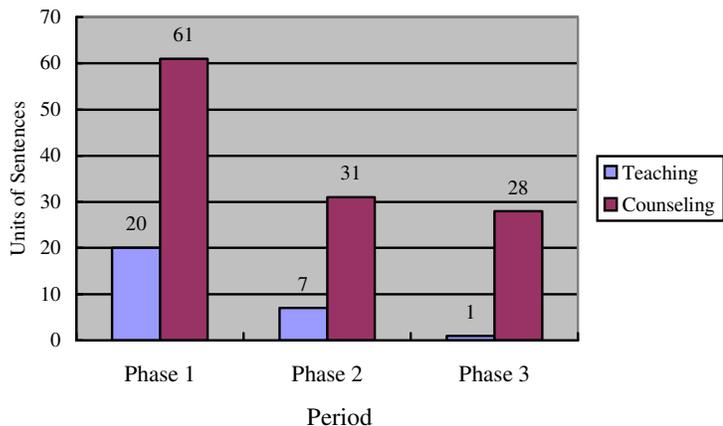
**Table 4.2 - Number of Coding in Counseling Roles**

Number of Coding in	Phase 1	Phase 2	Phase 3
1. Eliciting and fulfilling aims, needs and wishes	7	5	0
2. Clarifying and giving information for course content	1	1	0
3. Suggesting materials and other sources	2	0	1
4. Suggesting time line and organization procedures	0	1	0
5. Suggesting learning tasks	1	4	4
6. Being willing to listen and respond	6	7	5
7. Interpreting learning situation	4	0	0
8. Presenting learning materials	0	0	0
9. Analyzing learning technique	0	0	0
10. Offering alternative procedures and choices	7	2	2
11. Suggesting self-assessment techniques and tools	1	0	0
12. Supporting and being considerate	13	6	7
13. Giving positive or negative feedback	19	5	9
Total	61	31	28

Notes: The unit of coding is by sentence.

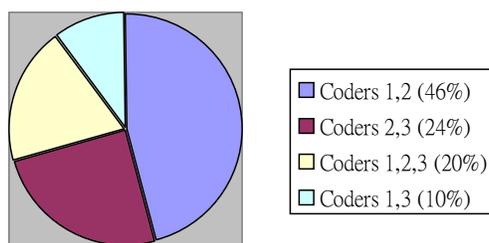
Phase 1 = Email 027 – 080; Phase 2 = Email 149 – 201; Phase 3 = Email 263 – 328

Figure 4.1 illustrates the total codings of roles in teaching and counseling for each phase of the instructional period. In *teaching* roles, the total codings decreased from 20 (Phase 1) to 7 (Phase 2) and 1 (Phase 3). With respect to *counseling* roles, the total number of codings also decreased, but to a much lesser degree. As was true for the roles in teaching, the counseling roles in Phase 1 significantly outnumbered the total codings for the other two phases (61, 31, 28). Nevertheless, the difference in the total codings in counseling roles between Phase 2 and Phase 3 was much less than that for teaching roles. In other words, the teacher's teaching roles became less active as the course progressed whereas the counseling roles remained active throughout the instructional period.



**Figure 4.1: Total Codings in Teaching & Counseling Roles by Phase**

Figure 4.2 shows the ratios of coding agreement reached by the three coders during the inter-coder treatment. A high percentage (46%) of coding was matched between coder 1 (male) and coder 2 (female) who both came from Taiwan. The coding matches between coder 2 (Taiwanese female) and coder 3 (Egyptian male) were 24%. The least agreement was reached between coder 1 (Taiwanese male) and coder 3 (Egyptian male) at only 10%. This difference seemed to suggest that the descriptions of teacher acts and discourse were not universal, although this study involved such a small sample of coders that nothing significant can be assumed. The coding agreement reached by the three coders together comprised 20% of the total.



**Figure 4.2 Coder Agreement**

Coder 1 was a Taiwanese male PhD student.  
 Coder 2 was a Taiwanese female PhD student.  
 Coder 3 was an Egyptian male visiting scholar.

## **4.2. Discourse Analysis**

Linked to the content analysis, a follow-up discourse analysis was conducted to investigate how the two EFL adult learners responded to the teacher roles in the e-mail interactions. Therefore, the discourse analysis focused on the three frequent teaching roles identified in the content analysis: 1) *deciding on learning tasks and allocating homework*, 2) *presenting grammar and vocabulary* and 3) *deciding on time, place and pace*, and the following six frequent counseling roles: 1) *giving positive or negative feedback*, 2) *supporting and being considerate*, 3) *being willing to listen and respond*, 4) *eliciting and fulfilling aims, needs and wishes*, 5) *offering alternative procedures and choices*, and 6) *suggesting learning tasks*. These role categories were ordered in coding frequency and included examples from the beginning, middle and end phases of the instructional period.

### **4.2.1. The Discourse of Teaching Roles**

#### *Teaching Role 1: Deciding on Learning Tasks and Allocating Homework*

The five sentences highlighted in E-mail 37 were coded as the teaching role, *deciding on learning tasks and allocating homework*. These teacher instructions required the Thai learner, Jing to complete a grammar lesson and a narrative composition. She was asked to complete an exercise on the past tense and to use past tense in the composition whenever necessary. Several imperatives were found in these teacher instructions. By using imperatives, I directed the learner in exactly what to do and showed her the focus of each task. Therefore, an important linguistic characteristic in realizing this teaching role was through the use of imperatives.

E-mail 37

From: George  
Date: 3/13/2001  
To: Jing  
Subject: lesson 1

“Dear Jing,

Now we only have 2 weeks left for March. Lesson 1 is for the following week and lesson 2 for the last week.

Lesson1: Regular Vs Irregular Verb

Lesson2: To be (linking verb)

Please go to this link and read Lesson1. #1 There is an introduction of verbs. You can skip the exercise of irregular past verb tense if you are not interested. Let me know if you prefer a doc file. In that case, I'll send you doc file next time.

<http://www.personal.psu.edu/users/c/y/cyc109/lesson1.htm>

1. Send me your answers if you do the exercise on past tense verb.
2. Write a story about yourself. It can be a trip, your family, or even English learning experience. Focus on one thing and be sure to use past tense when it is

appropriate. No limit with words. Be informal. Let me know if the lesson is too hard or too easy for you.  
George”

Notes: #1 was coded as a different teaching role, *selecting materials* by coder 3.

In her return mail, Jing followed all of my instructions coded as the teaching role of *deciding on learning tasks and allocating homework*. She submitted her answers for the grammar exercise on the past tense. Choosing one of the topics I provided, she wrote a composition about the trip to Chiang Mai with her husband. Throughout her composition, past-tense verbs were used wherever appropriate (see highlighted verbs in E-mail 44). The content of Jing’s e-mail was simply on the task, and it made her e-mail look like an answer sheet for a written examination. There was nothing related to free-topic conversation or dialogue. Her response seemed to reflect the role of teacher as authority figure who exerted control to learners.

E-mail 44

From: Jing  
Date: 3/20/2001  
To: George  
Subject: Re. Lesson1

“Dear teacher George,  
(...)

Here are the exercises of lesson 1 below:

1. Suddenly, Bill's horse (see) saw a shadow. It (spring) sprang sideways and (rise) rose on its hind legs, raking at its invisible foe with its hooves. Bill, (catch) caught unprepared, was (fling) flang flung to the rocky ground with a sickening thud. At that moment lightening (split) split the heavens, and an image of Bill as he (lie) lay crumpled and bleeding was (freeze) froze in the ice-blue light. I (leap) leap leapt down from my saddle but as I (bend) bent close to Bill, I (can) could see that for him life's long journey (is) was all over.

Sorry, there are two mistakes and I've corrected it into black letters at once.

2. I'll never forget the day we, only my husband and I, went to Chiang Mai, the second largest city in Thailand. It's 18th Nov 1993, not long after I came to Thailand. It's also one of the most important Thai national festivals, called Loy Krathong. We started at afternoon on Nov 17, prepared a lot of snack, drinking water and record tapes for the long trip. We drove through the city of Bangkok, busy traffic made me have a chance to enjoy its old and modern streets, buildings, temples, etc. After driving out of the city, the view changed, wide rice field, green woods and blooming flowers came into our sights. How beautiful! I asked my husband about the festival--Loy krathong, he told me Thai people laid floating floral tributes lit by candles on the water--pond, river, lake or sea, and they believed it could take the past year's bad things away and bring their luck for future. We talked and talked, when it became dark, we found we were missing. We should drove to north, but turned east

unconsciously. We **had** to stay at a hotel in a very small city--Sinburi. The next day, we **went** back to the right way and **drove** up straightly, **arrived** in Chiang Mai late afternoon finally.”

(....)

#### Teaching Role 2: Presenting Grammar and Vocabulary

E-mail 281 provides an example of *presenting grammar and vocabulary*, a teaching role assigned by coders 2 and 3. I consulted my co-worker in the inter-library loan office at the Pennsylvania State University to clear up a misunderstanding originated from a form feedback I provided to Jing’s description of her job (as a nurse) in a multiple-draft essay. After receiving comments from my friend, I confirmed with Jing who proposed that it did not seem to be grammatically accurate to describe her job as “mix the medicine or mix the pills” in a previous e-mail.

E-mail 281

From: George

Date: 11/2/2001

To: Jing

Subject: prescribe the pills

“Hi Jing,

I have asked my library colleague and he cannot think of a good term for collect the medicine. However, he suggested to say "mutleplize the pills" or "give the pills" or make multiple pills and so on. He said it's right to say, "compound the medicine but he seldom puts it this way. **But it's wrong to say, "mix the medicine or mix the pills."** #2 Also you may say prescribe the pills. I think they have different kinds of saying because in America, they have pharmacists who are responsible to prescribe pills for the patients. According to my understanding, a good way to describe, "compound the medicine" in America is "prescribe the pills." Let me know if anything sounds confusing to you.

George”

Notes: #2 was coded as a counseling role, *giving positive or negative feedback* by coder 1.

In E-mail 283, Jing replied to me that she had looked over again for a right phrase to explain her job in the hospital, but without success. She finally looked it up in the dictionary and found the same expression “make up the prescription” that she used to describe her job originally in a previous draft. However, Jing was not completely confident herself but asked me for confirmation on the usage. Jing’s response to the form feedback seemed to reflect a role of teacher as provider of *correct answers* which was very typical in Taiwan where the control of *form* was always on the side of the teacher

(Wang, 2000).

E-mail 283

From: Jing

Date: 11/4/2001

To: George

Subject: Re. prescribe the pills

“Dear teacher George,

I've just sent my job 3 to you. I tried to search the answer about the question, but I failed.

However, I found it in my dictionary, and it's "make up the prescription". Do you think it's right?

Yours, Jing”

*Teaching Role 3: Deciding on Time, Place, and Pace*

E-mail 53 has a coding on the teaching role of *deciding on time, place, and pace*. I decided a due date for Mick to submit his essay 3. The time control was pretty loose because I did not set up an exact time. “The week after next week” in fact could range from two weeks to three weeks. Writing a multiple-draft essay was a very time-consuming process, and that was why I decided to control the pace of the assignment.

E-mail 53

From: George

Date: 3/30/2001

To: Mick

Subject: Re. feedback3 to essay2

“Dear Mick,

(...) As for your third essay, "fishing" is fine but I would suggest you some direction to write about it. Think about how this hobby means to you. Do you consider it a good hobby or bad hobby? How often do you go fishing? With whom? Maybe you can also tell some stories that happened at the time when you were doing this hobby. And you can say more! Same requirement: no limits with words. **Due date: the week after next week.** By the way, I include lesson 6 in the attachment and the answers for the exercise are listed in the end.

George”

Mick sent his essay 3 two weeks later than the due date (see E-mail 103). He apologized for the delay because he had been very busy at that time. Mick's delay was unusual because he often turned in his written assignments sooner than I expected. However, if I had had institutional power to enforce the due date, “being busy” would not have become a justifiable reason for Mick to miss it.

E-mail 103

From: Mick

Date: 5/2/2001

To: George

Subject: essay3

Attachment: fishing2.doc

“Hi George,

I apologize for my delay, as you know I was very busy recently.  
I finished my third essay and I am sending it to you.

(...)

Best regards

Mick”

#### **4.2.2. The Discourse of Counseling Roles**

##### *Counseling Role 1: Giving Positive or Negative Feedback*

The texts highlighted in E-mail 40 were all coded in the category of the counseling role, *giving positive or negative feedback*. They established the teacher’s feedback in the brief e-mail response to the Hungarian learner, Mick, who had submitted essay 2 earlier on during that same day. The comments were all positive feedback which served as a general evaluation of the essay. In trying to convince him of the good quality of the essay, I attempted to give Mick confidence in his own writing ability. The linguistic realization of the counseling role in this e-mail suggested that I tended to provide only positive feedback to help the learner. This supported Regent’s (1993) inclusion of “being positive” as one category of counseling roles.

E-mail 40

From: George

Date: 3/14/2001

To: Mick

Subject: Re. Essay2

“Mick,

This essay is really nice! The description of your job is very clear and the information is very thoughtful. I love reading this.#3

(...)

George”

Notes: #3 was coded as a different counseling role, *supportive and being considerate* by coder 1.

Mick expressed his appreciation for the praise in his response (see E-mail 41). He was motivated to re-read the essay and make more revisions. Yet that same day, March 14<sup>th</sup>, he sent a revised draft of essay 2 to me, thanks to the technology of e-mail, a tool for free and efficient delivery. Mick’s immediate response suggested that he was encouraged by the praise to take the initiative and do more to improve the quality of his essay.

E-mail 41

From: Mick

Date: 3/14/2001

To: George

Subject: Re. Essay2  
Attachment: aboutmyjob.doc

“Hi George,  
Thank you very much for your praise about my second essay. I read it again and I found some small mistakes what I repaired. I am sending you it and please read this version when you will check it.  
(...)  
Best regards  
Mick”

*Counseling Role 2: Supporting and Being Considerate*

The sentence highlighted in E-mail 149 was coded both as *supporting and being considerate*, and *giving positive or negative feedback*. I told Mick that he did a great job on the essay of Hungarian food. Without much knowledge of Hungarian cuisine, I had a question for the last paragraph of the essay. I was not able to understand the recipe for a particular dish and wondered whether the food was some kind of paste. My provision of this counseling role seemed to create a supportive or non-threatening working climate for Mick so that he would not feel too bad when I expressed difficulty in comprehending his writing. The counseling role therefore accords with the psycho-social support of a facilitator who motivates learners by being empathic and supportive (Voller, 1997).

E-mail 149  
From: George  
Date: 6/23/2001  
To: Mick  
Subject: feedback to essay4  
Attachment: feedback1toessay4.doc

“Dear Mick,  
**You did a great job on the essay about Hungarian food.**<sup>#4</sup> I do not know Hungarian cuisine is so famous in European countries. As you say, French food is famous as well, but I don't like it so much. Maybe it's because the Fresh bread. I did see two children fight each other with French bread on the street. Their bread is so hard and flavorless when compared to English cake. I do not quite understand the last paragraph of this essay since I am very unfamiliar with Hungarian food. Is the food some kind of paste? Please read it over before doing any revision. Thanks!  
George”

Notes: <sup>#4</sup> was coded in this role category by coders 1 & 3 but it was also coded as another counseling role, *giving positive or negative feedback* by coders 1 & 2.

In two days, Mick sent a revised draft of the essay to me and hoped the new draft would be more comprehensible. Mick clarified that the food was not any kind of paste. Because of cultural differences, he could not find a good term to translate it into English.

E-mail 153

From: Mick

Date: 6/25/2001

To: George

Subject: Re. Feedback to essay4

Attachment: feedback1hungarianfood.doc

“Hi George,

(...)

I did over my essay. I think it became more comprehensible. I agree with your opinion that writing about food is difficult. The Hungarian cuisine is really famous in Europe. It is an attraction for many European tourists.

The mentioned food in the last paragraph is not a kind of paste. I didn't find a correct word in dictionaries to translate it exactly. Perhaps I would name it potato-squash or potato-mush. It is a folk dish from West Country of Hungary. I think it is true that must go to Hungary if someone wants to eat a pure Hungarian food. I had bad experience when I ate Hungarian food in Poland. It was terrible.

I am sending improved version of my essay.

Best regards, Mick”

### *Counseling Role 3: Being Willing to Listen and Respond*

The sentence highlighted in E-mail 158 were assigned to the counseling role of *being willing to listen and respond*. I introduced Jing to the procedure of an e-mail writing project after gaining her feedback on the project. I planned to let her develop her multiple-draft essays from her previous e-mail pal exchanges. In other words, the topics of her new multiple-draft essays would be selected from the e-mail pal exchanges, which served as preliminary drafts. In this project, I would provide feedback through attachment like what I did in the multiple-draft written assignments.

E-mail 158

From: George

Date: 6/29/2001

To: Jing

Subject: email project

“Dear Jing,

I have read the feedback for email project and also have more ideas with the project. I am thinking to use one of the emails in the emails you sent each other weekly to give my feedback through doc. file. In this way, the piece can replace our essay assignment and ease the workload. (...) **You can either agree or tell the reasons for disagreement.**

(...)

George”

Jing liked my ideas for the e-mail writing project; however, she was a little confused about the procedure of the new activity. She was not sure how the feedback would be provided and whether she should continue the email pal exchanges (See E-mail 167).

Through the provision of this counseling role, I was able to hear what Jing thought about the new activity and make sure whether the procedure was clear for her to understand. However, the counseling role did not engage Jing in decision-making of the e-mail writing project and did not provide evidence for transference of control to the learner.

E-mail 167

From: Jing

Date: 7/5/2001

To: George

Subject: Re. Email project

“Dear teacher George,

I like your idea, and your feedback must be useful and helpful to both of us. Thank you very much! Only one question, you said we would go on to next email when one was satisfactory, does it mean that we'll begin to write next email or you'll begin to give next feedback when one is satisfactory?

Best Wishes,

Jing”

*Counseling Role 4: Eliciting and Fulfilling Learner Aims, Needs and Wishes*

One sentence was coded under the counseling role, *eliciting and fulfilling learner aims, needs, and wishes* in E-mail 29, where Jing was requested to give a self-introduction. I asked Jing why and what she wanted to learn in this cyber course offered on EFI. The use of WH question suggested my willingness to hear Jing's voice and have dialogue with her. By asking *why*, I was able to investigate her learner needs, interests and wishes. The WH question thus served as an important medium of technical support (Voller, 1997) for the teacher to transfer the locus of control, letting the learner have her say in the course.

E-mail 29

From: George

Date: 3/6/2001

To: Jing

Subject: self-introduction

“Hi Jing

(...) Before we start our lessons, I would like you to give a brief introduction about yourself. Something related to your country, family, favorite stuffs, and your school. **I'm also interested to know why and what you want to learn English in EFI.**  
**#5** There is no limit with words for the self-introduction. Please be informal! I look forward to hearing from you.

George”

Notes: **#5** was coded as a teaching role, *setting objectives* by coder 3.

The WH question allowed Jing to express her reasons for participating in the cyber writing course. Three of the reasons she gave had to do with her needs to use English for communication (see E-mail 33). For example, she had a close friend who was married to

an English gentleman; however, Jing could hardly talk with him when they visited her on holidays. Therefore, she applied to this class immediately upon learning the address of EFI from a book about the Internet. As we can see, this counseling role not only empowered Jing to recognize her communicative needs but also helped me to interpret what she wished to gain from this class. The counseling role of *eliciting and fulfilling learner aims, needs, and wishes* seemed to serve as a prior step for the teacher to hand over the locus of control to the learner.

E-mail 33

From: George

Date: 3/11/2001

To: Jing

Subject: Re. Self-introduction

“Dear teacher George,

(...)

About the reason of learning English in EFI, there are so many. At first, I think learning will not be end or over for a man, and his heart will never be old if he studies new knowledge constantly. I have been learning English myself after graduated, and do my best to keep what I've learned though there are not many chances for me to practice. I'm not good at it, particularly at speaking and listening. The second, I want to look for a job when my kids get older, a good level in English will be more helpful. The third, English is used and learned by people all over the world, it 's a very useful language, especially on line. Some friends of mine always write to me in English, and I have to write back in English too since they can't read my message in Chinese or Thai in their computer. The more reason is one best friend of mine, her husband is a very lovely and kind English gentleman, they often come and spend holidays here, but I nearly can't talk with him. When I got the address of EFI from a book about Internet, I made an application at once.

(...)

Best Regards,

Jing”

#### *Counseling Role 5: Offering Alternative Procedures and Choices*

The text highlighted in E-mail 49 offers an example of the counseling role, *offering alternative procedures and choices*. Instead of giving Mick a final topic for essay 3, I allowed him to have a topic of his own choice. By offering this alternative procedure for the topic selection, I was able to share with Mick the decision as to the topic for essay 3.

E-mail 49

From: George

Date: 3/26/2001

To: Mick

Subject: Re. Book review

“Dear Mick

(...) I think you are doing very well with this class and as for essay 3, I haven't

decided a final topic. Maybe you can suggest something you are interested to write about such as your favorite sports, foods, holidays or anything else. (...)  
George”

In E-mail 50, Mick conveyed his decision to write about his hobby for essay 3. However, he wanted to make sure that he could begin his writing. Although Mick was given the locus of control as to the topic for essay 3, there was still some concern on his part for the need to seek permission from the teacher.

E-mail 50  
From: George  
Date: 3/28/2001  
To: Mick  
Subject: Re. Lesson5 and feedback2;  
Attachment: feedback2toessay2repaired

“Hello George,  
(...)  
Can I start to write my third essay about my hobby?  
I am waiting for your reply.  
Best regards  
Mick”

In E-mail 51, I asked Mick what kind of hobby he was going to write about for essay 3. By asking this WH question, I was able to help Mick narrow the topic. In other words, even after the learner had chosen to write about his hobby, I still had a role to play in negotiating a more specific topic.

E-mail 51  
From: George  
Date: 3/29/2001  
To: Mick  
Subject: feedback3 to essay2;  
Attachment: feedback3toessay2.doc

“Hi Mick  
(...) As for essay3, can you tell me what kind of hobby you'd like to write about? Let me know more about what you want to write about this topic because it must be interesting for me to read. I'll send you Lesson 6 right after we together design the topic for next essay. Thanks!  
George”

In E-mail 52, Mick finally decided on the more specific topic of (line) fishing, a topic I would have never thought about because of my vegetarian diet. Nevertheless, as in E-mail 50, the learner expressed concern for my opinion. He told me to suggest other topics if his own choice did not prompt my interest. Here again we see that Mick still considered the choice of essay topic to be subject to the teacher's control even though he had made the

final decision to write about his favorite hobby, fishing.

E-mail 52

From: Mick

Date: 3/29/2001

To: George

Subject: Re. Feedback3 to essay2

“Hi George,

(...) I think I'll write about line fishing, it is my hobby. If it isn't interesting for you, please suggest me other topic. I'm sending you version 3 of essay 2 on Monday.

Have a nice weekend.

Mick”

#### *Role 6: Suggesting Learning Tasks*

In E-mail 57, a sentence was coded under the counseling role, *suggesting learning tasks*. Here I proposed an e-mail pal activity to the students in the cyber course. This activity was aimed at developing the learners' communicative ability in writing. Participants in this activity needed to send an e-mail message to each other on a weekly base, including a copy for me because I agreed to read every copied message in the e-mail exchanges.

E-mail 57

From: George

Date: 4/2/2001

To: Jing; Arturo; Mick

Subject: grammar & writing learning group

“Hi folks,

**I'm designing an interactive class to develop my EFI students' communicative ability in writing.** The class is more like an email learning group and very informal. The only thing to participate in this class is to send or respond at least an email each week to two other students of you three. Of course, you can send more when you get excited. You should also send a duplicate to me whenever you send an email to other students. I'll read every message from your interaction. (...)

George”

Both Mick and Jing considered it a good idea to participate in the e-mail pal activity. They appeared to be happy and motivated in their responses (see E-mails 61 & 64). Mick, in particular, was curious about the English level of his e-mail pal and hoped that together with his partner they could find common topics. In other words, Mick felt that the locus of control as to topic decision was between him and his partner in the communication-based activity of e-mail exchanges. This was an example of the connection between computer-mediated communication and the intention of taking control in language learning. It supports Benson and Nunun (2003) who assert that the intention of learners to take control of their language learning almost always arises out of their experiences of

communication. Through the e-mail pal activity, the two EFL adult learners were able to be in charge of their learning while the teacher became a reader or a secondary participant.

E-mail 61

From: Mick

Date: 4/3/2001

To: George

Subject: Re. Grammar & writing learning group

“Hi George,

I think, it is a good idea. I like it and I gladly participate in this group. I don't know the others, but I hope we will find common topics. Which level do they speak English?

What are their opinions about this group?

I am curious to see the rules.

Best regards

Mick”

E-mail 64

From: Jing

Date: 4/6/2001

To: George

Subject: Re. Grammar and writing learning group

“Dear teacher George,

What a good idea! I like it, so certainly I'll participate.

Jing”

In the beginning of the cyber course, I often played a managerial role of *deciding on learning tasks and allocation of homework*. However, my class became less structured as soon as the e-mail pal activity was included. As Van Lier (2005) suggests, learning activities should be carefully structured earlier on and less structured later in the process of handing over the locus of control to the learners. However, learner autonomy does not imply learning without a teacher. The teacher should step in and take control whenever there is a need to play a teaching or managerial role.

Whether their contributions are direct or indirect, counseling roles are essential for the teacher to promote learner autonomy. The provision of a counseling role such as *suggesting learning tasks* had directly helped me to transfer the locus of control to the learners. As soon as I introduced the e-mail pal activity, the learners were engaged in decision-making and became more self-directed. Other types of counseling roles might indirectly contribute to the development of learner autonomy. For example, the counseling role of *giving positive feedback* helped me to create a safe and supportive learning environment in which the learners became more comfortable to take over the control. The counseling role of *eliciting and fulfilling aims, needs and wishes* in this class served as a prior step to the development of learner autonomy. As Nunan (1996) argues in his

discussions of autonomous language learning, at the very least, teachers should find out what learners think about what they want to learn and how they want to learn it.

### **5. Implications**

Autonomy is enhanced when language learners are actively involved in the productive use of language, and when they are given opportunities to engage in the learning tasks that activate their language outside the classroom (Nunan, 1996; Savignon, 1997). Teachers should be encouraged to introduce language learners to *task-based learning activities* (Ellis, 2003) that foster autonomy, for example: group presentations, conversational partners, language games, online discussion board, e-mail pal project, and learning in tandem, etc. By promoting these learning activities, teachers will be able to transfer the locus of control to learners and help them become autonomous language learners who use language to learn and communicate, thereby demonstrating a capacity to take control of their own learning.

This study provides evidence that supports the connection between computer-mediated communication and learner autonomy in language learning. In his response to a counseling role of *offering alternative procedures and choices*, Mick seemed to feel that the decision as to a composition topic was subject to the teacher's control even after he was given the opportunity to have a topic of his own choosing. However, as soon as the e-mail pal activity was introduced, Mick actively expressed his intention to find common topics between him and his partner. This implies that his intention of taking over the locus of control as to topic decision is connected to the use of language for communication. Different from a composition assignment, the e-mail pal activity provides the learners with genuine audience to establish two-way communication.

As an important element in a writing activity, topic or theme directs or even constrains the content and organization of writing. In a composition task, the decision-making of a topic is usually controlled by the teacher and that leaves the learners little room for negotiation. In contrast, an e-mail pal activity affords the learners freedom in the decision-making of a topic in a process of writing for the purpose of communication. The free or negotiable topic choice in e-mail exchanges thus contributes to the promotion of autonomy and helps learners to gain more control of their own writing. This ownership of learner text is essential in the development of an independent writer. As we enter the electronic age of teaching L2 writing, the use of e-mail pal activity is beneficial to create an authentic discourse community that enhances the social dynamics of a L2 writing class.

In this study, the CMC technology of e-mail created a democratic climate for decentralization. The free topic choice in e-mail pal exchanges was asserted to promote learner autonomy in L2 writing since it gave learners more control of their own writing. The e-mail pal activity was self-directed because it was led by the learners. The language learning in this activity was inter-dependent as long as the decision-making of a topic was shared by the learners. Besides e-mail, MOO is another CMC technology that has been associated with the promotion of autonomy in language learning. Other new CMC technologies such as instant messaging, Wikis and blogs, and chat rooms require further investigation to examine how they are related to the promotion of learner autonomy in language learning.

Regent (1993) and Riley's (1997) categorization of discourse in teaching and counseling was beneficial to identify the teacher roles in this cyber course and might be further applied to investigate the teacher roles in other settings of communicative language teaching or CLT. As CLT is now popular, especially in East Asian countries, there is a need to look more thoroughly at the role of the teacher in this major approach to language learning. For example, now that CLT is officially promoted in Taiwan, the roles of Taiwanese English teachers await investigation. Their roles need to be adjusted to take on a learner-centered approach in an environment where learner competition tends to be high and grammar-based instruction remains prevalent (Wang, 2000).

With respect to the inter-coder treatment, the results suggest that the descriptions of teacher acts and discourse are not universal, and the judgments of teacher roles can be influenced by the cultural factors of the evaluators. In fact, the teacher roles that have been established in the literature of autonomous language learning are basically dominated by Western scholars and may not represent the perspectives of teachers from other cultural backgrounds. However, explanation of this assertion of cultural bias requires further research with teachers or coders from different cultural backgrounds to see to what extent they agree or disagree with the categories of teacher roles that appear in discussions of autonomous language learning.

## **6. Conclusion**

This study has provided evidence of the relationship between teacher roles and learner autonomy in a cyber English course of grammar and writing. Assertions drawn from the findings imply that teaching roles need to be reduced whereas counseling roles are essential to the promotion of autonomy in language learning. In addition, this study relates

the concept of learner autonomy to the nature of language learning by asserting a connection between computer-mediated communication and the intention of taking control in language learning. However, the assertion is made in a specific pedagogical context of grammar and writing where the issue of topic control comes to the fore. The promotion of this assertion requires further investigation into other contexts of language teaching and learning. The results of the content analysis shows that the teacher's teaching roles become less active as the course progresses whereas the counseling roles remain active throughout the instructional period. Content analysis also calls into question the universality of established categories of teacher roles, suggesting that cultural context and experience need to be taken into consideration.

### **References**

- Appel, D., & Mullen, T. (2000). Pedagogical considerations for a web-based tandem language learning environment. *Computers and Education, 34*, 291-308.
- Balester, V., Halasek, K., & Peterson, N. (1992). Sharing authority: Collaborative teaching in a computer-based writing course. *Computers and Composition, 9*(3), 25-39.
- Barker, T., & Kemp, F. (1990). Network theory: A postmodern pedagogy for the written classroom. In C. Handa (Ed.), *Computers and community: Teaching composition in the twenty-first century*. Portsmouth, NH: Heinemann.
- Barnes, D. (1992) *From communication to curriculum*. 2<sup>nd</sup> ed. Portsmouth, NH: Heinemann.
- Beatty, K. (2003). *Teaching and researching computer-assisted language learning*. London: Longman.
- Benson, P. (2001). *Teaching and researching autonomy in language learning*. Harlow: Longman.
- Benson, P. (2002). Autonomy and communication. In P. Benson, & S. Toogood (Eds.), *Learner autonomy 7: Challenges to research and practice*. Dublin: Authentik.
- Breen, M. P. & Candlin, C. (1980). The essentials of a communicative curriculum in language teaching. *Applied Linguistics, 1*(2), 89-112.
- Camilleri, A. (1999). *Introducing learner autonomy in teacher education*. Strasbourg: Council of Europe Publishing.
- Benson, P. & Nunan, D. (2003). *What Hong Kong University students know about English and learning English*. English Centre Monograph No.3. Hong Kong: English Centre.

- Crabbe, D. (1999). Defining the field: Introduction. In S. Cotterall, & D. Crabbe (Eds.), *Learner autonomy in language learning: Defining the field and effecting change*. Frankfurt: Peter Lang GmbH.
- Curran, C. A. (1976). *Counselling-learning in second languages*. IL: Apple River.
- Ellis, R. (2003). *Task-based language learning and teaching*. Oxford: Oxford University Press.
- Faigley, L. (1990). Subverting the electronic workbook: Teaching writing using networked computers. In D. Baker, & M. Monenberg (Eds.), *The writing teacher as researcher: Essays in the theory and practice of class-based research*. Portsmouth, NH: Heinemann.
- Gardner, D. (1994). Student-produced video documentary: Hong Kong as a self-access resource. *Hong Kong Papers in Linguistics and Language Teaching*, 17, 45-54.
- Gardner, D., & Miller, L. (1999). *Establishing self-access: From theory to practice*. Cambridge: Cambridge University Press.
- Graddol, D. (1997). *The future of English*. London: British Council.
- Gremmo, M. J. & Abe, D. (1985). Teaching learning: Redefining the teacher's role. In P. Riley (Ed.), *Discourse and learning*. London: Longman.
- Guillot, M. (1996). Resource-based language learning: pedagogic strategies for Le Monde sur CD-ROM. In E. Broady, & M. Kenning (Eds.), *Promoting learner autonomy in university language teaching*. London: Association for French Language Studies/CILT.
- Hammond, M. & Collins, R. (1991). *Self-directed learning: Critical practice*. London: Kogan Page.
- Hill, B. (1994). Self-managed learning: state of the art. *Language Teaching*, 27, 213-223.
- Holec, H. (1985). On autonomy: some elementary concepts. In P. Riley (Ed.), *Discourse and learning* (pp. 173-90). London: Longman.
- Kern, R., & Warschauer, M. (2000). Theory and practice of network-based language teaching. In M. Warschauer, & R. Kern (Eds.), *Network-based language teaching: Concepts and practice*. Cambridge: Cambridge University Press.
- Knowles, M. (1975). *Self-directed learning: A guide for learners and teachers*. Cambridge: The Adult Education Company.
- Knowles, M. (1986). *Using learner contracts*. San Francisco: Jossey-Bass.
- Levy, M. (1997). *Computer-assisted language learning: Context and conceptualization*. Oxford: Oxford University Press.

- Lewis, T., Woodin, J., & St. John, E. (1996). Tandem learning: Independence through partnership. In E. Broady, & M. Kenning (Eds.), *Promoting learner autonomy in university language teaching*. London: Association for French Language Studies/CILT.
- Little, D. (1991). *Learner autonomy 1: Definitions, issues and problems*. Dublin: Authentik.
- Little, D., & Brammerts, H. (1996). *A guide to language learning in tandem via the Internet*. Dublin: Trinity College.
- Makin, L. (1994). Learner telesupport: Language advising by email. In E. Esch (Ed.), *Self-access and the adult language learner*. London: CILT.
- Milton, J. (1997). Providing computerized self-access opportunities for the development of writing skills. In P. Benson, & P. Voller (Eds.), *Autonomy and independence in language learning*. Harlow: Longman.
- Motteram, G. (1997). Learner autonomy and the web. In V. Darleguy, et al. (Eds.), *Educational technology in language learning: Theoretical considerations and practical applications*. Lyons: INSA.
- Nunan, D. (1996). Towards autonomous language learning: some theoretical, empirical and practical issues. In R. Pemberton, E. Li, W. Or, & H. Pierson (Eds.), *Taking control: autonomy in language learning*. Hong Kong: Hong Kong University Press.
- Papert, S. (1980). *Mindstorms: Children, computers, and powerful ideas*. New York: Basic Books.
- Regent, O. (1993). Communication, strategy and language learning. In J. L. Otal, & M. L. Villanueva (Eds.), *Primeves Jornades Sobre Auto-aprenentage de Llengues*. Castello Publicacions de la Universitat Jaume I.
- Richards, J. C., & Rodgers, T. S. (1986). *Approaches and methods in language teaching*. Cambridge: Cambridge University Press.
- Riley, P. (1997). The guru and the conjurer: aspects of counselling for self-access. In P. Benson, & P. Voller (Eds.), *Autonomy and independence in language learning*. Harlow: Longman.
- Savignon, S. J. (1997). *Communicative competence, theory and classroom practice: Texts and contexts in second language learning*. 2<sup>nd</sup> edition. New York: McGraw-Hill.
- Schwienhorst, K. (1997). *Talking on the MOO: Learner autonomy and language learning in tandem*. Paper presented at the CALLMOO: Enhancing language learning through Internet technologies, Bergen, Norway.

- Stevens, V. (2003). Blending communities of practice and CMC tools in teacher training. Retrieved March 1<sup>st</sup>, 2004 from [http://www.geocities.com/vance\\_stevens/papers/apacall/profdev\\_cop.htm](http://www.geocities.com/vance_stevens/papers/apacall/profdev_cop.htm)
- Sturtridge, G. (1992). *Self-access: Preparation and training*. Manchester: British Council.
- Tough, A. (1971). *The adult's learning projects*. Toronto: Ontario Institute for Education Studies.
- Van Lier, L. (2005, May). *The ecology and semiotics of online communication*. Paper presented at 2005 CALICO Conference, Michigan, USA.
- Voller, P. (1997). Does the teacher have a role in autonomous language learning? In P. Benson, & P. Voller (Eds.), *Autonomy and independence in language learning*. Harlow: Longman.
- Wang, C. C. (2000). *A social-linguistic profile of English in Taiwan: Social context and learner needs*. Unpublished Doctoral Dissertation.
- Warschauer, M., Turbee, L., & Roberts, B. (1996). Computer learning networks and student empowerment. *System*, 24(1), 1-14.



## **Mentoring Beginning EFL Teachers at Tertiary Level in Vietnam**

**Nguyen Thi Mai Hoa,**  
*University of Queensland, Australia and Vietnam National University, Vietnam*

### **Bio Data:**

Nguyen Thi Mai Hoa received her MA in Tesol (University of Queensland) and M.Ed. in Educational Management and Leadership (RMIT). She has worked as an EFL teacher, teacher trainer, materials writer, and researcher at the College of Foreign Languages, Vietnam National University in Hanoi for several years. She is currently working towards her Ph.D. at the University of Queensland, Australia.

### **Abstract**

Mentoring is a prominent approach used in teacher professional development in the world, especially in the U.S.A, U.K. and other Western countries but is not fully employed in beginning teacher education in Vietnam. Therefore, this study examines the current situation of mentoring at university and investigates beginning English language teachers' attitudes and feelings toward their mentoring experiences. It is argued that conducting such studies can raise the awareness of mentoring in the process of guiding beginning English language teachers in their new profession. Data for the study was collected by a survey questionnaire which was given to a number of beginning teachers. Descriptive statistical of frequency, mean, standard deviation were applied to examine the data from the survey. The results of the study suggested that although beginning English language teachers in Vietnam valued the role of mentoring in their professional development, it is not fully employed in teacher education for many reasons. Their thoughts about their mentoring can be a blueprint for future implementation of a formal mentoring program.

### **Introduction**

Among different approaches such as mentoring, coaching and supervision, mentoring seems to be very effective in enhancing teachers' professional development (Carter & Francis, 2001; Johnson, 2002; Portner, 2002; Yost, 2002). Mentoring is a process in which a more skilled or more experienced person, the mentor, nurtures someone less skilled or experienced, the mentee. This may involve a mentor adopting a variety of roles of modeling, teaching, sponsoring, encouraging, counseling, or befriending (Anderson and Shannon, 1988), coaching, supporting, promoting (Lacey, 1999), or even supervising and providing situational leadership (Colwell, 1998) to a mentee for the purpose of promoting the latter's professional and/or personal development (Colwell, 1998; Anderson and Shannon, 1988; Peterson & Williams, 1998).

Inexperienced teachers are often regarded as teachers who have less than three years of teaching experience (Troutman, 2002). This group of teachers always struggle for survival in their early career. Additionally, first-year teaching experiences have powerful influences on their practice and attitudes throughout the remainder of their careers. It can be seen that, without help, most first-year teachers often “lose their enthusiasm, ambition, and idealism and start getting lost in the flurry of a challenging beginning” (Saban, 2002, p.33). Therefore, providing some kind of support to them is essential to retain them within the profession and to develop them as potential professionals.

However, the emergence of mentoring as a professional development strategy has not been paid due attention in Vietnam. Moreover, there has been almost no research into the field of mentoring in the context of the Vietnamese educational system. No study exists that examines the presence and benefits of mentoring for beginning English language teachers in Vietnam. Therefore, this study set out to investigate the use of mentoring as a professional growth strategy and the beginning English language teachers’ opinions about their mentoring experience at their workplace.

### **Benefits of mentoring in beginning teachers’ professional development**

Literature reveals that there tends to be a general acceptance that mentoring yields benefits for all involved parties – the organization, mentees and mentors. However, this part only discusses the benefits of mentoring in developing beginning teachers’ professional development.

Through mentoring, mentees have opportunities to learn, grow and move along life’s professional and psychological pathways (Galbraith & Cohen, 1995, Chao, 1997, Smith & Ingersoll, 2004). Studies of mentoring (Galbraith & Cohen, 1995; Kohler et al, 1997; Street 2004), in many settings, especially in schools, colleges and universities have showed that mentoring enhances learning. More specifically, mentoring has a great potential of supporting teacher learning because it occurs in the direct context of teaching and learning and provides real learning environments for teachers to develop (Meyer, 2002). In a mentoring process, new teachers require the assistance of more experienced colleagues as they enter the teaching profession as a student teacher. Huberman (1993) has characterized the beginning stage of teacher professional development as “a period of survival and discovery”(p.3) during which beginning teachers must face the challenges of their teaching career, such as “emotional and psychological stress”, “the lack of support”, and “conceptual struggles about teaching and learning”(Wang & Odell, 2002, p.514). As a

consequence, they may seek help from mentor teachers who provide information and opportunities to help the novice to practice, explore and improve the current situation. In a recent study of new teachers, Maynard (2000) discovered that what the new teachers wanted “above anything else was to be made to feel welcome” (p.21). In effect, beginning teachers are learning to do the job of teaching while experiencing a mentoring relationship. In recent studies (Street, 2004; Forbes, 2004; Woullard & Coats; 2004), researchers have proved that mentoring new teachers assists their teacher learning. More specifically, mentoring programs provide structure and support in helping new teachers to learn, thereby promoting a higher level of teaching skills, and pedagogical knowledge (Saban, 2002; Hargreaves & Fullan, 2000; Forbes, 2004; Evertson & Smithey, 2000; Woullard & Coats, 2004) as well as teaching efficacy (Yost, 2002). In this perspective, mentoring has often functioned as a process of transmission in teacher professional learning (Carter and Francis, 2001).

Miller (2002) argues that using isolated approaches such as coaching, tutoring or training which aim only at improving the trainees’ technical skills or knowledge is not sufficient to promote trainees’ personal and professional growth. For this reason, the importance of establishing mentor-mentee relationships is emphasized in order to understand the mentees’ levels of personal and professional development (Hargreaves & Fullan, 2000; Miller, 2002; Lacey, 1999). The purpose of this is to help the mentor “gain knowledge” or understand the protégé’s ego, moral and conceptual levels so as to design appropriate mentoring models responsive to the mentees’ levels of development (Reiman & Thies-Spintall, 1998, p.178). This supporting relationship is also helpful to develop mutual respect on both sides (Gaskin, Lumpkin & Tennant, 2003), thus facilitating the process of professional development, especially for beginning teachers.

Mentoring also brings a dramatic improvement in beginning teachers’ self-confidence and job satisfaction. Several studies have investigated the relationship between mentoring, teacher confidence and job satisfaction. Feiman-Nemser (1998) argues that “the promise of mentoring lies not in easing novices’ entry into teaching but in helping them confront difficult problems of practice and the use of their teaching as a site for learning” (p.19). The learning ability to overcome problems and to reflect upon them develops their confidence and their job satisfaction. In this regard, Douglas (1997) and Walker & Stott (1994) believe that mentoring can increase teachers’ self confidence about their teaching competence. In other studies, it is seen that employees who are more satisfied with their job tend to be more committed to their work (Burke & McKeen, 1997), thus making

efforts in developing their profession. In brief, confidence and job satisfaction can serve as a catalyst for the growth within teacher professions.

Although there are still many other benefits of mentoring for mentees, the above-mentioned literature has shown that mentoring in general and mentoring for beginning teachers in particular is a powerful way to foster beginning teachers' professional development for their positive career-related outcomes. Current research also suggests that mentoring not only supports teacher learning but also serves as important strategies for the professional development of new teachers. Blair-Larsen (1998) also asserts that a mentoring process is considered to be productive when it provides support to new teachers and opportunities for personal and professional reflection and development.

Although there is a large body of literature about the benefits of mentoring, there are also studies which investigate the barriers to mentoring. The next section explores further insights into this aspect.

### **Barriers to mentoring**

Theoretically, mentoring is a very positive model for personal and professional development. However, the practice of mentoring faces barriers and the problems of potential misuse. These concerns have been revealed in Long's study (1997) in which the researcher has shown a cautious attitude toward mentoring. According to Long, "under various conditions, the mentoring relationship can actually be detrimental to the mentor, mentee or both" (p.115).

The first barrier is related to the lack of understanding of the mentoring process such as the role of mentoring, the goal of mentoring (Gratch, 1998), and the planning of mentoring process (Holt, 1982; Tellez, 1992). The goal of mentoring is to enhance the mentee's ego, moral and conceptual development (Reiman & Thies-Spinthall, 1998). But in reality, it is shown that the practice of mentoring is, in many cases, not to enhance mentees' knowledge but rather to socialize them into the conservative social organization and the norms tied to this organization (Gratch, 1998). The same author also points out that the early stages of mentoring mostly focus on providing information about the system rather than consultation on curriculum and instruction, and also much emphasis is placed on "comfort and harmonious relations along with the norms of schools"; as a consequence, such factors "constrain mentors from posing tough questions about practice" (p. 222). A problem of poor planning is expressed in Johnsrud's (1991) study in which the researcher has found that mentoring did not meet with success due to personal and organizational

barriers. This is partly shared by Balassa et al (2003) who also prove that participants in a mentoring program had strong concerns with institutional and organizational problems inherent in the present system of education. Additionally, Wang (2001) also believes that “the reformed-minded teaching practice that the mentors developed does not necessarily guarantee the effective mentoring that supports teacher learning and teaching reform” (p.51). The reason is that both the mentor’s and mentee’s practices are constrained by the mentoring contexts such as the structure of the curriculum and assessment, the organization of teaching and the student population. Because of this, Wang (2001) suggests considering how to restructure school contexts and help mentors learn how to mentor before designing mentoring programs and arranging mentoring relationships.

The second concern of mentoring practice in educational settings is the potential lack of skills of the mentors themselves. Many researchers (Gratch, 1998; Rowley, 1999; Colwell, 1998) show that most mentors lack formal training in mentoring. Mentorship is a unique role where not everyone who volunteers may be suitable (Daresh & Playko, 1990). Although there is no best approach for training mentors, Holloway (2001) believes that even a nurturing mentor is not enough to effectively help novice teachers unless the mentors have received training on how to support new teachers. In a study, Rowley (1999) argues that mentor teachers who have not participated in a quality training program often display a lack of dedication to the role and responsibilities of mentoring. In addition, Colwell (1998, p.318) indicates that untrained mentors may become only “poorly trained buddies” unless there is an extensive system of in-service mentor training. For Gratch (1998, p.224), those “teachers who are good at teaching children may not be qualified to teach teachers”. Such perspectives show that there is a crucial need for an increased training for mentor teachers. Without substantial support for developing mentoring skills, mentor teachers may act more like guides who help novices adapt to the existing school systems, rather than as efficient supporters to assist novices in developing ambitious teaching practice (Wang, 2001).

Furthermore, several studies (Gerstein, 1985; Madison et al, 1993) have identified that a poor relationship between the mentor and mentee can often lead to a breakdown in the mentoring relationships. This can occur because mentoring process involves interpersonal relationships (Madison et al, 1993). Where a mentoring program is forced, the mentoring relationship formed may “lead to kind of contrived collegiality” (Long, 1997, p.122). This relationship is not inherently beneficial because, to some extent, it can undermine trust and openness, which are of great importance to the establishment and the maintenance of a

real collaborative culture.

Another barrier is that mentoring requires enormous investment of time, money, effort, and resources (Manson, 1990; Blank & Sindelar, 1992; Saban, 2002). Concerning time, a group of studies (Noe, 1988; Redmond, 1990; Saban, 2002) have shown that it is a critical factor for the effectiveness of mentoring. Ganser (1999) concurs with this, adding that successful mentoring also depends on the resources allocated to it and time made available for mentoring activities.

It is obvious that mentoring is a beneficial process for beginning teachers, but needs further research into how to overcome its barriers. Among many factors leading to its success, mentors play an important role. The next section discusses the functions and characteristics of mentors.

### **Mentor qualities and expectations**

According to Hutto et al, (1991, as cited in Sinclair, 2003, p.79), a mentor is defined as an “experienced, successful and knowledgeable professional who willingly accepts the responsibility of facilitating professional growth and support of a colleague through a mutually beneficial relationship”. It can be clearly seen that mentors should have outstanding knowledge, skills, professional competence in a particular field. It is expected that mentors pass on their accumulated knowledge to the less experienced teachers as well as being available for general discussion. The mentors are required to help mentees develop teaching skills, giving constructive feedback and having more general discussions about teaching with mentees, meaning that they should be highly proficient teachers with a strong base of pedagogical knowledge and successful experience (Blank & Sindelar, 1992; Kay & Hinds, 2002; Evertson & Smithey, 2000).

Mentoring has been defined as a “nurturing process” which occurs in a caring and ongoing relationship between more experienced teachers with those who are less experienced (Peterson & Williams, 1998; Colley, 2003). To achieve effective mentoring, Galbraith & Cohen (1995) and Sinclair (2003) emphasize the importance of interpersonal skills and communication skills. In this respect, several interpersonal skills are identified, including the ability to motivate, listen, influence, counsel, manage time, and to exhibit trustworthiness (Kay and Hinds, 2002; Beyene, et al, 2002). Feiman-Nemser & Parker (1992) also add that mentoring roles and functions should include those of a motivator, teacher, role model, supporter, counselor, advisor, demonstrator, guide, change agent, companion, and coach.

Not every one possesses the personality, professional knowledge, and numerous skills to be a mentor. Fundamentally, they should be personally committed to the mentoring role and believe in the potential of the mentee (Beyene, et al, 2002; Sinclair, 2003). They should be able to provide information and assistance, model appropriate practice and provide positive, sensitive feedback regarding the mentee's development and progress. Therefore, a mentor's ability to adapt to different situations is critical (Sullivan, 1992; Klausmeier, 1994), as are skills in modeling effective teaching strategies and managing time (Klausmeier, 1994), and giving feedback (Beyene, et al, 2002), all of which are regarded as indispensable in increasing the effectiveness of mentoring.

In conclusion, there is no ideal type of mentor. According to Blank & Sindelar (1992), "mentors appear to be a unique blend of intuitive sensitivity and technical expertise" (p.23). The presence of a mentor alone is not enough; the mentor's knowledge and skills of how to mentor are also crucial.

### **The relevance of current research in the literature**

Literature has revealed that mentoring has enormous potential to bring about learning, professional growth, and development for teachers, especially beginning teachers. Although there have been barriers, with appropriate understanding of the mentoring process, it has emerged as a potentially effective means of facilitating the entry of beginning teachers into the profession. The literature review of previous studies in the field has provided the background in the field of mentoring research and can be considered as basis for the investigation of mentoring in the context of the four major colleges at Vietnam National University, Hanoi. Located in Hanoi, the capital of Vietnam, VNU is the largest multidisciplinary higher educational and research centre in Vietnam. It consists of four major colleges, several schools, faculties, and institutes. In response to the call for innovations in education in Vietnam from the Ministry of Education and Training, VNU has proposed several schemes to develop teachers' professional development, in which mentoring seems to be under-explored.

Additionally, there has not been any research into this issue in the context of Vietnam, leading to the inadequate use of mentoring as a tool for professional development for beginning teachers in Vietnam. Besides, while mentoring has been recommended, research provides little information regarding how beginning teachers experience informal mentoring. In other words, the literature review indicates the need for more research in the area of mentoring for beginning teachers, particularly in the area of mentoring in Vietnam.

This study adds to the existing body of literature by investigating the current situation of mentoring for beginning English language teachers at a leading university in Vietnam and examining the beginning EFL teachers' opinions about their mentoring experiences.

## **Method**

### **Objectives of the study**

The research attempts to answer the following research questions:

1. To what extent is mentoring for beginning English language teachers employed at their workplace in Vietnam?
2. What are the benefits of mentoring in terms of their professional development, as perceived by beginning English language teachers?
3. What are the barriers to mentoring, as perceived by beginning English language teachers?
4. How do the beginning English Language teachers evaluate their mentors?

### **Participants**

The researcher employed a stratified random sampling technique to select the data. This technique was employed because the population of this study is stratified on more than one variable (it includes 4 groups of beginning English teachers from 4 different colleges of VNU). Therefore, a random sample of a proportionate size was selected from each group to ensure its reliability. According to Dornyei (2003), this kind of sampling can "minimize the effects of any extraneous or subjective variables that might affect the outcome of the survey study" (p.73).

The data for this study were obtained from 31 beginning English language teachers at four major colleges which are part of Vietnam National University, Hanoi (VNU). Respondents used for this study met the following criteria: 1) having to be employed as full-time teachers at VNU, 2) having less than 3 years of teaching experience, 3) volunteering to participate in the study. Beginning EL teachers were chosen for the survey because the study attempts to investigate their opinions and feelings about their mentoring experiences. Additionally, every year, VNU recruits a large number of English teachers. Therefore, the number of young and new English Language teachers is larger than that of other subjects.

The survey was mailed to 40 EFL beginning teachers. Accompanying the survey was a consent form for the participants, a letter outlining the purpose of the survey, the intent of

using their response in the data collection, a statement of anonymity and confidentiality and an explanation of related terms. After the first mailing, 20 questionnaires were returned. After the second mailing, 20 questionnaires were returned. Nine incomplete responses were deleted. Therefore, 31 questionnaires were used for data analysis.

Respondents in the study included 29 females (93 %) and 2 males (approximately 7 %). Among these respondents, 10 teachers from the College of Foreign Languages (VNU) accounted for 32 % of the total. Approximate 23% and 26% were from the College of Natural Sciences and the College of Social Sciences and Humanities (VNU) respectively. Six teachers (19%) were from the College of Technology (VNU).

### **Instruments**

A survey in the form of a questionnaire was developed on the literature reviewed and associated with the related objectives. The researcher used a Likert Scale and a Semantic Differential Scale to develop the questionnaire. According to Dornyei (2003), it is the most popular technique in requiring the respondents to make an evaluation judgment of the target by means of a scale. The design of this questionnaire considered such matters as avoiding “leading” questions, as suggested by Oppenheim (1992) and Dornyei (2003), in order to increase the reliability of the data. The survey solicited attitudes towards various aspects of mentoring: benefits, barriers, types, mentoring activities, mentors, which attempted to answer four research questions.

Forty five questions were developed based on an extensive review of the pertinent literature. This survey has identified major issues concerning the research questions. Question 1 and 2 identified the role of mentoring in beginning teachers’ professional development. Question 22-28 helped to recognize the type of mentoring occurring in Vietnam. Question 29-35 measured the frequency of mentoring activities they had experienced. These questions required respondents to use the following semantic differential scale: 1= always, 2= frequently, 3=sometimes, 4=infrequently, 5=never. These groups of questions helped answer research question 1. Questions 3-21 gathered information regarding the benefits as well as the barriers, which answers the research question 2 and 3. These questions required respondents to use the following Likert scale: 1= Strongly agree, 2= Agree, 3=Slightly agree, 4= Disagree, 5= Strongly disagree. Question 36-45 investigated their opinions about their mentor, which is the focus of research question 4. These questions asked respondents to use the following Likert Scale: 1=very good, 2=good, 3=satisfactory, 4=poor, 5= very poor. This survey instrument

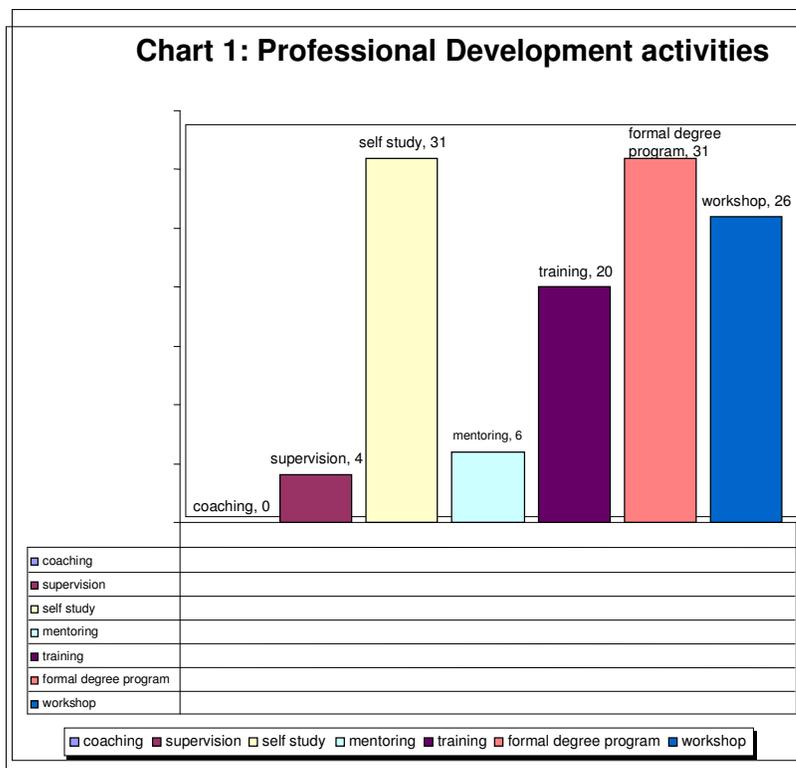
provided the opportunity for quantitative feedback from the respondents. All of these questions are found in table 1 to 5.

## **Results**

### ***Activities for beginning teachers' professional development (PD) and the role of mentoring in their professional development***

It can be seen from chart 1 below, among many mentioned professional development activities asked such as coaching, supervision, self-study, mentoring, training, formal degree program, workshops, and others (Question 1), self-study and formal degree program rank the first. Interestingly, all respondents said that they had used these activities, but three important strategies (coaching, supervision, and mentoring) for professional development were under-explored regardless of their perceived roles in their professional development. It is noted that apart from coaching and supervision, mentoring seems to be a missing link in their professional development as only 19 % teachers employed this as a professional development activity. The data in the next section also confirms that this type of mentoring is informally self-initiated.

Although mentoring is not largely used by the teachers, they still value its importance in their professional development. When asked about its role (Question 2), most of the respondents said that it was important, even very important in their professional development.



***Benefits of mentoring in their professional development***

Items in Table 1 reported beginning English language teachers’ (BELTs) responses regarding their opinions about the benefits of mentoring. All subjects were asked to respond to the questions on a Likert Scale: 1= strongly agree, 2= agree, 3=slight agree, 4= disagree, 5= strongly disagree. The scores for each item were ranked by mean and standard deviation to indicate how BELTs felt about it.

Respondents indicated that overall they agreed that mentoring helped them feel more confident about their professional competence and acquire more teaching knowledge and teaching skills. In addition, they also believed that teacher learning and teaching improvements were benefits of their mentoring experiences. They slightly agreed that it had positive influences on their teaching efficacy, their job satisfaction and their career-long profession. Responses to these issues yielded results that ranged from agree to slightly agree with standard deviation of approximately 0.5. Most respondents also slightly agreed that mentoring established a foundation for their career-long profession. However, there is a variation in this answer as its standard deviation is 0.68. In general, the perceived benefits of mentoring support the major findings in the literature. This can create good conditions to continue the development of the mentoring relationship in their

workplace for their positive attitudes toward its role in their professional development.

**Table 1. Opinions towards the benefits of mentoring**

	N	Minim um	Maxi mum	Mean	Std. Deviation
<b>3. Mentoring helps you feel more confident about your professional competence</b>	31	1.00	3.00	1.7419	.51431
<b>4. Mentoring helps you acquire more teaching knowledge</b>	31	1.00	2.00	1.5484	.50588
<b>5. Mentoring helps you acquire more teaching skills</b>	31	1.00	3.00	2.1613	.45437
<b>6. Mentoring supports teacher learning</b>	31	1.00	3.00	1.8387	.52261
<b>7. Mentoring helps you gain more teaching experience</b>	31	1.00	3.00	1.8065	.60107
<b>8. Mentoring helps you improve your teaching</b>	31	1.00	3.00	2.0323	.48193
<b>9. Mentoring helps you increase your teaching efficacy</b>	31	2.00	3.00	2.4839	.50800
<b>10. Mentoring increases your job satisfaction</b>	31	2.00	3.00	2.4839	.50800
<b>11. Mentoring establishes a foundation for your career-long profession</b>	31	2.00	4.00	2.8387	.68784
<b>Valid N (listwise)</b>	31				

### **Barriers to mentoring**

Items listed in Table 2 reported BELTs' responses toward the barriers to their mentoring process. All subjects were asked to respond to questions regarding this issue on a Likert scale: 1= strongly agree, 2= agree, 3=slight agree, 4= disagree, 5= strongly disagree. The scores for each item were ranked by mean to indicate level of agreement. From the table, it can be seen that BELTs agreed that lack of awareness of the role of mentoring, understanding about the mentoring process, and access to mentors represented barriers to their mentoring process. Additionally, there was agreement particularly in the area of resources associated with mentoring, limited time for mentoring and unsuccessful matching of mentor and mentee. They only slightly agreed that their mentoring process was hindered by the lack of administrative support. However, this needs further investigation because its standard deviation is high (0.8), which means there is a fluctuation in their answers. Also, responses to barriers such as mentor's poor skills, poor planning and insufficient funding yielded scores that ranged from "slightly agree" to

“agree”. In general, most of the respondents slightly agreed or agreed that mentoring in their workplace still faces many barriers. These findings are of practical value in taking into consideration these factors in an attempt to extend mentorship for beginning EFL teachers.

**Table 2. Opinions towards the barriers to mentoring**

	<b>N</b>	<b>Minimu m</b>	<b>Maxi mum</b>	<b>Mean</b>	<b>Std. Deviation</b>
<b>12. Lack of awareness of the role of mentoring</b>	31	1.00	3.00	1.7742	.56034
<b>13. Your mentors’ poor skills in mentoring</b>	31	2.00	3.00	2.1935	.40161
<b>14. Limited time for mentoring</b>	31	1.00	3.00	1.7419	.57548
<b>15. Poor planning of the mentoring process</b>	31	2.00	4.00	2.3871	.55842
<b>16. Lack of understanding about the mentoring process</b>	31	1.00	3.00	1.9032	.59749
<b>17. Lack of access to mentors</b>	31	1.00	3.00	1.8065	.60107
<b>18. Insufficient funding</b>	31	1.00	3.00	2.2258	.66881
<b>19. Lack of resources associated with mentoring</b>	31	1.00	3.00	1.8387	.52261
<b>20. Unsuccessful matching of mentors and mentees</b>	31	1.00	3.00	2.0000	.57735
<b>21. Lack of administrative support</b>	31	2.00	4.00	3.1935	.83344
<b>Valid N (listwise)</b>	31				

### **Mentoring types**

The items in Table 3 investigated the types of mentoring the respondents experienced. The respondents agreed or even strongly agreed that they themselves initiated their mentoring relationship and this relationship seemed to be supported by the organization. Additionally, they indicated that there was no procedure to match mentees with mentors and they did not attend a defined mentoring program at their college. This was consistent with the next two items which indicated that respondents disagreed that they experienced a mentoring program with specific tasks for short periods and clear purposes. These results have showed that they experienced informal mentoring.

**Table 3 - Opinions towards their mentoring types**

	N	Min.	Max.	Mean	Std. Deviation
<b>22. Your mentoring relationship is created spontaneously or informally without any assistance from the organization.</b>	31	1.00	2.00	1.5484	.50588
<b>23. You select the mentors for your self</b>	31	2.00	4.00	2.7097	.78288
<b>24. You have attended a defined mentoring program at your working place</b>	31	4.00	5.00	4.3548	.48637
<b>25. There is a procedure to match mentees with mentor</b>	31	3.00	5.00	4.0968	.59749
<b>26. Your mentoring programs have clear purposes</b>	31	3.00	4.00	3.4839	.50800
<b>27. Your mentoring program focuses on specific tasks for short periods</b>	31	3.00	5.00	3.9032	.65089
<b>28. Your mentoring relationship is facilitated and supported by the organization</b>	31	2.00	5.00	2.9677	.91228
Valid N (listwise)	31				

**Mentoring activities**

The items listed in Table 4 examined the frequency of the mentoring activities they had experienced. Overall, most of the respondents reported that they never experienced mentor modeling and class observation. Some of them experienced these activities infrequently. Additionally, most of the respondents also indicated that they infrequently shared their experience with the mentors and counseled them. Mean score for guidance and training from mentors were respectively 3.1 (sd 0.61) and 3.3 (sd 0.54), which showed that they sometimes employed these mentoring activities. The sole activity with higher frequency was asking for mentor’s advice. In general, asking for a mentor’s advice was the most frequent mentoring activity while other activities seemed not to occur frequently. This seems to verify that a need exists for formalizing more mentoring activities to develop beginning EFL teacher’s professionalism.

**Table 4 - Mentoring activities**

	N	Min.	Max.	Mean	Std. Deviation
<b>29. Asking for mentor's advice</b>	31	1.00	4.00	2.6452	.66073
<b>30. Mentor modeling</b>	31	4.00	5.00	4.7097	.46141
<b>31. Sharing experience with the mentors</b>	31	3.00	5.00	3.5806	.62044
<b>32. Class observation( mentor observes your class)</b>	31	4.00	5.00	4.6452	.48637
<b>33. Counseling</b>	31	2.00	5.00	3.6774	.79108
<b>34. Guidance</b>	31	2.00	4.00	3.1290	.61870
<b>35. Training</b>	31	2.00	4.00	3.3226	.54081
Valid N (listwise)	31				

### **Evaluation on mentors**

The items listed in Table 5 reported respondents' evaluation on their mentor. All subjects were asked to respond to questions regarding their evaluation on several qualities of their mentor on a Likert scale ranging from 1- 5 (1= very good, 2= good, 3= satisfactory, 4= poor, 5= very poor). The score for each item was ranked by mean to indicate the level of their evaluation.

Overall, most of the respondents indicated that their mentor's ability to manage time and give feedback was poor. Additionally, they also reported that they had poor accessibility to the mentor. Although the mean of items 36 and 37 was almost the same, its range was different. It can be inferred that most of the respondents were satisfied with their mentor's ability to model effective teaching strategies. However, their opinions about their mentor's ability to work in a collaborative manner ranged from "satisfactory" to "poor". Responses to the mentor's ability to adapt to different situations were also ranged from "satisfactory" to "good". There was not much agreement in this item. They also felt satisfied with their mentor's verbal communication. Additionally, they also thought that their mentor's professional competence, interpersonal skills, and teaching experience were good.

**Table 5 - Opinions towards the mentors**

	N	Min.	Max.	Mean	std. Deviation
<b>36. Ability to model effective teaching strategies</b>	31	2.00	4.00	3.2903	.52874
<b>37. Ability to work in a collaborative manner</b>	31	3.00	4.00	3.2258	.42502
<b>38. Ability to adapt to different situations</b>	31	2.00	3.00	2.5161	.50800
<b>39. Ability to manage time for mentoring effectively</b>	31	3.00	5.00	4.0968	.53882
<b>40. Ability to give feedback</b>	31	3.00	5.00	4.0645	.57361
<b>41. Professional competence</b>	31	1.00	3.00	2.0256	.50588
<b>42. Verbal communication</b>	31	2.00	4.00	2.9355	.57361
<b>43. Interpersonal skills</b>	31	1.00	3.00	2.1194	.56416
<b>44. Teaching experience</b>	31	1.00	3.00	2.0000	.36515
<b>45. Accessibility of the mentor</b>	31	3.00	5.00	3.8065	.54279
<b>Valid N (listwise)</b>	31				

### **Discussion**

Overall, the research revealed that there is a lack of awareness about mentoring, especially formal mentoring at tertiary level in Vietnam. The result of the present study showed that beginning English language teachers at their workplace experienced only informal mentoring. This type of informal mentoring was identified in their opinions about their mentoring experiences, types of mentoring activities, and their evaluation of their mentor. The study confirmed findings from other research in relation to the benefits of mentoring in teacher professional development. Although these benefits were perceived by only BELTs in Vietnam, it supports the use of mentoring as a strategy to help beginning teachers in their early career. In addition, they also identified barriers to their mentoring process such as lack of awareness of the role of mentoring, understanding about the mentoring process, mentors, time, resources, poor mentoring skills, planning and matching between mentors and mentees. Although the administration supported their mentoring, it failed to go beyond informal support. These findings are of practical importance for future implementation of a mentoring program at tertiary level in Vietnam. Additionally, data analysis from the survey indicated that the current beginning English language teachers at four colleges of Vietnam National University, Hanoi did not feel satisfied with their mentor's mentoring skills in such areas as giving feedback, working in a collaborative

manner, modeling, effective teaching strategies, managing time for mentoring. The BELTs found it hard to access their mentor. However, they have a high regard for their mentor's professional competence, teaching skills, teaching experience, and interpersonal skills. These findings are also of practical value in examining the current situation for the future design and implementation of a mentoring program for beginning English language teachers.

### **Recommendations**

The findings suggest that mentorship is a potentially important component of an initial teacher's professional development for its tremendous benefits; however, mentoring is not formally advocated as a tool for professional development for beginning English language teachers. It occurred spontaneously to fill a need that existed. Therefore, it is highly recommended that a mentoring program be implemented at this research site to assist, facilitate, and ease the transition of the beginning teachers in to their potential profession. This is essential in providing beginning EFL teachers with opportunities to learn 'how to learn' through the process of trial and error, feedback and reflection. As a result, this approach is effective in helping improve beginning teachers' professional development. Additionally, mentoring can show its greater effectiveness under two conditions: firstly, if the university envisages the creation of learning workplace-based culture in which its "members have learned conscious, communal processes for continually generating, retaining and leveraging individual and collective learning to improve the performance of the organizational system in ways important to all stakeholders" (Teare & Dealtry, 1998, p.47); and secondly, if that new culture can monitor and improve performance (Drew & Smith, 1995). A learning culture can encourage all the teachers, especially beginning teachers to continue developing their professional skills and participate actively in teacher collaboration. It can be sure that mentoring will be welcome in this environment.

Furthermore, in order to promote both formal and informal mentorship, it is necessary to take all the aforementioned barriers into consideration. More specifically, the desired mentoring programs should have a clear communication of the program's objectives, clarification of roles such as training mentors, a setting of expectations, suggesting activities, creating procedures to match mentor with mentee, and the appropriate allocation of resources, and planning.

Through the beginning English language teachers' evaluation of their mentors, the study emphasizes the importance of training prospective mentors of new teachers. The

study provides further insight into the issue of the identification and selection of mentors for beginning teachers. The following criteria are recommended for choosing a mentor for beginning English language teachers. Firstly, they need to possess abilities to model effective teaching strategies, to work in a collaborative manner and to manage time effectively. These abilities are of great importance because the main job of a mentor is to provide expertise and ongoing support and professional development opportunities to enhance skills and effectiveness of beginning English language teachers (Brooks, 1999). Additionally, they need to have a good knowledge of teaching and have more than three years of successful teaching experiences. This parameter is necessary to help the beginning teachers acquire more teaching knowledge. More importantly, they should display good professional competence and effective verbal communication, which enable them to share their expertise with their mentees and to communicate effectively. Furthermore, interpersonal skills are also significant in effective mentorship because mentoring is a nurturing process of guidance and the mentor's role is non-evaluative and non-judgmental. With good interpersonal skills, mentors can build a better relationship with their mentees. However, all of these qualities alone are not enough to develop effective mentorship; the mentor's knowledge and skill of how to mentor are also crucial. In other words, they need adequate training in the role of mentorship. In a recent study, Evertson & Smithey (2000) found that "protégés of trained mentors showed increased evidence of developing and sustaining more workable classroom routines, managed instruction more smoothly, and gained student cooperation in academic tasks more effectively" (p.301). Similarly, according to Evertson & Smithey (2000) and Carter and Francis (2001), the training of mentors in more useful mentoring skills and the preparation of mentors for their task enables them to be more successful in their role. Consequently, it is recommended that mentoring programs should place more emphasis on the training of mentors. This should involve formal training on the goals and objectives of mentoring; what a teacher mentor is supposed to do; which role they play, and what skills they need to demonstrate. Zachary (2000) believes that not only do mentors need training prior to working with their mentees, but also the opportunities to discuss ideas, problems, and solutions with other mentors while continuing to support their mentees.

## **Conclusions**

Even though this research has showed several limitations associated with a relatively small number of respondents (N=31) from a small number of single organizations, thus

limiting the scope of the study, it draws an overview picture of this issue at four colleges of Vietnam National University in Vietnam. Data from this research demonstrated that mentoring for beginning English language teachers at Vietnam National University was missing in initial teacher education and was just employed informally. In this respect, it outlines the need to develop a formal mentoring program for beginning EFL teachers. In the study itself, findings confirmed those from other research in relation to the benefits of mentoring in teacher professional development. In addition, barriers to their mentoring at the research site were also identified by beginning EFL teachers, which is of practical importance for future implementation of a mentoring program. Regarding the evaluation of their mentors, it is noted that they did not feel satisfied with their mentoring skills and their mentor's availability. This could be explained by the fact that their mentoring relationships were informally initiated. Thus, their mentors were not trained in mentoring. From what has been found about the current situation, the above mentioned recommendations have been put forward to support the call for a formal mentoring program for beginning EFL teachers, as well as to create conditions which can facilitate its success.

Although this study's findings demonstrated that mentoring at VNU requires further improvement and attention, the researcher suggests a further longitudinal study be conducted on beginning teacher mentoring experiences. Furthermore, an additional means of analysis of the situation through observation and the compilation of personal records could be conducted. Future studies should analyze a larger sample of beginning English language teachers, thus extending the scope of the context and strengthening the validity of its findings.

## **References**

- Anderson, E. M., & Shannon, A. L. (1988). Toward a conceptualization of mentoring. *Journal of Teacher Education, 39*(1), 38-42.
- Beyene, T., Anglin, M., Sanchez, W., & Ballou, M. (2002). Mentoring and relational mutuality: Protégés' perspectives. *Journal of Humanistic Counseling, Education and Development, 41*(1), 87-102.
- Balassa, K., Bodoczky, C., & Saunders, D. (2003). An impact study of the national Hungarian mentoring project in English Language Training. *Mentoring and Tutoring, 11*(3), 307-320.
- Blair-Larsen, S. M. (1998). Designing a mentoring program. *Education, 118*(4), 602- 605.

- Brooks, M. (1999). Mentors matter. In M. Scherer (Ed.), *A better beginning: Supporting and mentoring new teachers*. Alexandria, Va: Association for Supervision and Curriculum Development.
- Burke, R. J., & McKeen, C. A. (1997). Benefits of mentoring relationships among managerial and professional women: A cautionary tale. *Journal of Vocational Behavior, 51*, 43-57.
- Carter, M., & Francis, R. (2001). Mentoring and Beginning teachers' workplace learning. *Asia-Pacific Journal of Teacher Education, 29*(3), 249-261
- Colley, H. (2003). *Mentoring for social inclusion: A critical approach to nurturing mentor relationships*. London: Routledge Falmer.
- Colwell, S. (1998). Mentoring, Socialization and the mentor/protégé relationship [1]. *Teaching in Higher Education, 3*(3), 313- 325.
- Chao, G. T. (1997). Mentoring Phrases and Outcomes. *Journal of Vocational Behavior, 51*, 15-28.
- Dareh, J., & Playko, M. (1990). Mentor programs focus on the beginning principal. *NASSP Bulletin, 74* (September), 73-77.
- Dornyei, Z. (2003). *Questionnaires in second language research: Construction, administration, and processing*. New Jersey: Lawrence Erlbaum Associates.
- Drew, S., & Smith, P. (1995). The learning organization: change proofing' and strategy. *The Learning Organization, 2*(1), 4-14.
- Evertson, C. M., & Smithey, M. W. (2000). Mentoring effects on protégés' classroom practice: An experimental field study. *The Journal of Educational Research, 93*(5), 294-304.
- Feiman-Nemser, S., & Parker, M. B. (1992). *Los Angeles mentor: Local guides or educational companions*. MI: National Center for Research on Teacher.
- Galbraith, M. W. & Cohen, N. H. (1995). *Mentoring: New strategies and challenges*. San Francisco: Jossey-Bass Publishers.
- Gerstein, M. (1985). Mentoring: An age old practice in a knowledge-based society. *Journal of Counselling and Development, 64*(Oct), 156-157.
- Kay, D., & Hinds, R. (2002). *A practical guide to mentoring*. Oxford: Howtobook.
- Klausmeier, R. L. (1994). Responsibilities and strategies of successful mentors. *The Clearing House, 68*(1), 27-29.
- Kohler, F. W., Crilley, K. M., Shearer, D. D., & Good, G. (1997). Effects of peer coaching on teacher and student outcomes. *The Journal of Educational Research, 90*(4), 240.

- Johnsrud, L. (1991). Mentoring between academic women: the capacity for interdependence. *Initiative, 54*, 7-17.
- Johnson, H. (1997). *Mentoring for exceptional performance*. CA: Griffin Publishing.
- Lacey, K. (1999). *Making mentoring happen: A simple and practical guide to implementing a successful mentoring program*. Australia: Tim Edwards.
- Long, J. (1997). The dark side of mentoring. *Australian Educational Research, 24*, 115-123.
- Madison, J., Knight, B., & Watson, K. (1993). Mentoring amongst academics in Australia: a case study. *Australian Educational Researcher, 20*(1), 77-91.
- Manson, N. (1990). Great Lakes Colleges Association. Program foster faculty mentoring. *Liberal Education, 76*(May), 34-36.
- Blank, M. A., & Sindelar, N. (1992). Mentoring as professional development: From theory to practice. *The Clearing House, 66*(1), 22-27.
- Meyer, T. (2002). Novice teacher learning communities: An alternative to one-on-one mentoring. *American Secondary Education, 31*(1), 27-43.
- Noe, R. (1988). An investigation of the determinants of successful assigned mentoring relationships. *Personnel Psychology, 41*, 457-471.
- Oppenheim, A. N. (1992). *Questionnaire design, interviews, and attitude measurement*. London: Printer Publishers.
- Peterson, B. E., & Williams, S. R. (1998). Mentoring beginning teachers. *The Mathematics Teacher, 91*(8), 730-.
- Portner, H.(1998). *Mentoring new teachers*. CA: Corwin Press.
- Redmond, S. (1990). Mentoring and cultural diversity in academic settings. *American Behavioral Scientist, 34*(2), 188-200.
- Reiman, A., & Thies-Spinthall, L. (1998). *Mentoring and supervision for teacher development*. New York: Longman.
- Saban, B. (2002). Mentored teaching as (more than) a powerful means of recruiting newcomers. *Education, 122*(4), 828- 840.
- Sinclair, C. (2003). Mentoring online about mentoring possibilities and practices. *Mentoring and Tutoring, 11*(1), 79-94.
- Smith, T. M. & Ingersol, R. M. (2004). What are the effect of induction and mentoring on beginning teachers' turnover? *American Educational Research Journal, 41*(3), 681-714.
- Sullivan, C. G. (1992). *How to mentor in the midst of change*. Alexandria: Association for Supervision and Curriculum Development.

- Teare, R., & Dealtry, R. (1998). Building and sustaining a learning organization. *The Learning Organization*, 5(1), 47.
- Tellez, K. (1992). Mentor by choice, not design: Help-seeking by beginning teachers. *Journal of Education*, 43(3), 214-221.
- Troutman, Y. R. H. (2002). *Effectiveness of teacher mentoring as perceived by proteges*. Unpublished Ph.D., University of Southern Mississippi, Mississippi.
- Wang, J., & Odell, S. J. (2002). Mentored learning to teach according to standards-based reform: A critical review. *Review of Educational Research*, 72(3), 481-546.
- Walker, A., & Stott, K. (1994). Mentoring programs for aspiring principals: Getting a solid start. *NASSP Bulletin*, 78, 72-77.
- Woullard, R., & Coats, L. T. (2004). The community college role in preparing future teachers: the impact of mentoring program for pre-service teachers. *Community College Journal of Research and Practice*, 28, 609-624.
- Yost, R. (2002). "I think I can": Mentoring as a means of enhancing teaching efficacy. *The Clearing House*, 75(4), 195-197.
- Zachary, L. J. (2000). *The mentors' guide*. San Francisco: Jossey-Bass.



## **Is Syntactic Maturity a Reliable Measurement to Investigate the Relationship Between English Speaking and Writing?**

**Zhu Xinhua**

*The North University for Ethnicity, China*

### **Bio Data:**

Zhu Xinhua obtained two master's degrees in English, one of which was in China with a focus in English language and culture and the other in TESL/Linguistics at Oklahoma State University, USA. To date, she has worked in three universities in China. Currently, she works in the English department of the North University for Ethnicity of China with her fields of interest being EFL teaching and second language acquisition.

### **Abstract:**

Numerous studies have been carried out to study the relationship between English speaking and writing. Researchers have adopted a great many measurements to investigate this relationship. However, the reliability of these measurements has been questioned. So, this study is especially devoted to examining whether syntactic maturity can be used as a reliable measurement to investigate the relationship between English speaking and writing. Specifically, the researchers have mostly attended to the correlation and differences between speaking and writing. In this study, selected 40 randomly selected college-level ESL students who studied in one American university participated in the study. The subjects' written and spoken samples were collected for the analysis in this study. The subjects' written samples were their diagnostic essays written for a college-level ESL composition course at one American university. Their spoken samples were gathered through their participation in a semi-direct, tape-mediated oral proficiency test, the Video Oral Communication Instrument (VOCI). The measures of syntactic maturity were used to examine the lexical development of speaking and writing of a group of 10 students [out of the 40 students] enrolled in a college-level ESL composition course. The rank-order of the subjects' spoken and written data showed that the same subjects belonged to the high and low-rated subjects in terms of both their speaking and writing performance. Most importantly, the study revealed that the measures of syntactic maturity can differentiate between proficiency levels but cannot indicate the differences between two modes.

**Keywords:** syntactic maturity, English as second language, Video Oral Communication Instrument (VOCI), relationship between speaking and writing

### **1. Introduction**

Numerous studies have been conducted on the relationship between speaking and writing development in L1 acquisition from different perspectives (e.g. Harrell 1957; Gleason 1965). Cramer (1978) stressed that speaking influenced writing positively since written language derives from oral language and stated that the teaching of reading and writing could maximize the connections between spoken and written language. Kroll (1981) stated there was a developmental trend on speaking and writing, which progressed through

different phases of development. These studies have provided valuable information for L1 instructors to facilitate their students' language acquisition.

However, fewer studies have been written about this issue on L2 acquisition. L2 instructors therefore have to turn to L1 acquisition studies for information on the relationship between speaking and writing for their teaching strategies. Nevertheless, Vann (1981) states that many questions need to be answered concerning this relationship. For example, "How much transfer of learning can we expect from one language skill area to another?" "Will fluent speakers make good writers?" "Why do some students have particular difficulty with one mode or another? (p.167)" Similarly, Kim (2000) points out the stages in L1 and L2 acquisition might not be the same. Thus, much research is needed to explore this relationship so as to enrich L2 acquisition theories and provide a guide for L2 instructors when they practice their teaching in classrooms. Since 1920s, researchers have undergone this kind of research with different measurements. To illustrate, some studies have adopted the measurement of syntactic maturity (Hunt, 1965; O'Donnell, et al. 1967; Price & Graves, 1980) while others have used that of Flesch Reading Ease, Flesch Human Interest and type-token ratios (Gibson et al, 1966). Still others have preferred the measurement of lexical density (Chafe & Tannen, 1987; Beaman, 1984; Eggin, 1994 et al). However, what measurement should researchers choose to adopt remains very essential in conducting research on the relationship between English speaking and writing. This study, therefore, has especially attended to the use of syntactic maturity in the examination of the relationship between English speaking and writing. As shown in the literature review on the relationship between English speaking and writing, most studies of this kind have fell into two categories: those on the correlation between two modes and those on the difference between two modes.

### **1.1 Measuring Speaking and Writing Development with T-unit Analysis**

Since Hunt (1965) and O'Donnell (1970) developed objective measures of syntactic maturity, they have been successfully used in both first and second language research. Syntactic maturity is defined as the ability to manipulate the syntax of the language, which is positively related with students' age (Hunt 1965). Syntactic maturity can be measured by objective measures such as mean T-unit length (MTUL), mean error-free T-unit length (MEFTUL) and other indices. In terms of Hunt, T-unit is "minimal terminal syntactic maturity" i.e., any main clause with all of its modifiers (Nystrand 1982). Since T-unit can be objectively identifiable in both speech and writing, it has been used to measure the

relationship between writing and speaking development in both L1 and L2 research.

### **1.1.1. L1 Research**

A few studies have used T-units to explore the relationship between speaking and writing in L1 acquisition. Cayer and Sacks (1979) studied both the spoken and written samples of eight basic writers with T-unit count procedures in an effort to describe some of the similarities and differences between their oral and written discourse. They found that “some analogues and parallels do appear to exist between the two modes of discourse, particularly when the writer has not yet fully achieved mastery of skills needed to generate written language effectively” (p.126). O’Donnell and his colleagues’ study revealed that the average length of T-units was significantly greater in the speech than in the writing of third graders, but not of fifth and seventh graders. As students got older, written units became longer and more complex than spoken ones (O’Donnell, Griffin, & Norris 1967). O’Donnell (1974) divided both the written and spoken sample produced by an author and editor into T-units to test the value of a method of analysis and the hypotheses about the syntactic differences between the speech and writing. His findings showed that the median length of the written T-units is greater than that of the spoken T-units and there are more T-units with dependent clauses in writing than in speech. Most important of all, this study demonstrated that T-unit can be objectively identified and is useful as a unit for syntactic analysis.

### **1.1.2. L2 Research**

T-unit analysis has been used widely in L2 research on either speaking or writing development (Cooper 1976; Gaies 1979; Kameen 1983; Halleck 1995). But a few studies shed light on the relationship between speaking and writing in L2 development, let alone the use of T-unit analysis. As far as the literature I’ve reviewed, Vann (1979) examined paired oral and written discourse of a group of 28 adult native speakers of Arabic studying English in the United States with the methodology based on previous research using native speakers. She found that “Oral compositions were almost twice as long as written ones (p.9)” and Both mean T-unit length and mean error-free T-unit length were longer in written than in oral discourse. Kim (1996, 1998 & 2000) conducted a study to compare speaking and writing development in a L2 college student over a two-year period. In this study, she analyzed errors and syntactic maturity in the subject’s writing development and his formal language institute oral interviews to determine whether the L2 college student

followed the same pattern of writing and speaking development reported for L1 acquisition. Her findings indicate that errors decrease and syntactic maturity increases in the same way that it occurs in L1 development. Additionally, her study indicates that writing is of greater syntactic maturity than speaking in an L2, which agrees with the research in an L1 (e.g. Harrell 1957). The review of literature on syntactic maturity indicates that most L2 research using the T-unit analysis has focused on writing rather than speaking. For instance, twenty-one studies reviewed by Ortega (2003) are also concerned with writing. Very few studies have paid attention to speaking in L2, which should be an important part in L2 development. The earliest study that examined the spoken data with syntactic maturity measures was conducted by Thornhill (1970), who discovered that remarkable parallels existed between the trends in the language behavior in both L1 acquisition and L2 learning. Halleck (1995) compared holistic ratings of ACTFL Oral Proficiency Interviews (OPIs) and objective measures of syntactic maturity to investigate the syntactic basis of ACTFL's holistic ratings and examine the relationship between oral proficiency and syntactic maturity. Her findings demonstrate that levels of syntactic maturity vary according to interview tasks and objective measures of syntactic maturity overall correlate with holistic evaluation. In other words, her study provides some empirical evidence that measures of syntactic maturity can be used to demonstrate second and foreign language learners' oral proficiency. Therefore, further studies should attend to the use of syntactic maturity measures in L2 speaking in order to get an overall picture of L2 development.

## **2. Purpose**

The purpose of the present study is to examine the reliability of syntactic maturity as the measurement of the relationship between speaking and writing. Specifically, this study used the measures of syntactic maturity to investigate the syntactic development of speaking and writing. I was interested in finding answers to the following research questions. Is syntactic maturity a reliable measurement to examine the relationship between English speaking and writing?

## **3. Materials and methods**

### **3.1. Subjects**

The 40 students who participated in this study were enrolled in a college-level ESL composition course, *International Freshman Composition I* at Oklahoma State University,

majoring in public administration, international business, communication sciences and disorders, electrical engineering, architecture, computer science and accounting. These subjects came from different countries such as Korea, Japan, Uzbekistan, Taiwan, Nigeria, Indonesia, Myanmar, Nepal and Bangladesh. The length of their stay in the United States ranged from 5 days to 1.5 years with a mean of 4.3 months. The average number of years spent studying English was 8.1 years with a range of 2 to 15 years.

### **3. 2. Instrument**

The instrument, the Video Oral Communication Instrument (VOCI) for ESL/EFL was used to assess the oral proficiency of the subjects. The measures of syntactic maturity were employed to examine the syntactic development of both the subjects' oral and written language.

#### **3.2.1. VOCI**

VOCI is a semi-direct and tape-mediated speaking test, which is used as an alternative for the OPI (Oral Proficiency Interview) to determine level of oral proficiency. This study particularly used the English version, the VOCI for ESL/EFL, which was developed at San Diego State University's Language Acquisition Resource Center (LARC) by Halleck and Young (1995). Like other semi-direct speaking tests such as Test of Spoken English, the VOCI uses video stimuli to elicit samples of oral performance from the subjects. In the VOCI test, the participants create a context or situation and then ask the test-taker a question related to the situation or the context. The test-taker then has to respond to the question accordingly and the response is recorded with a manually operated audio recorder. Specifically, the test-taker watches the audio-visual stimulus and pauses the VCR with a remote control to respond to the question asked by the participant. After the test-taker finishes answering the question, he or she restarts the video and watches the next video. The VOCI takes two forms: one with time constraints (a timed version) and the other without time constraints (an untimed version). The present study used the untimed version.

The VOCI for ESL/EFL used in this study consists of a total of 23 questions (Please see Appendix A for a complete transcript of the VOCI questions), which assess the four proficiency levels: novice, intermediate, advanced, and superior levels defined by the ACTFL (American Council on the Teaching of Foreign Languages) guidelines (1986). The speech tasks elicited by the VOCI vary from describing, comparing and contrasting to

supporting an opinion and hypothesizing. Table 1 presents examples of such questions and the level of difficulty represented by each question.

Table 1: Examples of questions on the VOICI

Question Level	Example
Novice	What's your name? (Q1) Where are you from? (Q2)
Intermediate	Tell us about your hometown. (Q3) Describe one of your friends. (Q5)
Advanced	Compare your hometown with a city you visited or know well. (Q7) Describe an unforgettable experience. (Q9)
Superior	If you were a teacher and you discovered one of your students had cheated on a test by copying from another student's paper, what would you do? (Q 17) What do you think about the portrayal of violence and crime on TV? (Q21)

(Adapted from McCrackin, 1998)

### 3.2.2. Measures of syntactic maturity

The measure of syntactic maturity was used to investigate the relationship between speaking and writing in college-level ESL. Seven measures of syntactic maturity were used in this study to examine the syntactic development of both the subjects' speaking and writing. Specifically, these measures are total number of words (TNW), mean T-Unit length (MTUL), mean error-free T-Unit length (MEFTUL), and percentage of error-free T-unit (%EFTU), subordination ratio (SR), mean clause length (MCL) and dependent clauses per clause (DC/C). Of these measures, TNW, MTUL and MCL tap length of production at the clausal and phrasal level. MEFTL and %EFTU measure the accuracy of both spoken and written performance. SR and DC/C gauge the amount of subordination in spoken and written data. As far as the calculation of these indices were concerned, the total number of words included in each language sample, and the total number of T-units were counted, as well as the total number of dependent clauses and the total number of error-free T-units. On this basis, six indices were then determined through the following formulas:

MTUL=number of words/number of T-units

MCL = number of words/number of T-units+ number of dependent clauses

MEFTL=number of words/number of error-free T-units

%EFTU=number of error-free T-units/number of T-units

SR=number of T-units+number of dependent clauses/number of T-units

DC/C=number of dependent clauses/ number of T-units + number of dependent clauses

### **3.3. Data collection**

The written samples were collected from all the subjects and the spoken samples were gathered from those subjects who participated in the VOCI test, which will be explained later in the article. The written samples were taken from the first draft of the diagnostic essay the subjects wrote for the course of *International Composition I*. The essay prompt is to describe three underlying rules shaping student behavior in American classrooms which they had discovered since they came to Oklahoma State University. Compared to other essays written for this course, the diagnostic essay is written in class, so it has less planning time than other essays that are written after class by the students. Therefore, the use of the diagnostic essays increases the comparability between the spoken and written data.

#### **3.3.1. Written sample**

40 essays were originally collected from the second written assignment of the class. Then, according to the rating of the instructor, 20 students took part in the study: 10 high-rated and 10 low-rated students. After the written samples were collected, three raters graded the papers holistically on a scale of 100. A higher interrater reliability was found between Raters 1 and 2 ( $r=0.91$ ;  $p=0.0002$ ) than between Rater 3 and the other two raters, respectively. As Davies, et al. (1999) suggests, a correlation coefficient of more than 0.8 indicates a good interrater reliability. So, the rating between Raters 1 and 2 was used as the evaluation of the written samples. The ratings of these two raters were averaged, which were then transferred to the rank-order of the written samples from 1 to 10. Accordingly, 5 high-rated and 5 low-rated students were finally chosen to take the spoken test in this study in order to obtain the corresponding spoken sample. The average of 5 high-rated students' written score was 90.1 and that of the 5 low-rated students' written score was 77.1. So, there were 13 points apart between the written samples of 5 high-rated and 5 low-rated students.

#### **3.3.2. Spoken Sample**

After the written data of these 10 students were collected, the spoken data were gathered through the students' participation in the VOCI. The subjects took the VOCI in a testing room alone in order for them to feel less nervous. As noted earlier, the VOCI is a semi-direct, tape-mediated oral proficiency test, which lasts approximately 30 minutes and consists of a total of 23 questions (Please see Appendix A for the VOCI questions). Of

these 23 questions, the subjects' answers to 17 questions were used for analysis mainly according to the familiarity of the topics of these questions. These 17 questions are questions 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 15, 16, 17, 18, 21, 22 and 23. In a specific manner, questions 1 and 2 function as the practice questions in order for the students to familiarize themselves with the use of the machines. Questions 13, 14, 19 and 20 were not used for data analysis because the researcher felt that the subjects might not be familiar with the topics such as "lasting peace" (Q 13), "abstract painting" (Q 14), "free trade" (Q 19), "televising trials" (Q 20). After the collection of the spoken samples, they were rated and rank-ordered from 1 to 10 by a certified OPI tester according to the ACTFL Guidelines (1986). After the spoken samples were gathered, they were transcribed. In regard to the spoken sample, each subject's answers to all the above-mentioned 17 VOICI questions were taken as a whole to be analyzed in comparison with the subjects' written samples.

### **3.4. Research Hypotheses**

As mentioned earlier, this relationship is mostly concerned with two aspects: one is about the differences between speaking and writing and the other about the correlation. Consequently, this study was conducted according to the following hypotheses. Hypotheses one and two investigate whether students who write well speak well and vice versa, which is about the correlation between two modes. Hypotheses three to five examine the differences between two modes.

Specifically, hypothesis one concerns the rank-order of the subjects' spoken and written data. The subjects were categorized into high-and low-rated subjects depending on the rank-order of their spoken and written data. If hypothesis one is sustained, there will be five high-rated subjects and five low-rated ones in regards to both samples. Specifically speaking, five high-rated students will be ranked highly in both samples and five low-rated students will be placed in the low rank area in both samples. Based on hypothesis one, hypotheses two to five are raised in relation to the six indices of syntactic maturity. These hypotheses are as follows:

***Hypothesis 1:*** Students who were ranked highly in the spoken samples will be

those who were ranked highly in the written samples and vice versa.

***Hypothesis 2:*** The written samples will positively correlate with the spoken samples in relation to the measures of syntactic maturity for all subjects.

***Hypothesis 3:*** The written samples will contain longer T-units and longer clauses than spoken samples.

**Hypothesis 4:** The written samples will have longer error-free T-units and more error-free T-units than the spoken samples.

**Hypothesis 5:** The written samples will contain more dependent clauses than the spoken samples.

### **3.5. Data Analysis**

Two procedures were carried out to analyze statistical data in regard to the purpose of this study. First, the objective measures of syntactic maturity were marked and tallied in both the spoken and written samples. Second, the data were analyzed using version 3.03 of the statistical software, GraphPad Prism. Pearson product-moment correlation was carried out to demonstrate how each measure in the spoken and written data correlates with each other so as to investigate the relationship between speaking and writing. Finally, the level of significance was computed with version 8 of SAS software to examine whether the results obtained were statistically significant or not. Results were considered significant at the  $p < .05$  level.

## **4. Results and Discussion**

Given the purpose of the study, this section first demonstrates the rank-order of the subjects in terms of their speaking and writing performance. And then, results of the measures of syntactic maturity are analyzed to further explore this relationship.

### **4.1. The Rank-Order of the Subjects**

Table 2 provides an overview of how the subjects' speaking and writing related to each other in terms of the rank-order of the subjects in accordance to their speaking and writing performance. Both the spoken and written samples were ranked from 1 to 10. In order to protect the confidentiality of the subjects, pseudonyms were used to identify each subject. With regard to the spoken samples, Melody, Andrea and Polina were rated as No 1, since these subjects had the highest level of oral proficiency. Similarly, John and Kala were both ranked as No 4 and Anna and Tom were rated No 8. As revealed in the table, Melody, Andrea, Polina, John and Kala were rated highly in both spoken and written samples as compared to Anna, Emily, Tom, Jack and Mike, who were ranked in both

Table 2 - The rank order of subjects (identified by pseudonyms) in both spoken and written samples (N=10)

Ranking	Spoken Sample	Written Sample
High-rated Ones		
1	Melody, Andrea, Polina	Melody
2		Andrea
3		Polina
4	John, Kala	Kala
5		John
Low-rated Ones		
6	Emily	Jack
7	Jack	Tom
8	Anna, Tom	Emily
9		Anna
10	Mike	Mike

samples as the low-level group. In other words, Melody, Andrea, Polina, John and Kala belonged to the high-rated group while Anna, Emily, Tom, Jack and Mike were in the low-rated group in both samples. In general, this finding indicates that subjects who write well also speak well and vice versa.

Concerning the spoken samples, the difference is related to the holistic nature of the oral proficiency rating scales. With a close look at the table, one could find that Melody, Andrea and Polina were rated No1 in the spoken sample as compared to No 1, 2 and 3, respectively in the written sample. So, they were almost ranked the same in both the spoken and written samples. The same observation could also apply to John and Kala, who were ranked No 4 in the spoken sample and No 4 and 5 in the written sample. As far as the low-rated subjects were concerned, Mike was ranked No 10 in both samples. Thus, the rank-order of the subjects in light of their speaking and writing performance indicated that there seemed to exist a positive relationship between the subjects' speaking and writing. Stated differently, if students are found to write well, they are judged to speak well too and vice versa.

#### **4.2. The Correlation between Speaking and Writing regarding High-and Low-rated Subjects with Measures of Syntactic Maturity**

As mentioned earlier, the measures of syntactic maturity were used to investigate the syntactic development of the subjects' speaking and writing. The analysis of this relationship with measures of syntactic maturity was particularly concerned with the correlation between spoken and written sample in terms of the measures of syntactic

maturity. If such correlation can be found between two modes, it may be concluded that the one who speaks well also writes well and vice versa.

Table 3 reflects this correlation with measures of syntactic maturity. In Table 3, a Pearson product-moment correlation coefficient was computed to show the relationship between the spoken and written data. As far as this correlation is concerned, TNW was not considered because the topics of the written and spoken data were different from each other. Besides, there were 17 topics in the spoken data as opposed to 1 topic in the written data. So, the spoken and written data were not of comparability in terms of this index. Thus, all the other measures except TNW were calculated to exhibit this correlation. Of all the measures, MTUL showed the strongest correlation with a correlation coefficient  $r$  being 0.85 followed by DC/C ( $r=0.75$ ), %EFTU ( $r=0.72$ ), SR ( $r=0.71$ ), MCL ( $r=0.67$ ) and MEFTUL ( $r=0.67$ ).

As indicated in this table, the spoken and written data were found to positively and significantly correlate with each other in light of all the measures. So, it supports the assumption made above that the one with good speaking skills has good writing skills and vice versa. When examined closely, it can be noted that MTUL demonstrates the strongest correlation between the spoken and written data followed by DC/C, %EFTU,

Table 3 - The correlation between the spoken and written data with regard to the measures of syntactic maturity for all subjects (\*significant at  $p < .05$ )

Measures	Correlation Coefficient (r)
MTUL	0.85*
MEFTU	0.67*
%EFTU	0.72*
SR	0.71*
MCL	0.67*
DC/C	0.75*

SR, MEFTUL and MCL, respectively. Therefore, it can be concluded from Table 3 that the students with good speaking skills have good writing skills and vice versa. Besides, MTUL is the most sensitive index to exhibit the relationship between speaking and writing with respect to high-and low-rated students.

To sum up, significant correlation between the two samples was found in regard to all the measures of syntactic maturity. Therefore, it can be concluded that measures of syntactic maturity are good indicators of levels of proficiency.

### 4.3. The differences between Speaking and Writing

Table 4 presents the means of all measures of syntactic maturity of spoken and written data. From this table, it can be observed that all the measures except %EFTU were higher in the written data than in the spoken data. Specifically, the average of MTUL in the written samples was 14.87 words per T-unit as opposed to 11.11 words per T-unit in the spoken samples. Similarly, the written samples showed a higher index with the mean of MCL being 8.10 words per clause than the spoken samples with that of MCL being 7.11 words per clause. Regarding MEFTUL, the average of it in the written

**Table 4 - Average of All Measures of Syntactic Maturity of Spoken and Written Data**  
(\*significant at  $p < .05$ )

Measures	Spoken	Written
<b>*MTUL</b>	11.11	14.87
<b>*MCL</b>	7.11	8.10
MEFTU	14.21	23.89
%EFTU	0.78	0.65
SR	1.54	1.81
DC/C	0.34	0.38

samples (mean=23.89) was also higher than that of it in the spoken samples (mean=14.21). Likewise, the mean values of SR and DC/C exhibited a higher index in the written samples (mean for SR =1.81; mean for DC/C=0.38) than in the spoken samples (mean for SR =1.54; mean for DC/C=0.34). However, given the level of significance, a significance difference was found only with the measures of MTUL and MEFTUL. Nevertheless, unlike other indices, %EFTU indicated a lower index in the written sample (mean=0.65) than in the spoken sample (0.78). Besides, this index approached the significant difference with  $p=0.06$ .

According to the above results, in general, Table 4 demonstrated that the written samples contained longer T-units, longer clauses and longer error-free T-units and incorporated more dependent clauses than the spoken samples. However, since only MTUL and MEFTUL showed a significant difference between the spoken and written data, it may be said these two indices can function as good indicators of the differences between these two modes.

However, one thing worth noting is that %EFTU showed that the subjects' spoken discourse included more error-free T-units than their written discourse. This finding is very unexpected, because it might be assumed that written discourse should contain more error-free T-units than spoken discourse since more planning time is available in writing than in speaking. However, a close look at %EFTU of each subject's spoken and written data demonstrated that all the subjects produced more error-free T-units in speaking than writing. Besides, this index approached significance. So, this finding would probably not be obtained by chance. With a larger sample, it may be found that this index would indicate a significant difference between these two modes of communication. With a close examination of the spoken transcripts of the subjects, it was found that the subjects spoke in short sentences, which made it unlikely for them to make errors. However, when they wrote, they tended to write longer sentences with more dependent clauses. In this way, it is very likely for them to make errors in writing. As shown in the table, SR and DC/C exhibited that the subjects used more dependent clauses in writing than in speaking. To illustrate, approximately 100 words from the same subject's spoken and written data were taken to demonstrate this difference between the two modes.

*The Written Sample:*

In my country, Indonesia, there are many culture characteristics *which are respected. Because Indonesia is an Asian country and we always call Asian countries "East Countries".* So, there are many "rules" in our country.\*/ There are three culture characteristic *which are respected in my country* \*/and those are always giving something with your right hand, don't ever call a person *who is older than you only with his or her name* and always walk on the left side./ Giving or receiving something with your left hand is unpolite.\*/ We are always taught with our parents *since we were a child* to give and receive with right hand./

As observed from the written sample, it consists of 6 T-units (between the virgules), 5 dependent clauses (italicized) and 3 error-free T-units (with an asterisk). According to the guideline for T-units, clauses, word counts and errors provided by Polio (1997), if a subordinate clause is standing alone, attach it to the preceding or following sentence and count as a 1 T-unit with an error. So, the first T-unit (*In my country, Indonesia, there are many culture characteristics which are respected. Because Indonesia is an Asian country and we always call Asian countries "East Countries".*) is counted as a T-unit with an error. As for the third T-unit (*There are three culture characteristic which are respected in my country*) and the fifth T-unit (*Giving or receiving something with your left hand is*

*unpolite.*), the spelling mistakes are not taken into account. So, these two T-units are error-free. In comparison to the written sample, the spoken sample contains 8 T-units, 3 dependent clauses and 6 error-free T-units.

*The Spoken Sample*

Ten years is a very very long time.\* / I mean, ok honestly, I have nothing in my mind *what I'm gonna be next ten years, because it's a lot of time.*\* / But, my goal in the next four years absolutely is *I want to graduate from here as soon as possible.*\* / I want to graduate from universities.\* / I don't know, I want to graduate from here or another university./ But I want to graduate as soon as possible.\* / And I want to get great job in the future./ And I also plan to get my master's degree.\* /

To illustrate and explain the reason why %EFTU is generally higher in the spoken than in the written samples, Table 5 shows the MTUL, MCL, MEFTUL, %EFTU, SR and DC/C in both samples. As noted from Table 5, the %EFTU was obviously higher in the spoken data (%EFTU=0.75) than in the written data (%EFTU=0.50). Concerning other indices, all except MCL showed a lower index in the spoken sample than in the written sample. As observed from the table, the subject produced fewer words per T-unit in the spoken (MTUL =12.50) than the written sample (MTUL=16.70). Regarding MEFTUL, the spoken sample demonstrated a lower index (MEFTUL=16.67) than the written sample (MEFTUL=33.33). Besides, the subject produced fewer dependent clause per T-unit in the spoken (SR=1.38) than in the written sample (SR=1.83). Similarly, DC/C indicated a lower index in the spoken sample (DC/C=0.27) than in the written sample (DC/C=0.45).

**Table 5: Measures of Syntactic Maturity of Selected Spoken and Written Data from the Same Subject (TNW=100 words)**

<b>Measures</b>	Spoken	Written
<u>MTUL</u>	12.50	16.70
MCL	9	9
MEFTU	16.67	33.33
%EFTU	0.75	0.50
SR	1.38	1.83
DC/C	0.27	0.45

In a discussion of the above results, the subject produced fewer words per T-unit and

per error-free T-unit, fewer dependent clauses per T-unit and per clause in the spoken than in the written samples. So, when the subject speaks, he speaks in shorter sentences than he writes. Besides, he does not incorporate as many dependent clauses in a T-unit in speaking as in writing. Therefore, it's less likely for him to make errors when he speaks. On the other hand, since his writing include many long sentences and dependent clauses, he is very likely to make errors in writing. That's why his spoken sample shows more error-free T-units than his written sample.

## **5. Conclusion**

The conclusion of the study is presented by a discussion of outcomes of the hypotheses and implications of the results. Hypothesis one deals with the rank-order of the subjects in both the spoken and written samples. The results of the present study show that the same subjects were ranked as high- and low-rated groups in both spoken and written samples. That is to say, the subjects who were ranked highly in the spoken samples were also ranked highly in the written samples and vice versa. Therefore, this hypothesis is proved. Based on hypothesis 1, hypothesis 2 investigates whether the spoken and written samples for all subjects correlate with each other. They take into account all subjects' written and spoken performance. The results of the study demonstrated a positive and significant correlation between spoken and written data with respect to all the measures of syntactic maturity. This meant that students with good speaking skills had good writing skills and vice versa. So, this hypothesis is sustained.

The findings of the present study seemed to indicate that there was a positive relationship between speaking and writing according to the rank-order of the subjects and the analysis of syntactic maturity of both the spoken and written samples. However, this implication cannot be generalized, because other aspects of this relationship have to be considered to get an overall picture of this relationship. In addition, with regard to the measures of syntactic maturity, the present study found that the measures of syntactic maturity could be used to investigate the correlation between speaking and writing.

Concerning hypothesis three, the study showed that the written sample did contain longer T-units and clauses than spoken samples. In addition, MTUL was found statistically significant. Therefore, it can be said that this hypothesis is sustained with respect to MTUL but not to MCL. With reference to hypothesis four, as noted from the findings of the study, more error-free T-units appeared in the spoken sample than in the written sample regarding %EFTU in both samples. This finding was unexpected. However,

this index was not found statistically significant. With respect to MEFTUL, the written sample contained more words per an error-free T-unit than the spoken sample and the difference between the two modes was statistically significant. So, this hypothesis is proved concerning MEFTUL but not %EFTU. Regarding hypothesis five, as noted earlier, this hypothesis was concerned with SR and DC/C. The finding of the study demonstrated that both indices were higher in the written sample than in the spoken sample. However, neither of these two indices was found to reach the significant level. Thus, this hypothesis is rejected. Overall, the present study seems to show that syntactic maturity could not indicate the differences between the two modes.

In conclusion, according to the findings of this study, the measures of syntactic maturity can be adopted to study the correlation between speaking and writing. However, they cannot be used to distinguish between speaking and writing. This finding should be of significance to other researchers in this area.

## **References**

- Beaman, K. (1984). Coordination and subordination revised: Syntactic complexity in spoken and written narrative discourse. In D. Tannen (Ed.), *Coherence in spoken and written discourse*. Norwood, NJ: Ablex.
- Cambourne, B. L.(1981). Oral and written relationship: A reading perspective. In B.M. Kroll & R.J. Vann (Eds). *Exploring speaking-writing relationships: Connections and contrasts* (pp.82-88). Illinois: National Council of Teachers of English.
- Cayer, R. L. & Sacks, R.K. (1979). Oral and written discourse of basic writers: similarities and differences. *Research in the Teaching of English*, 13(2), 121-128.
- Chafe, W. & Tannen, D. (1987). The relation between written and spoken language. *Annual Review of Anthropology*, 16, 383-407.
- Cooper, T. (1976). Measuring written syntactic patterns of second language learners of German. *Journal of Educational Research* 69,176-183.
- Cramer, R. L. (1978). *Writing, reading, and language Growth: An introduction to language arts*. Columbus, Ohio: Charles E. Merrill.
- Egins, S. (1994). An introduction to systemic functional linguistics. London: Printer Publishers.
- Gaies, S. J. (1976, May). *Sentence-combining: A technique for assessing proficiency in a second language*. Paper read at the conference on perspectives on language, University of Louisville, Louisville, Kentucky, USA.

- Gaies, S. J. (1980). T-unit analysis in second language research: Applications, problems and limitations. *TESOL Quarterly*, 14, 53-60.
- Gibson, J.W., Gruner, C.R., Kibler, R.J. & Kelly, F.J. (1966). A quantitative examination of differences and similarities in written and spoken messages. *Speech Monograph*, 33, 444-451.
- Gleason, H.A., Jr. (1965). *Linguistics and English grammar*. New York: Holt.
- Halleck, G. B.(1995). Assessing oral proficiency: A comparison of holistic and objective measures. *The Modern Language Journal*, 79(2), 223-234.
- Halleck, G. B. & Young, R. F. (1995). *Video oral communication instrument (VOCI) for ESL/EFL*. San Diego, CA: Language Acquisition Research Center.
- Harrell, L. (1957). A comparison of oral and written language in school-age children. *Monographs of the Society for Research in Child Development*, 22, no 3, serial no. 66. Lafayette, Ind.: Child Development Pubs.
- Hunt, K. W. (1965a). Grammatical structures written at three grades levels. *Research Report No 3*. Urbana, Illinois: NCTE.
- Hunt, K. W.(1965b). Sentence structures used by superior students in grades four and twelve and superior adults. *Cooperative Research Projects, No.5-0312*. Tallahassee, Florida: Florida State University. Available from: ERIC database (ERIC Document Reproduction Service No. ED 010 047). [ Accessed 16 January 2004].
- Kameen, P. (1983). Syntactic skill and ESL writing quality. In A. Freedman, I. Pringle, & J. Yalden (Eds.), *Learning to write: First language/second language* (pp. 139-158). London: Longman.
- Kantor, K.J. & Rubin, D.L. (1981). Between speaking and writing: Processes of differentiation. In B.M. Kroll & R.J. Vann (Eds.), *Exploring speaking-writing relationships: Connections and contrasts* (pp. 55-81). Illinois: National Council of Teachers of English.
- Kim, A. (1996). A comparison of oral and writing development in a second language college student. Available from ERIC database (ERIC Document Reproduction Service No. ED169 790). [Accessed 13March 2003].
- Kim, A. (1998). A case study of error analysis/syntactic maturity as indicators of second language writing development. *Research and Teaching in Developmental Education* 1(1), 79-88.
- Kim, A. (2000). A comparison of oral & writing development in a second language college student. *Research and Teaching in Developmental Education*, 16(2). 73-82.

- Kroll, B. (1981). Developmental relationships between speaking and writing. In B. M. Kroll. & R. J. Vann (Eds.), *Exploring speaking-writing relationships: Connections and contrasts* (pp.32-54). Illinois: National Council of Teachers of English.
- Larsen-Freeman, D. & Strom, V. (1977). The construction of a second language acquisition index of development. *Language Learning*, 27, 124-134.
- McCrackin, K. (1998). *The effects of learning strategies and motivation on the oral proficiency of ESL students*. Unpublished master thesis, Oklahoma State University, USA.
- Nystrand, M.(1982). An analysis of errors in written communication. In M. Nystrand (Ed.), *What a writer knows: The language, process and structure of written discourse* (pp.57-74). New York: Academic Press.
- O'Donnell, R.C., Griffin, W.J. & Norris, R.C. (1967). A transformational analysis of oral and written grammatical structures in the language of children in grades three, five, and seven. *Journal of Educational Research*, 61, 36-39.
- O'Donnell, R. C. (1974). Syntactic differences between speech and writing. *American Speech*, 49, 102-110.
- Ortega, L. (2003). Syntactic complexity measures and their relationship to L2 proficiency: A research synthesis of college-level L2 writing. *Applied Linguistics*, 24(4), 492-581.
- Price, G.B., & Graves, R.L. (1980). Sex differences in syntax and usage in oral and written language. *Research in the Teaching of English*, 4(2), 47-153.
- Scott, M. & Tucker, G. R. (1974). Error analysis and English-language strategies of Arab students. *Language Learning*, 24, 69-97.
- Sharma, A. (1979). Syntactic maturity: Assessing writing proficiency in a second language. Available from: ERIC database (ERIC Document Reproduction Service No. ED 185 105). [Accessed 16 January 2004]
- Thornhill, D. (1970). *A quantitative analysis of the development of syntactic fluency of four young adult Spanish speakers learning English*. Unpublished doctoral dissertation, Florida State University.
- Vann, R. J.1979. Oral and written language relationships: The question of universals. Available from: ERIC database (ERIC Document Reproduction Service No. ED 191 051). [Accessed 16 January 2004]
- Vann, R. J. (1981). Bridging the gap between oral and written communication in EFL. In B.M. Kroll & R.J. Vann (Eds.), *Exploring speaking-writing relationships: Connections and contrasts* (pp.154-167). Illinois: National Council of Teachers of English.

**Appendix: VOCI Questions**

Q1: Hello, my name is Gene and this is Ron. What's your name?

Q2: I am from New York and Ron is from Wisconsin. Where are you from?

Q3: This is a picture of my hometown. Tell us about your hometown?

Q4: Instead of writing letters, you have decided to send a cassette message to a friend back home. Describe where you are living now and what you've been doing recently.

Q5: I'm so happy my best friend just got back from vacation. I really missed him a lot. My best friend moved away and she's impossible to replace because she's so special. Describe one of your friends.

Q6: Because of a last minute problem you missed a dinner engagement with a friend. You called to apologize, but your friend is not yet home, so you need to leave a message on the answering machine apologizing for the date and explaining why you were not there.

Q7: Did you know that I went to New York last month? It sure is an interesting city. What's so special about it? The entire time I was there I tried to compare it with our city. There are lots of differences, but on the other hand, lots of things are similar. Can you compare your hometown with a city that you visited or you know well?

Q8: One thing that I didn't like about New York was that it is so big. I never really feel comfortable in big cities anymore. Really, I love city life. There's nothing more fascinating than a really big city. Not me. There are too many problems I guess. What do you think? What are the advantages or disadvantages of big city life?

Q9: Yes, that's just really unbelievable. It was a really terrific experience. There are some experiences you just can't forget. That's true. Have you ever had such an experience? An experience that you'll never forget. It can be something positive or it can be something negative. Tell us about it.

Q10: So, you finally made up your mind? Yes, and I'm really excited about it. Then you must have pretty concrete plans for the next few years? I have a good idea about what my life might be like. And you, what are your plans? What do you need to reach your goals? How might your life look ten years from now?

Q11: You have a summer job selling great books. I'm a potential customer. Convince me why I should buy the books from you.

Q12: Gene, did you read about the student who took one of these Swiss army knives to school with him in his pocket? No, what happened? Well, when he was using the scissors

part of it, his teacher caught him and she took the knife away from him and they expelled him from school. I don't get it. It looks like an innocent tool to me. Well, their school has a zero tolerance policy and they considered a Swiss army knife as a weapon. If you were the principal of this school, what would you do about this issue?

Q13: Wow, look at the headlines, another war. There have always been wars. It's nothing new. It's just human nature. Not necessarily. How do you feel about this issue? How do you think we could create a lasting peace?

Q14: I really love this painting. I don't understand it at all. Tell us why you think this is or isn't art.

Q15: My computer is broken again. Man, what a disaster. I feel so dependent on this machine. Yeah, modern technology can make life easy, but sometimes it can cause a lot of frustrations too. Discuss the positive benefits and the negative consequences of our dependence on such machines.

Q16: Some undergraduates at American universities think that native speakers of English make the most effective teachers. On the other hand, some people think the advantages of having an international teacher outweigh the disadvantages. What do you think?

Q17: If you were a teacher and you discovered one of your students had cheated on a test by copying from another student's paper, what would you do?

Q18: In many countries, higher education is for an elite group of students. Not everybody can go to the university. That certainly isn't the case in this country. Our universities are open to almost everyone regardless of their background. I can see the pros and cons of both types of educational systems. Discuss the advantages and disadvantages of both types of educational systems.

Q19: You know, I'm reading an article here on free trade in Europe and in America and it says that everybody benefits from having a free trade. No, I don't know. There's still an awful lot of opposition in a few countries to the whole issue of free trade. Take one position and defend your opinion regarding the issue of free trade.

Q20: Did you know that US laws allow trials to be televised? Yes, several high profile trials have been televised recently because of the freedom of information act. I wonder if that's such a good idea? What do you think about televising criminal trials?

Q21: Have you noticed how many shows on TV portray violent crimes? Pretty hard not to notice. Some people feel that this creates violence in our society. Yeah, but other people feel it doesn't have any effect at all on our young people. In fact, they're proud of this country's freedom of expression. What do you think about the portrayal of violence and

crime on TV?

Q22: There must be problems in your country too. What are some of the problems in your country? Suggest some solutions and discuss the implications of these solutions.

Q23: This is the last question. If you've gotten this far, you've probably taken other English tests. If so, how does this test compare to other English tests you have taken.



## **Another Look at the C-Test: A Validation Study with Iranian EFL Learners**

**Mahmood Rouhani**  
*Mashad, Iran*

### **Bio Data:**

Mahmood Rouhani holds a masters degree in ELT from the University of Isfahan, Iran. He has been teaching for seven years in high schools and private English schools in Iran. He currently teaches English in high schools in Mashad, Iran.

### **Abstract**

This study probes into the validity and discrimination power of the C-Test for the assessment of overall language proficiency. A total of 144 university students participated in this study. A Michigan Test of English Language Proficiency (MTELP) and a C-Test developed by the researcher were administered. The results indicated that the C-Test enjoyed high reliability and acceptable content relevance. Also the C-Test proved to have fairly high criterion-related validity. The extracts used in the C-Test turned out to measure, to a large extent, the same underlying trait as the MTELP –significant evidence of construct validity for the C-Test. However, the C-Test texts did not prove to behave consistently with examinees of different proficiency levels. Also it came out that the C-Test could not consistently classify the subjects in their appropriate proficiency levels. This finding was further affirmed by an ANOVA whose results demonstrated that the C-Test had difficulty discriminating between participants of lower and upper intermediate levels.

**Key terms:** language proficiency, reliability, discriminatory/discrimination power, construct validity, criterion-related validity, content validity, cloze test, C-Test.

### **1. Introduction**

Cloze test is now a well-known and widely-used integrative language test. Wilson Taylor (1953) first introduced the cloze procedure as a device for estimating the *readability* of a text. However, what brought the cloze procedure widespread popularity were the investigations with the cloze test as a measure of *ESL proficiency* (Jonz, 1976, 1990; Hinofotis, 1980; Bachman, 1982, 1985; Brown, 1983, 1993; Laesch & van Kleek 1987; Chapelle & Abraham 1990; see also Oller, 1979 for an overview). The results of the substantial volume of research on cloze test have been extremely varied. Furthermore, major technical defects have been found with the procedure. Alderson (1979, 1980, 1983), for instance, showed that changes in the starting point or deletion rate affect reliability and validity coefficients. Other researchers like Carroll (1980), Klein-Braley and Raatz

(1984), Klein-Braley (1983, 1985), Farhady (1983b), and Brown (1993) have questioned the reliability and different aspects of validity of cloze tests. In view of all the criticisms made against the cloze procedure, Klein-Braley and Raatz proposed the C-Test as a modified form of the cloze test.

The C-Test consists of four or five short texts in each of which the first sentence is left intact, then the *C-principle* (or *the rule of two*) is applied: the second half of every second word is deleted, beginning with the second word of the second sentence. If a word has an odd number of letters, the 'larger' half is omitted. Numbers, proper names, abbreviations, and one-letter words such as 'I' are ignored in the counting. In the canonical C-Test each text will have either 20 or 25 blanks. The students' task is to restore the missing parts. Only entirely correct restorations are counted as correct (i.e., spelling problems are considered errors). The testees would have roughly five minutes to answer each text, so that a test with five parts would take twenty five minutes to complete.

The C-Test is believed to have a number of advantages over the cloze test (Klein-Braley & Raatz, 1984; Klein-Braley, 1997). Some of the most important rewards of the C-Test are as follows:

1. The use of a variety of passages allows for a better sampling and representation of the language and content. Also, a person with special knowledge in a certain field cannot have an unfair advantage all through the test.
2. Since every second word is damaged, it is possible to obtain a better sampling of all the different language elements in a text.
3. C-Tests are very easy for native speakers. But someone who doesn't know the language at all normally scores zero or close to zero.
4. C-Tests are easy to construct, administer, and score.
5. As there is only one acceptable solution in most cases, the scoring is more objective.

Ever since it was introduced, the C-Test has been the subject of many research studies and scholarly controversies. On one hand, some researchers have found the C-Test a highly integrative, reliable and valid measure of overall language ability (Klein-Braley & Raatz, 1984; Cohen, Segal & Weiss, 1984; Klein-Braley, 1985, 1997; Chapelle & Abraham, 1990; Dörnyei & Katona, 1992; Huhta, 1996; Connelly, 1997; Ikeguchi, 1998; Babaii & Ansary, 2001; Eckes & Grotjahn, 2006; see Sigott, 2004 for an extensive review). More specifically, Klein-Braley (1997) empirically compared the C-Test with a group of other reduced redundancy tests – classical cloze test, cloze elide test,

multiple-choice cloze test, and standard dictation. She found that the best test to represent general language proficiency was the C-Test. Also, Eckes and Grotjahn (2006), using a Rasch model and confirmatory factor analysis, found clear evidence that the C-test is a highly reliable and unidimensional measure of general language proficiency. They found that “lexis and grammar are important components of general language proficiency as measured by C-tests” (p. 316).

However, research findings have not always been very consistent. Cohen et al. (1984), for instance, reported acceptable reliability and validity indices for their Hebrew C-Test, but they could not find any clear pattern for macro-level processing in the C-Test, though they found indications of micro-level processing (i.e., language processing at or below sentence level).

Dörnyei and Katona (1992) validated a C-Test against four different language tests including an oral interview, and a TOEIC (the Test of English for International Communication). Their results confirmed that the C-Test is a reliable and valid instrument. They also reported that their C-Test was a random and representative sample of the original text. Nevertheless, they noted that the C-Test was less efficient in testing grammar.

Analysing retrospective verbal protocols of 32 C-Test takers, Babaii and Ansary (2001) found four categories of cues used by the participants: automatic processing (16.6%), lexical adjacency (54.9%), sentential cues (22.4%), and top-down cues (6.1%). They reported that the test takers did fully exploit macro-level cues (though less frequently (28.5%) than micro-level cues (54.9%)) to restore the mutilations and concluded that the C-Test taps various aspects of language proficiency to varying degrees and, as such, it is a valid operationalization of the reduced redundancy principle. Notwithstanding, they maintained that their subjects mostly relied on their grammatical judgments to restore the mutilations –a finding which was contradictory to that of Dörnyei and Katona (1992).

On the other hand, researchers like McBeath (1989, 1990), Hughes (1989), Weir (1990), and Jafarpur (1995, 1999a, 1999b, 2002) have doubted some of the claims made on the part of the C-Test. In more specific terms, Hughes (2003, p. 195) referred to the puzzle-like nature of the C-Test as a disadvantage: “It is harder to read than a cloze passage, and correct responses can often be found in the surrounding text. Thus the

candidate who adopts the right puzzle-solving strategy may be at an advantage over a candidate of similar foreign language ability.” Along the same lines, Weir (1990) believes that the face validity of the procedure is low as it is irritating for students to have to process heavily mutilated texts.

Jafarpur’s (1995) findings showed that C-testing is suffering from the same shortcomings as the cloze procedure. He found that “the rule of two” is not a proper tool to obtain a representative sample of the basic elements of a text. He was able to show that different deletion starts and deletion ratios produce different tests with different results –which he interpreted as suggestive of the invalidity of the procedure (but see Hastings, 2002 for counter arguments). More interestingly, his analysis of his subjects’ answers to 10 attitudinal questions after taking the C-Test led him to the conclusion that “C-Tests do not possess face validity” (p. 209). The subjects on the whole believed that the C-Test is more of an IQ test or a test of spelling than a test of overall language ability. They believed it is more like a puzzle and is basically good for children.

In much the same vein, Jafarpur (1999b) substantiated that there is nothing magical about the rule of two. He was able to empirically demonstrate that other deletion rates and deletion starts yield more or less similar results.

In another study, Jafarpur (1999a) pretested a C-Test comprising 5 texts and 126 items with 146 subjects. On the basis of a classical item analysis, he discarded unsatisfactory items and developed a ‘tailored’ C-Test version with 100 items and tried it with 60 other subjects. The results indicated that classical item analysis does not improve the psychometric and statistical characteristics of the C-Test.

Furthermore, Jafarpur (2002) compared the performance of a C-test and a cloze test against a standardized criterion measure. The results showed that the C-Test enjoys a high reliability and concurrent validity and the deletions in a C-Test represent a more comprehensive coverage of different language elements than the cloze test. Yet he concluded that: (a) the C-Test is not an easily constructed, automatically reliable and valid measure of language competence, (b) the application of the ‘rule of two’ does not guarantee acceptable discrimination power for all items, (c) scoring does not offer any advantage over the cloze.

In the light of the variability and inconsistency of the results obtained with the C-Test, it seemed to the researcher that replicative investigations of the qualities of this testing

device are in order before definitive decisions can be made as to its credibility as a measure of overall language ability. Therefore, the current study set out to empirically explore aspects of validity and discriminatory power of the C-Test among Iranian EFL learners.

## **2. Method**

### **2.1. Instrumentation**

*a. The C-Test:* To construct the C-Test, thirteen texts were chosen from various EFL/ESL materials. The excerpts were authentic and self-contained and they varied in subject matter. The texts were of different levels of difficulty as judged by the Flesch Reading Ease readability scale (Microsoft Word, 1983–99) and a group of eight university EFL instructors. Every first sentence of each passage was left intact to provide a complete context. Beginning from word two of sentence two, the second half of every other word was deleted. In each mutilation, exactly half of the word was omitted, but if the number of letters was uneven, one extra letter was left out. Numbers, proper names and one-letter words were ignored in the counting and thus were not mutilated either. In this way, thirteen mutilated texts were produced with each one containing 20 gaps.

To facilitate pretesting, the extracts were randomly divided into two C-Tests which were, then, randomly given to 49 Iranian foreign language learners of English, 6 Iranian EFL teachers, and 3 native speakers. The completed test papers were scored giving one point for each exact restoration. The scores were item analyzed and five texts with superior discriminability and facility values were chosen. These texts were about culture, education, listening, bees, and underwater discoveries. They varied in difficulty with Flesch Reading Ease values of 62, 40, 75, 82, and 64, respectively. Dörnyei and Katona (1992) recommend the use of extracts with various difficulties in order to obtain equal measurement accuracy in both tails of a sample distribution.

The C-Test thus prepared comprised 100 gaps, fulfilling the recommended minimum number of mutilations (Klein-Braley, 1997; Raatz & Klein-Braley, 1995). The instructions were given in Persian along with a short English C-Test example and its restored answer. The final version of the C-Test can be found in Appendix I.

*b. The criterion measure:* The Form Q of the Michigan Test of English Language Proficiency (MTELP) (Corrigan, Dobson, Kellman, Spaan, & Tyma, 1979) was used as the criterion for determining concurrent validity coefficients. This test is a retired component of the Michigan English Language Assessment Battery (MELAB) which is a

discrete point language proficiency measure. The MTELP lasts 75 minutes to administer and comprises three subtests: 'Grammar', 'Vocabulary', and 'Reading comprehension'. The subtests contain 40, 40, and 20 four-choice items, respectively. The total score is the sum of the subtest scores. The manual reports reliability estimates of over .90 for the test and its subtests.

## ***2.2. Participants***

A total of 144 university students participated in this study. From these, 101 subjects took both the C-Test and the MTELP. They include: (a) 14 freshman, 22 sophomore, 23 junior, and 31 senior English majors studying at Khurasgan Azad University and the University of Isfahan, and (b) 11 engineering majors enrolled at an ESP course at Isfahan University of Technology. The other 43 subjects were all MA students of TEFL. They include 23 students at Najafabad Azad University, 14 students at Khurasgan Azad University, and 6 students at the University of Isfahan. These examinees took the C-Test only. The participants (mostly in their twenties) were of both sexes and enjoyed different levels of proficiency.

## ***2.3. Test administration and scoring***

In neither of the two tests had the participants been informed beforehand; so there was no preparation of any kind for the exam. The MTELP was first administered to the testees within the time limit of 75 minutes. The subjects were told that they would be informed of their grades, that their high scores on the test would affect their final term grades, and that high-ranking students would receive a prize. They were all informed that marks would be taken away for their wrong answers. The answer sheets were scored by the researcher. The MTELP scores were corrected for guessing in order to reduce the effect of chance (cf. Harris, 1969; Jafarpur, 1997). However, to remove the effect of practice, the subjects were not told that they were going to be tested again.

The C-Test was administered to the same subjects. However, since the subjects studied at different universities, the administration date varied from 10 to 14 days to cope with some limitations. It was assumed that the examinees' level of language proficiency had not changed significantly over the period. The completed C-Test papers were scored using the more convenient exact word scoring and counting spelling mistakes as incorrect. Alternative scoring procedures (acceptable word scoring, and tolerating spelling mistakes) have been shown to produce practically the same results as the one adopted in this study (Dörnyei & Katona, 1992; Huhta, 1996).

### 3. Results and discussions

The scores of the participants on all the tests and subtests were processed using the Statistical Package for the Social Sciences, Release 9.0.0 (SPSS, 1989-99). Table 1 shows descriptive statistics obtained from the C-Test, the MTELP, and their respective subparts, along with item facility (IF) and item discrimination (ID) indices of each C-Test text (C-Test, hereafter). In computing item facility and item discrimination indices each C-Test was considered a 'super-item' (see below). A sample separation procedure was adopted for computing item discrimination indices (Henning, 1987; Farhady, Jafarpur, & Birjandi, 1994).

**Table 1 - Descriptive statistics for the scores of the subjects on all measures**

Test	No. of Items	Mean	SD	Min.	Max.	IF	ID
(N = 144) C-test:	100	54.69	14.70	16	93		
C-Text 1	20	14.25	3.09	5	20	.70	.29
C-Text 2	20	11.79	3.49	3	20	.58	.30
C-Text 3	20	12.65	3.52	3	20	.62	.32
C-Text 4	20	9.45	4.36	0	20	.49	.47
C-Text 5	20	6.60	4.35	0	18	.35	.42
(N = 101) Michigan:	100	28.08	14.96	-3.33	70.33		
Grammar	40	15.94	8.84	-4.33	36		
Vocabulary	40	7.39	5.87	-.33	33.33		
Reading	20	4.75	4.18	-2.33	18.66		

The item discrimination values are in the range of .29 to .47 with mean value of .36 for the whole test. Jafarpur (1997, 2002) believes that item discrimination indices higher than .20 are acceptable. On this basis, texts in our C-Test demonstrate fairly low, yet acceptable item discriminability indices. The most attractive item facility and the highest item discrimination goes to C-Text 4 with an IF value of .49 and an ID value of .47.

The item facility indices for the five texts of the C-Test are in the acceptable range of .35 to .70 (cf. Raatz & Klein-Braley, 1995). The mean item facility for the whole C-Test is thus .55, which is very desirable (Henning, 1987). As far as item facility is concerned, except for C-Text 3, the other texts are arranged in an ascending order of difficulty.

As another index of relative difficulty, mean scores of the participants on each extract show the same pattern. They drop from 14.25 on C-Text 1, to 11.79 on C-Text 2, and after

a slight increase to 12.65 on C-Text 3 continue a descending route to 9.45 on C-Text 4 and then 6.60 on C-Text 5.

### **3.1. Reliability**

In order to allow better comparison, reliability coefficients for all the tests and subtests were estimated by the Kuder-Richardson Formula 21 (KR-21). The reliability estimate for the C-Test was also computed by the Cronbach's alpha formula. Both these formulas are measures of internal consistency.

Raatz and Klein-Braley (1995) suggest that it is possible to perform an inner consistency analysis on C-Tests. They agree that it is not permissible to define the individual blanks in the C-Test as items, since they are dependent on each other as a result of text structure and content. But they propose a practical solution: to consider each C-Test text as a super-item and then enter these four or five super-items into the Cronbach's alpha formula to estimate the reliability. Raatz (1985, p. 64) states:

Assuming that all the parts are independent of each other, but are equivalent and measure the same thing, then the total test score is the sum of the part scores. These parts can be viewed as superitems. In this case one can calculate intercorrelations and discrimination indices for the superitems without going inside the test parts. The reliability of the whole test can be calculated using Cronbach's alpha.

Table 2. Reliability indices for all tests

Test	KR-21	Cronbach's Alpha
C-test:	.90	.85
C-Text 1	.65	
C-Text 2	.63	
C-Text 3	.68	
C-Text 4	.78	
C-Text 5	.79	
Michigan:	.92	
Grammar	.90	
Vocabulary	.85	
Reading	.83	

Table 2 shows reliability coefficients of the two tests and their subparts computed by KR-21 formula. It also shows the reliability estimate for the C-Test computed by the Cronbach's Alpha formula. In doing so, each C-Test text was regarded as a 'super-item' and accordingly the alpha coefficient was calculated with five items.

Scores from both the C-Test and the MTELP show very high KR-21 reliability coefficients (.90 and .92, respectively). The reliability of the C-Test as estimated by the Cronbach's Alpha formula is also reasonably high (.85). The reliability coefficients of the scores obtained from the components of the MTELP are quite acceptably high too, the coefficients for each being over .83. However, only two subparts of the C-Test show satisfactory reliability indices, namely C-Text 4 (.78) and C-Text 5 (.79). The other three C-Tests demonstrate only moderately acceptable reliability with coefficients of .65 for C-Text 1, .63 for C-Text 2, and .68 for C-Text 3.

The fact that the whole C-Test is almost as reliable as the criterion (MTELP) appears to support claims concerning the high reliability of the C-Test (e.g. Klein-Braley & Raatz, 1984; Klein-Braley, 1985, 1997; Dörnyei & Katona, 1992; Connelly, 1997, to name a few).

### **3.2. Validity**

The primary concern for any test is that the interpretations and the uses we make from the test scores are valid. The evidence that we collect in support of the validity of a particular test can be of three general types: content relevance, criterion relatedness, and meaningfulness of construct (Bachman 1990). These categories have been separately discussed below with regard to the data presented in this study and the interpretations that can be legitimately made on their basis.

#### **3.2.1. Content validity**

A necessary stage in test validation is to investigate whether the test is relevant to a given area of content or ability. In the case of language tests, one principal concern of content validity is with the extent to which a test measures a representative sample of the language in question (Weir, 1990).

Table 3 represents the number and percentage of content and function words in the whole C-Test and each of its texts. In addition, it shows the number, percentage, and type of words mutilated in the same texts. In this analysis, auxiliary verbs, prepositions, conjunctions, pronouns, determiners, numbers, and adverbs (other than manner adverbs) have been counted as function words. The other words in the texts belong to categories of

nouns, verbs, adjectives, and adverbs of manner, which are typically considered content words.

Table 3 - Number and percentage of content and function words (mutilated) in each C-Text and in the whole C-Test

	Total (343 words)		Mutilated (100 words)	
	Content	Function	Content	Function
	Freq. %	Freq. %	Freq. %	Freq. %
C-Test	160 47	183 53	46 29	54 30
C-Text 1	27 42	37 58	8 30	12 32
C-Text 2	39 53	35 47	12 31	8 23
C-Text 3	25 42	35 58	7 28	13 37
C-Text 4	29 37	50 63	10 34	10 20
C-Text 5	40 61	26 39	9 23	11 42

The truncated words in each C-Text represent different parts of speech. In C-Text 1, as an example, four prepositions, three adverbs, two determiners, one pronoun, one auxiliary verb, and one numeric expression are mutilated. As for content words, there are five nouns, two verbs, and one adjective mutilated.

As is evident from the table, the percentage of content words mutilated in the whole C-Test (29%) is almost equal to the percentage of the function words mutilated (30%). Hence, the truncated words in the C-Test conform to the demands of content validity as they represent ‘a slice of reality’ (Raatz 1985, p. 63). Although, this finding does not accord with Jafarpur’s (1995) results, it compares very favorably with those of Dörnyei and Katona (1992) and Klein-Braley (1985) for it reveals that the C-principle is capable of obtaining a reasonably representative sample of all the word classes in a text.

### **3.2.2. Criterion validity**

Exploring the validity of a test by means of external criteria is seen as essential by many scholars (Weir, 1990; Bachman, 1990). Criterion-related evidence demonstrates a relationship between test scores and some criterion which is believed to be also an indicator of the ability tested. Concurrent validity is a kind of criterion-related validity which is obtained through concurrent administration of a newly developed test with another well-known standardized test of which the validity is already established (Hatch & Farhady, 1982; Brown, 1988).

Table 4 - Correlation coefficients among the scores of the two measures

Test	MTELP	C-Test
	(N = 101)	(N = 144)
C-Test:	.72	
C-Text 1	.54	.71
C-Text 2	.63	.73
C-Text 3	.63	.80
C-Text 4	.63	.85
C-Text 5	.45	.77
	(N = 101)	(N = 101)
MTELP:		.72
Grammar	.88	.70
Vocabulary	.81	.47
Reading	.59	.46

All correlations are significant at  $p < .01$  level (2-tailed).

Table 4 provides product-moment correlations among the scores from the C-Test and the MTELP. The table delineates that total C-Test scores correlate comparatively highly with total scores from the criterion (.72). The correlation coefficients between the C-Test and each of its C-Texts are quite high (.71, .73, .80, .85, and .77, respectively). There is also considerable correlation between the MTELP and the five C-Texts (.54, .63, .63, .63, and .45, respectively).

The C-Test shows a reasonably high correlation with the grammar subtest (.70). However, its correlations with the vocabulary and reading subtests are not very much promising (.47, and .46, respectively). These coefficients seem to contradict Dörnyei and Katona (1992) who found that the C-Test is less efficient in testing grammar. By contrast, these results are comparable with Chapelle and Abraham (1990) who concluded that the C-Test is more of a grammatically based test. Also Babaii and Ansary's (2001) finding

that their subjects mostly utilized their grammatical judgments to reconstruct the text is supported here.

Notice that the correlation of the C-Test with the MTELP was only moderately high (.72). A reasonable hypothesis is that the low face validity associated with the C-Test (Hughes, 2003; Weir, 1990; Jafarpur 1995) could most probably have affected the subjects' performance. If a test does not appear to the testees as face valid, then their adverse reaction to it results in a performance which is not a true reflection of their abilities. Weir (1990, p. 26) quotes Anastasi (1982, p. 136) who has argued:

Certainly if test content appears irrelevant, inappropriate, silly or childish, the result will be poor co-operation, regardless of the actual validity of the test. Especially in adult testing, it is not sufficient for a test to be objectively valid. It also needs face validity to function effectively in practical situations.

### ***3.2.3. Construct validity***

The main concern of language test makers is whether test performance truly reflects language abilities. Construct validation helps to substantiate the extent to which a testee's performance on a particular test can be indicative of his/her underlying competence. Construct validity, as characterized by Bachman (1990, p. 254), refers to 'the extent to which performance on tests is consistent with predictions that we make on the basis of a theory of abilities, or constructs'. In investigations of construct validity, therefore, we are concerned with empirically testing hypotheses about the relationships between test scores and underlying traits. Below there are reports on several analytical procedures conducted on the data obtained in this study to examine the construct validity of the C-Test.

#### ***3.2.3.1. C-Test and staged development of L2 competence***

One theory in second language learning holds that there is an orderly progress in L2 learning and learners go through a number of developmental stages, "from very primitive and deviant versions of the L2, to progressively more elaborate and target-like versions" (Mitchell & Myles, 1998, p. 10). In an attempt to establish the construct validity of the C-Test, Klein-Braley (1985) provides evidence that C-Tests support the theory of a regular progression in language learning. That is, since language competence increases progressively, if "the same C-Test is administered to the subjects at different stages of

language development, then the C-Test scores will become successively higher as the subjects become more proficient in the language” (Klein-Braley, 1985, p. 84).

To investigate the plausibility of this claim a special kind of subject grouping was required. Therefore, the undergraduate subjects were first classified into four proficiency groups based on the distance of their MTELP scores from the mean of the whole sample (the MA students were not included, for they had not taken the MTELP). The subjects whose scores were lower than  $-2/3$  SD below the mean were operationally classified as elementary level. Similarly, the lower intermediate level comprised examinees with scores between the mean and  $-2/3$  SD. Those whose scores were between the mean and  $+2/3$  SD were placed in the upper intermediate level. And finally, the advanced level contained examinees with scores more that  $2/3$  SD higher than the mean.

Test	Elementary (N = 26)	Lower Intermediate (N = 28)	Upper Intermediate (N = 23)	Advanced (N = 24)
	Mean SD	Mean SD	Mean SD	Mean SD
C-test:	36.38 10.59	51.79 7.78	56.39 8.84	65.21 15.44
C-Text1	11.08 3.46	13.75 3.92	14.35 2.60	16.04 3.16
C-Text 2	7.54 2.49	11.25 2.82	12.74 2.70	13.00 3.18
C-Text 3	8.88 3.22	12.11 2.48	13.83 2.55	14.83 3.41
C-Text 4	5.35 2.86	8.43 2.53	9.74 2.83	12.88 5.30
C-Text 5	3.38 2.84	6.25 2.82	6.22 4.28	8.50 4.86
MTELP:	10.40 4.51	23.41 2.91	32.36 3.03	48.58 9.81

Table 5 - Raw means and standard deviations for four proficiency groups

Table 5 presents means and standard deviations for the scores of the undergraduate subjects. As it is observed, the mean scores of both the criterion and experimental measures for the four groups increase progressively. Specifically, the mean scores on the C-Test become increasingly higher from a mean of 36.38 to 51.79 to 56.39 to 65.21, respectively. The mean scores on each C-Text behave in the same fashion, i.e., they

become successively higher as the level of proficiency increases. Though these means speak of validity for the C-Test, they should be subjected to further scrutiny to ensure their credibility. One way to do this is to examine the differences among the means of the four proficiency groups through an analysis of variance (ANOVA).

Table 6 - ANOVA results for the differences among means of four proficiency groups on the C-Test

Source of Variance	Sum of Squares	<i>df</i>	Mean Square	<i>F</i>	Sig.
Between Groups	10971.339	3	3657.113	30.480	.000
Within Groups	11638.305	97	119.983		

Table 6 shows the ANOVA results for the test of differences among the means obtained by the four proficiency groups on the C-Test. The obtained *F* ratio is significant at  $p < .000$  level suggesting that there *is* a difference among the means. However, it has to be noted that the significance of the *F* ratio in an analysis of variance merely indicates that there is a significant difference among the means of the compared groups as a whole; that is, it indicates that there is at least *one* significant difference between the means of at least *one* pair of the groups compared (Brown, 1988). All the same, it does not tell us where exactly this difference lies, i.e., exactly which two means are different. In order to determine exactly which means differ one has to resort to *pairwise multiple comparisons*, which are considered *post hoc* or *follow-up* tests (Hatch & Farhady, 1982). The only requirement for these tests is that the overall *F* in the ANOVA is statistically significant.

Table 7 represents the results of a Tukey's honestly significant difference (HSD) test conducted on the means of the four proficiency groups. Tukey's HSD test is a commonly used multiple comparison test which reveals the precise location of differences by analyzing every two means separately (Brown, 1988; Delavar, 2002). Table 7 denotes that there is significant difference between the means of every combination of two proficiency groups except for one: the upper and the lower intermediate groups. That is, the performances of the upper and the lower intermediate groups on the C-Test are not so much different that can be statistically acceptable.

Table 7 - Results of Tukey's HSD multiple comparisons on the means of the four proficiency groups

	Elementary	Lower Inter.	Upper Inter.	Advanced
Elementary	-----			
Lower Inter.	7.31*	-----		
Upper Inter.	9.02*	2.12	-----	
Advanced	13.16*	6.25*	3.90*	-----

\* Significant mean difference at  $p < .05$  level

The fact that the C-Test has not been able to produce significant distinction between the two middle groups in this study is indicative of a lucid shortcoming for the C-Test, namely a low classification power. These results are not only in clear contrast to claims about the measurement accuracy of the C-Test (Dörnyei & Katona, 1992) but they also challenge the dependability of using C-Tests for placement purposes (Klein-Braley, 1997). This interpretation is further supported by an investigation of decision consistency described below.

### **3.2.3.2. Decision consistency**

The scores from the C-Test were also studied for decision consistency. Decision consistency refers to the agreement between the classifications of the same examinees based on two tests of the same ability (Livingston & Lewis, 1995). In more practical terms, decision consistency is “the percent classifications of subjects by the experimental test that correspond correctly to those by the criterion” (Jafarpur, 2002, p. 42). Table 8 shows the percent correct classifications that are made if the C-Test was used as the criterion. As can be observed, the C-Test can on the average correctly place just over fifty percent of the subjects in their appropriate proficiency groups. It is by no means a promising quality for a test to fail to classify almost half of the examinees in their proper levels.

Table 8 - Percent of correct classification predicted by the C-test

Criterion for Placement	Elementary	Lower Inter.	Upper Inter.	Advanced	Average
C-test	69%	39%	35%	62.5%	51.5%

### 3.2.3.3. C-Test and text difficulty

In an attempt to establish the construct validity of the C-Test, Klein-Braley (1985, p. 88) claims that:

It is possible to show that while their empirically measured difficulty (as C-Test texts) varies according to the subject group involved, the group of texts used in any one C-Test remains more or less constant in terms of relative difficulty.

Therefore, one construct validity concern is to see whether C-Test texts (or C-Tests) function similarly across proficiency levels. In order to explore how similarly subjects from different levels of proficiency perform on each C-Text an ANOVA was performed on the scores obtained from the five C-Tests for the four proficiency groups and the MA students. It was assumed that the mean performance of a group of subjects on a C-Text can be a good index of the difficulty of that C-Text for that particular group.

Table 9 provides the outcome statistics of the ANOVA. The significance of the *F* ratio found for each group (at  $p < .000$  level) denotes that there are statistically meaningful differences among the means (i.e., average performances) of each group on the five C-Tests. Again, a Tukey's HSD test was carried out to specify on exactly which C-Tests the performances of each of the groups differ. Table 10 depicts the significant mean differences found among the five C-Tests for the four proficiency groups and the MAs.

Table 9 - ANOVA results for differences among means of five C-Texts for five proficiency groups

Proficiency Group	Source of Variance	Sum of Squares	df	Mean Square	F	Sig. (p)
Elementary (N = 26)	Between texts Residual	935.123 1119.000	4 125	233.781 8.952	26.115	.000
Lower Inter. (N = 28)	Between texts Residual	1006.857 897.286	4 135	251.714 6.647	37.871	.000
Upper Inter. (N = 23)	Between texts Residual	1056.617 1033.304	4 110	264.404 9.394	28.147	.000
Advanced (N = 24)	Between texts Residual	788.783 1918.917	4 115	197.196 16.686	11.818	.000
MA (N = 43)	Between texts Residual	1506.400 2480.233	4 210	376.600 11.811	31.887	.000

As observed, the mean performance of both the upper intermediate and the MA groups on the first three C-Texts are not significantly different. However, their means on C-Text 4 and C-Text 5 not only show a significant difference from the other three C-Texts but from each other as well. While the table indicates a similar pattern for the lower intermediate group, it is however visible that the performance of this group has changed noticeably from C-Text 1 to C-Text 2, too. On the other hand, the results obtained for the means of the advanced group on the five C-Texts represent a completely different pattern. For them it is simply the C-Text 5 which is significantly different from the other four C-Texts. The pattern of mean differences for the elementary group, however, is so complicated that it is almost impossible to interpret.

What is evident is that there are no less than *four patterns* of mean difference among these five groups. The fact that these five groups have performed differentially on the five C-Texts can be interpreted as counter evidence to Klein-Braley's (1985) claim concerning relative constancy of C-Test texts' difficulty independent of the subjects' proficiency level. These results are suggestive of the point that the C-Test suffers from one of the same problems as the cloze test does, namely the unpredictably variable nature of the cloze procedure (cf. Brown, 1993; see also Alderson, 1983; Klein-Braley, 1983). Jafarpur (1995) arrived at a similar conclusion as a result of comparing 20 C-Test versions developed based on the same text.

Table 10 - Results of Tukey's HSD for differences among the means of each proficiency group on the five C-Texts

		Elementary	Lower Inter.	Upper Inter.	Advanced	MA
C-Text 1	C-Text 2	*	*			
	C-Text 3					
	C-Text 4	*	*	*		*
	C-Text 5	*	*	*	*	*
C-Text 2	C-Text 1	*	*			
	C-Text 3					
	C-Text 4		*	*		*
	C-Text 5	*	*	*	*	*
C-Text 3	C-Text 1					
	C-Text 2					
	C-Text 4	*	*	*		*
	C-Text 5	*	*	*	*	*
C-Text 4	C-Text 1	*	*	*		*
	C-Text 2		*	*		*
	C-Text 3	*	*	*		*
	C-Text 5		*	*	*	*
C-Text 5	C-Text 1	*	*	*	*	*
	C-Text 2	*	*	*	*	*
	C-Text 3	*	*	*	*	*
	C-Text 4		*	*	*	*

\* Significant mean difference at  $p < .05$  level

### 3.2.3.4. Factorial validity

One of the most extensively used approaches in construct validation of language tests is factor analysis (Bachman, 1990). Factor analysis attempts to identify underlying variables, or factors, that explain the pattern of correlations within a set of observed variables (Farhady, 1983a; see also Oller & Hinofotis, 1980). Therefore, in order to further investigate the construct validity of the C-Test the scores of the subjects on the two measures were subjected to a factor analysis. To ensure higher precision, a principal axis factoring (PAF), as opposed to a principal components factoring (PCF), was employed to extract the initial factors (cf. Sharma, 1996; see also Carroll, 1983; Farhady, 1983a; Baker, 1989).

In order to determine the number of factors to be extracted, the eigenvalue-greater-than-one rule was utilized (Sharma, 1996). The eigenvalue-greater-than-one rule suggests that those factors whose eigenvalues (sum of

squared loadings) are less than unity be excluded from the analysis. It appeared that only the eigenvalue for the first factor exceeded unity. Accordingly, the one-factor solution was adopted as the most reasonable.

Table 11. Results of factor analysis (subtests only)

Test	Subtest	Factor 1
C-Test:	C-Text 1	.67
	C-Text 2	.75
	C-Text 3	.80
	C-Text 4	.80
	C-Text 5	.60
MTELP:	Grammar	.77
	Vocabulary	.57
	Reading	.49
Eigenvalue		4.283
Percent of total variance explained by the factor		53.539

Factor structure converged after 5 iterations

Table 11 presents the results of the factor analysis with loading patterns on the first factor (Factor 1). Almost all measures have high loadings on Factor 1 (i.e., have high correlations with it). The highest belongs to C-Text 3 and C-Text 4 (.80) and the lowest to the reading comprehension test (.49). Also Factor 1 explains 53.5% of the total variance, that is, more than half of the variance produced by the eight measures entered into the analysis is due to Factor 1, which probably can be best interpreted as accounting for overall proficiency of the subjects in English. These results can also be regarded as evidence that the tests to a large extent measure the same construct.

Another factor analysis was conducted, with MTELP entered as one single variable, so as to substantiate the results found above. The same methods were applied for factor extraction and for deciding on the number of factors. Table 12 shows the results of the second factor analysis where it was again only one factor whose eigenvalue was greater than one. That factor (once again termed Factor 1) could explain 62% of the variance with the MTELP heavily loading on it (.80) and with the C-Test texts demonstrating nearly the same loadings pattern as above. These figures serve to further confirm our conjecture that probably this first factor pertains to general proficiency in English. Given this conjecture is sustained, the comparatively high correlation of the C-Texts with the first factor can be

regarded as evidence that each of the C-Texts has a good claim to measuring language proficiency even on their own.

Table 12 - Results of factor analysis with MTELP as a single variable

Test (Subtest)	Factor 1
MTELP:	.80
C-Text 1	.66
C-Text 2	.75
C-Text 3	.81
C-Text 4	.81
C-Text 5	.61
Eigenvalue	3.732
Percent of total variance explained by the factor	62.199

Factor structure converged after 5 iterations

In view of the results of the two factor analyses just reported, it can be argued that the two experimental and criterion measures, to a great extent, tap the same underlying construct. Therefore, if what the MTELP measures is general language proficiency, then it is most probably what the C-Test measures as well. These results compare favorably with those of Jafarpur (2002) and Eckes and Grotjahn (2006), and provide support for Klein-Braley (1997) and Sigott's (2004) claims concerning the factorial validity of the C-Test.

#### **4. Conclusions**

The present study has primarily dealt with exploring the validity of the C-Test as reflected in domains of content relevance, criterion-relatedness, and construct meaningfulness. The reliability of the C-Test was also examined in the process. Reliability estimates found in this study confirmed earlier reports of high reliability coefficients associated with the C-Test.

A content/function word analysis was used to investigate the content validity of the C-Test. The C-principle showed a satisfactory method of sampling the linguistic elements in the text; hence, the claims of content validity made on its part are supported in this study.

As far as criterion-related validity is concerned, the C-Test scores correlated fairly highly with those of the MTELP. Not only that, but the C-Test's correlation coefficient with the MTELP was higher than with the grammar, vocabulary, and reading

comprehension tests. This is to be considered further evidence in favor of the claims that C-Tests measure general language ability.

The C-Test also was capable of fulfilling many of the requirements of a suitable test in terms of construct validity. The most important finding in this view was the factorial validity found with the C-Test scores. The subparts of the C-Test manifested the highest loadings on the same factor as the MTELP suggesting that not only the C-Test itself, but even the subparts of it had a substantial claim to measurement of general language proficiency.

The texts used in the C-Test, however, did not function uniformly with all proficiency groups. Each proficiency group found a different text or combination of texts more difficult. This finding reveals that when a text is turned into a C-Test, the C-Test text may unpredictably become more or less difficult for different proficiency levels. This is indicative of an unpredictable variability in C-Test results, a deficiency which has frequently been levelled against the cloze test as well.

As for the discrimination power of the C-Test, it came out that the C-Test did not perform very satisfactorily in differentiating subjects with different levels of linguistic ability. Specifically, the C-Test could not successfully discriminate between the subjects in lower and upper intermediate levels. In addition, a decision consistency analysis substantiated that the C-Test functioned poorly in classifying the participants in their appropriate proficiency levels. Therefore, contrary to C-Test proponents' claim (Klein-Braley & Raatz 1984; Klein-Braley 1997; Katona & Dörnyei 1993), the C-Test used in this study did not prove a very accurate and satisfactory placement test.

### **Acknowledgements**

My thanks are due to Dr. Mansour Koosha for his precious reviewing and constructive comments. I am also grateful to two Asian EFL Journal anonymous reviewers for their comments on an earlier version of this paper. Responsibility for any errors or deficiencies remains mine.

### **References**

- Alderson, J.C. (1979). The effect on the cloze test of changes in deletion frequency. *Journal of Research in Reading*, 2, 108-118.
- Alderson, J. (1980). Native and non-native performance on cloze tests. *Language Learning*, 30, 59-76.

- Alderson, J.C. (1983). The cloze procedure and proficiency in English as a foreign language. In J.W. Oller, Jr., (Ed.), *Issues in language testing research* (pp. 205-217). Rowley, MA: Newbury House.
- Anastasi, A. (1982). *Psychological testing*. London: Macmillan.
- Babaii, E. & Ansary, H. (2001). The C-test: a valid operationalization of reduced redundancy principle? *System*, 29, 209-219.
- Bachman, L. F. (1982). The trait structure of cloze test scores. *TESOL Quarterly*, 16, 61-70.
- Bachman, L. F. (1985). Performance on cloze tests with fixed-ratio and rational deletions. *TESOL Quarterly*, 19, 535-556.
- Bachman, L. F. (1990). *Fundamental considerations in language testing*. Oxford: Oxford University Press.
- Brown, J. D. (1983). A closer look at cloze: validity and reliability. In J.W. Oller, Jr., (Ed.), *Issues in language testing research* (pp. 237-250). Rowley, MA: Newbury House.
- Brown, J. D. (1988). *Understanding research in second language learning: a teacher's guide to statistics and research design*. Cambridge: Cambridge University Press.
- Brown, J. D. (1993). What are the characteristics of natural cloze tests? *Language Testing*, 10, 93-116.
- Carroll, B. J. (1980). *Testing communicative performance: An interim study*. London: Pergamon Institute of English.
- Carroll, J. B. (1983). Psychometric theory and language testing. In J.W. Oller, Jr., (Ed.), *Issues in language testing research* (pp. 80-107). Rowley, MA: Newbury House.
- Chapelle, C. A., & Abraham, R. A. (1990). Cloze method: what difference does it make? *Language Testing*, 7, 121-146.
- Cohen, A. D., Segal, M. & Weiss, R. (1984). The C-Test in Hebrew. *Language Testing*, 1, 221-225.
- Connelly, M. (1997). Using C-Tests in English with post-graduate students. *English for Specific Purposes*, 16, 139-150.
- Corrigan, A, Dobson, B., Kellman, E., Spaan, M., & Tyma, S. (1979). *Michigan test of English language proficiency (Form Q)*. Ann Arbor: Testing and Certification Division, the University of Michigan.
- Delavar, A. (2002). *Probabilities and applied statistics in psychology and educational sciences*. Tehran: Roshd.

- Dörnyei, Z. & Katona, L. (1992). Validation of the C-test amongst Hungarian EFL learners. *Language Testing*, 9, 187-206.
- Eckes, T. & Grotjahn, R. (2006). A closer look at the construct validity of C-tests. *Language Testing*, 23, 290-325.
- Farhady, H. (1983a). On the plausibility of the unitary language proficiency factor. In J.W. Oller, Jr., (Ed.), *Issues in language testing research* (pp. 11-28). Rowley, MA: Newbury House.
- Farhady, H. (1983b). New directions for ESL proficiency testing. In J. W. Oller, Jr, (Ed.), *Issues in language testing research* (pp. 253-269). Rowley, MA: Newbury House.
- Farhady, H., Jafarpur, A., & Birjandi, P. (1994). *Testing language skills: from theory to practice*. Tehran: SAMT.
- Harris, D. P. (1969). *Testing English as a second language*. New York: McGraw-Hill.
- Hastings, A. (2002). In defense of C-testing. In R. Grotjahn (Ed.), *Der C-Test. Theoretische Grundlagen und praktische Anwendungen [The C-Test: theoretical foundations and practical applications]* (Vol. 4, pp. 11-25). Bochum: AKS-Verlag.
- Hatch, E., & Farhady, H. (1982). *Research design and statistics for applied linguistics*. Rowley, MA: Newbury House.
- Henning, G. (1987). *A guide to language testing*. Cambridge, Mass.: Newbury House.
- Hinofotis, F. B. (1980). Cloze as an alternative method of ESL placement and proficiency testing. In J. W. Oller, Jr. & K. Perkins (Eds.), *Research in language testing* (pp. 121-34). Rowley, MA: Newbury House.
- Hughes, A. (2003). *Testing for language teachers* (2<sup>nd</sup> ed.). Cambridge: Cambridge University Press.
- Huhta, A. (1996). Validating an EFL C-test for students of English philology. In R. Grotjahn (Ed.), *Der C-Test. Theoretische Grundlagen und praktische Anwendungen [The C-Test: theoretical foundations and practical applications]* (Vol. 3, pp. 197-234). Bochum: Brockmeyer.
- Ikeguchi, C. B. (1998). Do different C-tests discriminate proficiency levels of EL2 learners? *JALT Testing & Evaluation SIG Newsletter*, 2, 3-8. Retrieved November 20<sup>th</sup>, 2006, from: [http://www.geocities.com/CollegePark/Field/1087/test/ike\\_1.htm](http://www.geocities.com/CollegePark/Field/1087/test/ike_1.htm)
- Jafarpur, A. (1995). Is C-testing superior to Cloze? *Language Testing*, 12, 194-216.
- Jafarpur, A. (1997). *An introduction to language testing*. Shiraz: Shiraz University Press.
- Jafarpur, A. (1999a). Can the C-test be improved with classical item analysis? *System*, 27, 79-89.

- Jafarpur, A. (1999b). What's magical about the rule-of-two for constructing C-Tests? *RELC Journal*, 30, 86-100.
- Jafarpur, A. (2002). A comparative study of a C-Test and a cloze test. In R. Grotjahn (Ed.), *Der C-Test. Theoretische Grundlagen und praktische Anwendungen [The C-Test: theoretical foundations and practical applications]* (Vol. 4, pp. 31-51). Bochum: AKS-Verlag.
- Jonz, J. (1976). Improving the basic egg: The multiple-choice cloze. *Language Learning*, 26, 255-265.
- Jonz, J. (1990). Another turn in the conversation: What does cloze measure? *TESOL Quarterly*, 24, 61-83.
- Katona, L. & Dörnyei, Z. (1993). The C-test: A teacher-friendly way to test language proficiency. *English Teaching Forum*, 31, 34-35.
- Klein-Braley, C. (1983). A cloze is a cloze is a question. In J.W. Oller, Jr., (Ed.), *Issues in language testing research* (pp. 218-228). Rowley, MA: Newbury House.
- Klein-Braley, C. (1985). A cloze-up on the C-test: a study in the construct validation of authentic tests. *Language Testing*, 2, 76-104.
- Klein-Braley, C. (1997). C-tests in the context of reduced redundancy testing: an appraisal. *Language Testing*, 14, 47-84.
- Klein-Braley, C. and Raatz, E. (1984). A survey on the C-test. *Language Testing*, 1, 134-146.
- Laesch, K. B., & van Kleek, A. (1987). The cloze test as an alternative measure of language proficiency of children considered for exit from bilingual education programs. *Language Learning*, 37, 171-189.
- Livingston, S. A., & Lewis, C. (1995). Estimating the consistency and accuracy of classifications based on test scores. *Journal of Educational Measurement*, 32, 179-197.
- McBeath, N. (1989). C-Tests in English: Pushed beyond the original concept? *RELC Journal*, 20, 36-41.
- McBeath, N. (1990). C-Tests – some words of caution. *English Teaching Forum*, 28, 45-46.
- Mitchell, R., & Myles, F. (1998). *Second language learning theories*. London: Arnold.
- Microsoft Word (1983–99). *Microsoft Word, Arabic Edition, Version 6.0*. Microsoft Corp.
- Oller, J.W., Jr. & Hinofotis, F. B. (1980). Two mutually exclusive hypotheses about second language ability: Indivisible and partly divisible competence. In J. W. Oller, Jr. & K. Perkins (Eds.), *Research in language testing* (pp. 13-23). Rowley, MA: Newbury

House.

Raatz, U. (1985). Better theory for better tests? *Language Testing*, 2, 60-75.

Raatz, E. & Klein-Braley, C. (1995). Introduction to language testing and C-tests. In Coleman (Ed.), *University language testing and the C-test*. Proceedings of a conference held at the University of Portsmouth in April 1995. Also retrievable from:  
<http://www.uni-duisburg.de/FB3/ANGLING/FORSCHUNG/HOWTODO.HTM>

Sharma, S. (1996). *Applied multivariate techniques*. USA: John Wiley & Sons Inc.

Sigott, G. (2004). *Towards identifying the C-Test construct*. Frankfurt am Main: Peter Lang.

SPSS (1989-99). *SPSS for windows, release 9.0.0*. SPSS Inc.

Weir, C.J. (1990). *Communicative language testing*. Hemel Hempstead: Prentice Hall.

## Appendix I

The C-Test developed by the researcher is presented here. The instructions were given in Persian along with an English example. Below, you can see the translation in English of the Persian instructions.

*In each of the texts below, the first sentence has been left intact. However, beginning with the second sentence, the second part of every other word has been omitted. In each mutilated word, the number of the deleted letters equals the number of remaining letters or one letter more. For instance, look at this text:*

There are usually five men in the crew of a fire engine. One of them drives the engine. The leader sits beside the driver. The other firemen sit inside the cabin of the fire engine.

*The above text is mutilated in this way:*

There are usually five men in the crew of a fire engine. One o\_\_\_\_ them dri\_\_\_\_ the eng\_\_\_\_. The lea\_\_\_\_ sits bes\_\_\_\_ the dri\_\_\_\_. The ot\_\_\_\_ firemen s\_\_\_\_ inside t\_\_\_\_ cabin o\_\_\_\_ the fi\_\_\_\_ engine.

*Now, try to complete the texts below. Marks will be taken off for spelling errors.*

-----

### *Text One*

In some cultures around the world, polygamy is recognized and accepted. This me\_\_\_\_(1) that a m\_\_\_\_(2) may ha\_\_\_\_(3) more th\_\_\_\_(4) one wi\_\_\_\_(5) or, i\_\_\_\_(6) some ca\_\_\_\_(7), a woman m\_\_\_\_(8) have mo\_\_\_\_(9) than o\_\_\_\_(10) husband a\_\_\_\_(11) the sa\_\_\_\_(12) time. Some\_\_\_\_(13) polygamous soci\_\_\_\_(14) occur wh\_\_\_\_(15), for so\_\_\_\_(16) reason, th\_\_\_\_(17) is a\_\_\_\_(18) imbalance bet\_\_\_\_(19) the num\_\_\_\_(20) of men and women, perhaps due to war, famine, or disease.

### *Text Two*

The way teachers teach is often a personal interpretation of what they think works best in a given situation. For ma\_\_\_\_(21) teachers, a teac\_\_\_\_(22) approach i\_\_\_\_(23) something uniq\_\_\_\_(24) personal, wh\_\_\_\_(25) they dev\_\_\_\_(26) through exper\_\_\_\_(27) and ap\_\_\_\_(28) in diff\_\_\_\_(29) ways acco\_\_\_\_(30) to t\_\_\_\_(31) demands o\_\_\_\_(32) specific situa\_\_\_\_(33). Teachers cre\_\_\_\_(34) their o\_\_\_\_(35) roles wit\_\_\_\_(36) the clas\_\_\_\_(37) based o\_\_\_\_(38) their theo\_\_\_\_(39) of teac\_\_\_\_(40) and learning and the kind of classroom interaction they believe best supports these theories.

**Text Three**

How can you learn to focus your attention better while listening? The mo\_\_\_\_(41) important th\_\_\_\_(42) is t\_\_\_\_(43) concentrate o\_\_\_\_(44) what t\_\_\_\_(45) speaker i\_\_\_\_(46) saying. Y\_\_\_\_(47), your da\_\_\_\_(48) tomorrow ni\_\_\_\_(49) is ve\_\_\_\_(50) important, b\_\_\_\_(51) right n\_\_\_\_(52) you mu\_\_\_\_(53) listen. B\_\_\_\_(54) firm wi\_\_\_\_(55) yourself. I\_\_\_\_(56) your mi\_\_\_\_(57) wanders, br\_\_\_\_(58) it ba\_\_\_\_(59) to list\_\_\_\_(60). You cannot concentrate and daydream at the same time.

**Text Four**

A family of bees is called a hive or a colony. It h\_\_\_\_(61) been sa\_\_\_\_(62) that i\_\_\_\_(63) a beehive, t\_\_\_\_(64) king i\_\_\_\_(65) actually a qu\_\_\_\_(66). The wor\_\_\_\_(67) are dev\_\_\_\_(68) to h\_\_\_\_(69), and th\_\_\_\_(70) wait o\_\_\_\_(71) her a\_\_\_\_(72) all ti\_\_\_\_(73) and sat\_\_\_\_(74) every ne\_\_\_\_(75) that s\_\_\_\_(76) might ha\_\_\_\_(77). They br\_\_\_\_(78) her t\_\_\_\_(79) best fo\_\_\_\_(80), and they lick her body whenever she passes by them in the hive. They guard her constantly so that no enemy can attack and harm her.

**Text Five**

Underwater archaeologists have it easy. Wrecks l\_\_\_\_(81) undisturbed f\_\_\_\_(82) centuries a\_\_\_\_(83) are pres\_\_\_\_(84) in go\_\_\_\_(85) condition. B\_\_\_\_(86) there i\_\_\_\_(87) one b\_\_\_\_(88) problem- i\_\_\_\_(89) is da\_\_\_\_(90) down th\_\_\_\_(91). Powerful lig\_\_\_\_(92) do n\_\_\_\_(93) help: ne\_\_\_\_(94) the bot\_\_\_\_(95), clouds o\_\_\_\_(96) tiny part\_\_\_\_(97) scatter t\_\_\_\_(98) light li\_\_\_\_(99) fog. A n\_\_\_\_(100) underwater TV system from Westinghouse Oceanics uses a fine beam of blue-green laser light to quickly scan the depths.



## **Lexical Collocations and their Relation to Speaking Proficiency of College EFL Learners in Taiwan**

**Jeng-yih Tim Hsu**

*National Kaohsiung First University of Science and Technology, Taiwan*

**Chu-yao Chiu**

*National Kaohsiung Hospitality College, Taiwan*

### **Bio Data:**

Jeng-yih Tim Hsu received his Master in English Language/Linguistics from University of Arizona, and holds a doctorate in Composition & TESOL from Indiana University of Pennsylvania. He is currently teaching at the Department of English, National Kaohsiung First University of Science and Technology, Kaohsiung, Taiwan.

Chu-yao Chiu is an adjunct lecturer in the Department of Applied Foreign Languages, National Kaohsiung Hospitality College, Kaohsiung, Taiwan. She has a Master in Applied Linguistics & TESOL from National Kaohsiung First University of Science and Technology. Her major field of research is L2 vocabulary acquisition.

### **Abstract**

The present study explores the knowledge and use of English lexical collocations and their relation to the speaking proficiency (assessed by two speaking tests) of Taiwanese EFL university learners. Data for the study were collected from 56 junior English majors at a national university of science and technology in southern Taiwan. Each student was asked to take three tests: (1) one lexical collocation test, measuring the subjects' knowledge of lexical collocations; (2) one English speaking test, administered to collect the subjects' use of lexical collocations and measure their speaking proficiency; and (3) PhonePass spoken English test, a standardized oral test, which was combined with the speaking test to measure the students' speaking proficiency. Test results were examined for correlations (1) between the subjects' knowledge of lexical collocations and their speaking proficiency, (2) between their use of lexical collocations and their speaking proficiency, and (3) between their knowledge of lexical collocations and their use of lexical collocations. The study findings showed that there was a significant correlation between Taiwanese EFL learners' knowledge of lexical collocations and their speaking proficiency. However, no significant correlation existed between the subjects' use of lexical collocations and their speaking proficiency. There was also no statistically significant correlation between the subjects' knowledge and use of lexical collocations. The current study concluded that knowledge of lexical collocations seemed to be a more significant indicator of speaking proficiency than the subjects' ability to use lexical collocations.

**Keywords:** lexical collocations, speaking proficiency, EFL education

## **1. Introduction**

Vocabulary and grammar are two basic elements of a language, but traditionally most EFL teachers often emphasize learning grammar more than vocabulary in their teaching. The role of vocabulary has long been underestimated in EFL education. With the approaching new millennium, pioneer scholars and researchers have started to direct people's attention to EFL vocabulary acquisition and began to emphasize the instruction of vocabulary in classroom practices (Channell, 1981; McCarthy, 1984; Nation, 1990; Nattinger, 1980). Among them, perhaps the most important one, Lewis (1993), proposed one groundbreaking teaching method, *the Lexical Approach*, and brought the field to systematically examine the nature of lexis in second language acquisition. He strongly argued that "language consists of grammaticalised lexis, not lexicalised grammar" (Lewis, 1993, p. vi). In Lewis' view, learning collocations, the key component of grammaticalised lexis, is equal to language learning.

A collocation, in its simplest definition, consists of two words which are linked together in the memory of native speakers and occur together with some frequency in both written and oral discourse (Aghbar, 1990). For example, *catch a cold* and *severe cold* are two commonly used word combinations that qualify as collocations. The verb *catch* and the adjective *severe* recurrently co-occur with the noun *cold*. In addition, many also believed that knowing a word includes knowing its collocations (Lewis, 2000; Nation, 1990, 2001).

Many scholars have maintained that collocational knowledge is one important factor that contributes to the differences between native speakers and foreign language learners (Aston, 1995; Fillmore, 1979; Kjellmer, 1991; Pawley & Syder, 1983). Particularly, failure to use collocations accurately for EFL learners is a major indicator of foreignness (McArthur, 1992; McCarthy, 1990; Nattinger, 1980; Wu, 1996). The strongest position held so far is that collocational competence is an indispensable component in the process of second/foreign language acquisition (Lewis, 1997, 2000; Nattinger & DeCarrico, 1992; Richards & Rogers, 2001).

### ***1.1 Purpose of the Study***

University students in Taiwan are currently required to pass an English proficiency test as one of the graduation requirements. Among various English proficiency tests available, GEPT (General English Proficiency Test), TOEFL (Test of English as a Foreign Language) and IELTS (International English Language Testing System) are the most

well-known among students. As a matter of fact, a speaking test is a basic component of these standardized exams. Nevertheless, acquiring speaking skills has been a challenge for most Asian students, and Taiwanese university students are no exception.

While searching for effective ways to improve EFL learners' speaking proficiency, many researchers have proposed that a good control of collocations can help language learners to speak more fluently (Brown, 1974; Ellis & Schmidt, 1997; Nation, 2001; Nattinger & DeCarrico, 1992; Pawley & Syder, 1983; Schmitt, 2000; Sung, 2003). Collocational knowledge could be a key element in enhancing the EFL learners' speaking ability. In Taiwan, a few researchers have carried out studies on collocational proficiency (Huang, 2001; Liu, 1999a; Tseng, 2002; Wu, 2005). However, most collocation studies in Taiwan investigate only the students' written production; few have looked into their speaking performance. Hence, the present study examines whether Taiwanese EFL learners' knowledge of collocations<sup>1</sup> is related to their speaking proficiency. The purpose of this study is to examine and describe the knowledge and use of collocations and their relation with English speaking proficiency among Taiwanese EFL university learners. To fulfill the purpose of the study, the experiment carried out in this study aimed to answer the following research questions:

- (1) To what extent is subjects' knowledge of lexical collocations (KLC) correlated with their general speaking proficiency (GSP)?
- (2) To what extent is subjects' use of lexical collocations (ULC) correlated with their general speaking proficiency (GSP)?
- (3) To what extent is subjects' knowledge of lexical collocations (KLC) correlated with their use of lexical collocations (ULC)?

### ***1.2 Significance of the Study***

The study is quantitative in nature and examines Taiwanese EFL learners' knowledge and use of collocations and further explores the relationship between collocations and their speaking proficiency. Since no study has been conducted to report such performance and the relationship between collocations and speaking in Taiwan, this study contributes directly to our understanding of the nature of collocations in speaking.

## **2. Review of Related Studies**

The review of literature introduces empirical studies particularly targeting EFL learners (both non-Taiwanese and Taiwanese) and examines the possible connection between lexical collocations and English proficiency.

### ***2.1 Studies of the Relation between Collocations and Language Proficiency of non-Taiwanese EFL Learners***

Zhang (1993) was the first scholar to explore the possible correlation between the knowledge and use of English collocations and the quality of college freshmen's writing. At a mid-size public university in Pennsylvania, in the United States, the 60 college freshmen in his study were categorized into two groups: 30 native and 30 non-native speakers of English. Within each group, two subgroups, i.e., Good writers and Poor writers, were established based on a writing test. Each subject completed one fill-in-the-blank collocation test and one writing task. The collocation test was used to measure the subjects' collocational knowledge; the writing task was used to elicit the subjects' use of collocations and writing proficiency. In this experiment, Zhang (1993) found that (1) native English writers performed significantly better than non-native writers on the collocation test, and Good writers within either group performed significantly better than Poor writers; (2) as for the use of collocations in their writing, native writers surpassed the non-native writers, and Good writers within either group surpassed Poor writers; (3) in terms of writing performance, a significant difference was found between non-native Good and Poor writers. Zhang drew two conclusions based on the observed correlations. Collocational knowledge is a source of proficiency in writing among college freshmen. Besides, quantity, but more important, quality for use of collocations distinguish between native and non-native college freshmen writing as well as between Good and Poor college freshmen writing.

After Zhang, Al-Zahrani (1998) investigated the knowledge of English lexical collocations among four academic levels of Saudi EFL university students and the relationship between the participants' collocational knowledge and their general language proficiency. In his study, the collocational knowledge of 81 Saudi male university English majors was measured by a cloze test, comprised of 50 "verb+noun" lexical collocations. Furthermore, the participants' general English proficiency was assessed by a writing test and a paper-and-pencil TOEFL test. Al-Zahrani found that there was a significant difference in his subjects' knowledge of lexical collocations among the different academic

years. The knowledge of lexical collocations increased with the subjects' academic years. Besides, he reported that there was a strong correlation between the subjects' knowledge of collocations and their overall language proficiency.

Sung (2003) examined the knowledge and use of English lexical collocations in relation to speaking proficiency of international students enrolled in a university in Pittsburgh area. A total of 72 non-native English speakers participated in her study. Each subject completed two tests: one collocation test and one speaking test. The first test was used to measure the subjects' knowledge of lexical collocations while the latter was used to elicit the subjects' use of lexical collocations and measure their speaking proficiency. Her results showed that there was a significant correlation between the knowledge of lexical collocations and the subjects' speaking proficiency.

Up to this point, three major studies have investigated the relation between EFL learners' collocational knowledge and their language proficiency (Al-Zahrani, 1998; Sung, 2003; Zhang, 1993). The results of the studies all showed similar positive correlations between these learners' collocational knowledge and their language proficiency.

## ***2.2 Studies of the Effect of Explicit Collocation Instruction on Language Skills of Taiwanese EFL Learners***

In Taiwan, while the majority of studies have looked at error analysis of collocations (Chen, 2002; Hsueh, 2003; Liu 1999a, 1999b; Tang, 2004), only six researchers have applied collocation instruction in classroom to observe the effects on students' language skills.

In one of her collocation studies, Liu (2000) led the field in investigating the effects of collocation instruction on students' writing performance. Forty-nine freshmen English majors at a Taiwanese university participated in the study. During an 18-week semester, in a three-hour weekly class, Liu (2000) gave her students a series of twenty-minute mini-lessons on collocations (e.g., the introduction of six major lexical collocation patterns, collocations without direct L1 equivalents, and de-lexicalized verbs as collocates of nouns). To assess their writing ability and use of collocations, the subjects were asked to write a composition in class at the beginning and the end of the semester without using a dictionary. The two compositions were analyzed and compared for the patterns of acceptable and unacceptable lexical collocations. It was found that the students in the second composition generated a greater number and variety of acceptable lexical collocations although they did not improve much in their writing of the second

composition. After Liu, Lin (2002) in a two-week course introduced 8 collocation activities (such as brainstorming for collocates of a word, underlining verb-noun collocations, and matching game) to her 89 senior high school students to explore the effects of collocation instruction on Taiwanese EFL learners' vocabulary development. Subjects also took a pre- and post-course collocation test (evaluating "productive collocation competence" (p. 84)) and also a pre- and post-course multiple-choice test (assessing just "recognitive collocation competence" (p. 85)). Lin reported that the students were found to have improved in both their receptive and productive competence after receiving the systematic instruction on collocations. In contrast with Liu (2000) and Lin (2002), Tseng (2002) examined the effects of direct collocation instruction on both vocabulary improvement and writing performance. In a 12-week period, Tseng taught ninety-four senior high school students in two groups, the experimental group (receiving collocation instruction) and the control group (no collocation instruction). She used some of Liu's (2000) activities on collocation, including introduction of collocations, collocations without L1 equivalents, and added two more teaching activities in her study, i.e., using collocation dictionaries and identifying collocations in the textbook. All the subjects in both groups did three kinds of tasks: (1) a questionnaire, (2) a pre- and post-instruction fill-in-the-blank collocation test, and (3) a pre- and post-instruction composition. Although the participants did not exhibit obvious improvement in their performance on the composition after receiving the collocation instruction, Tseng reported that the collocation instruction somewhat improved the subjects' vocabulary development. This finding is somehow different from Lin's (2002) study in which Lin reported her students' progress in "collocation competences" (p. 85). A possible explanation might be the instruments they used in assessing their subjects' improvement. In fact, unlike Tseng who analyzed an actual writing test, Lin relied on her fill-in-the-blank and multiple-choice tests only.

A teacher in a private Taiwanese university, Hsu (2002), also conducted an intensive business English workshop to investigate whether collocation instruction would strengthen Taiwanese college EFL learners' development of the knowledge and use of collocations, as well as promote language proficiency. Seven English and two Banking & Finance majors at a private Taiwanese university took part in this workshop. The workshop ran for a total of 48 hours in one month (i.e., three hours a day, four days a week). As a teacher-researcher, Hsu collected and analyzed students' writings, results of pre- and post-workshop collocation tests, teacher's class notes, student interviews, and videotapes

of the workshop sessions. After examining the data, he found that (1) collocation instruction appeared to have helped students acquire new collocations in both written and spoken discourse; and (2) there appeared to be a positive correlation between the students' use of lexical collocations and their language proficiency. Based on the findings, he concluded that explicit collocation instruction is highly valuable and recommended that it should be incorporated into EFL classrooms.

Investigating a specific language skill, Lien (2003) addressed the effects of collocation instruction especially on reading comprehension. A total of 85 Taiwanese university students at three academic levels participated in her study. In the pre-experimental stage, each student took one fill-in-the-blank collocation test and reading comprehension pre-test. The collocation test and reading comprehension pre-test were administered in order to investigate the relationship between knowledge of collocations and reading comprehension. In the experimental stage, all subjects at the three academic levels received three types of instruction (collocation instruction, vocabulary instruction, and no instruction); that is, subjects served as their own control group as well as participating in the experimental conditions. The subjects then took a reading comprehension post-test immediately after each type of instruction. The three reading comprehension post-tests were used to evaluate the effectiveness of the three types of instruction. In the post-experimental stage, a questionnaire was administered to collect the subjects' opinions and attitudes toward the three types of instruction. The major findings of the study included: (1) collocational knowledge was associated with reading comprehension; (2) the subjects performed better after receiving collocation instruction compared with the other types of instruction; and (3) the subjects also reported positive attitudes toward the direct teaching of collocations.

Looking at another specific language skill, L. Hsu (2005)<sup>2</sup> investigated the effects of explicit collocation instruction on EFL learners' listening comprehension. Thirty-four Taiwanese university students participated in her study. Over a three-week period, each subject received three different types of instruction (single-item vocabulary instruction, lexical collocation instruction, and no instruction), and took one listening comprehension test right after each type of instruction. In the fourth week, the subjects completed a questionnaire. The results indicated that the subjects performed best after receiving lexical collocation instruction. Furthermore, the students indicated in the questionnaire that: (1) collocation instruction was their preferred instruction type; (2) they were willing to learn

more about lexical collocations; and (3) they believed “their listening fluency can be improved as a result of instruction of lexical collocations” (p. 63).

In sum, empirical studies concerning the effects of explicit collocation instruction on Taiwanese EFL learners have been carried out to explore many aspects of language skills, including writing (Liu, 2000; Tseng, 2002), vocabulary development (Lin, 2002; Tseng, 2002), reading (Lien, 2003), listening (L. Hsu, 2005) and overall proficiency (Hsu, 2002). Among these studies, the results generally showed that direct collocation instruction was positively correlated with and possibly improved Taiwanese EFL learners’ language performance. In addition, the subjects in the majority of the studies displayed a positive attitude toward collocation instruction.

### **3. Research Method**

#### ***3.1 Research Design***

In this study, the data were gathered by three tests, including one written and two spoken tests. They are a fill-in-the-blank lexical collocation test, an English speaking test, and a PhonePass spoken English test. This study is also a quasi-experimental design; it is similar to true experimental designs, except that the participants are not randomly selected for the study (Mertens, 1997).

Three variables elicited from the three tests were used in the study. First, knowledge of lexical collocations (KLC) was measured by the subjects’ scores on the fill-in-the-blank lexical collocation test. Second, use of lexical collocations (ULC) was calculated by the quantity of lexical collocations collected from the subjects’ speaking test. Last, general speaking proficiency (GSP) was measured and combined from the average scores of the two speaking tests (i.e., English speaking test and PhonePass spoken English test). In this study, descriptive data were collected from the three tests. The Pearson correlation coefficient was adopted to determine the degree of a relationship between two quantifiable variables.

#### ***3.2 Study Setting and Subjects***

The subjects in this study were 56 EFL junior English majors at a university of science and technology in southern Taiwan. These subjects were chosen for the following reasons. Firstly, the purpose of the study is to examine the relation between knowledge of lexical collocations and general speaking proficiency of the EFL learners. Furthermore, with at least 8 years formal English training, the junior English majors should have learned a

certain number of collocations, and also would have had many opportunities to practice speaking English. Besides, no previous study has ever looked into this in Taiwanese university settings. Thus, the researchers chose only Taiwanese EFL learners as the target subjects. An investigation of these students' performance would provide the researchers with sufficient data to explore their collocational performance and speaking proficiency.

### **3.3 Instruments**

#### **3.3.1 Lexical collocation test**

The fill-in-the-blank lexical collocation test consisted of 50 items. The subjects were expected to provide the best answer for the target collocate in the blank. In this specific test, the first letter/phoneme was provided because it could help subjects trigger the appropriate target collocate, and reduce the possibility of guessing by test takers. A sample question for L2 type is "If you want to have a b\_\_\_\_\_ future, you need to set your goals in advance, and turn your plan into actions." The anticipated collocate for the question is "bright."

The test included five major types of lexical collocations as categorized by Benson, Benson, & Ilson (1997). Each type had 10 test items. Table 1 summarizes the distribution of the number for lexical collocations in the test. The underlined part of speech in Table 1 indicates the missing part in the test which the subjects needed to fill in.

Table 1 *Number of Items Used in the Test of Lexical Collocations*

Type	Pattern	Number of Questions
L1	verb + noun	10
L2	adjective + noun	10
L3	noun + verb	10
L4	adverb + adjective	10
L5	verb + adverb	10
Total Items		50

#### **3.3.2 Pear Film Speaking test**

Subjects took the Pear Film speaking test to elicit collocations, which would provide the quantity of their use of lexical collocations (ULC). The test results were later considered with those of the PhonPass speaking test as a measurement of the subjects' speaking proficiency. The *Pear Film* was produced under the direction of Chafe (1980) at the University of California, Berkeley. Chafe noted that the *Pear Film* was designed for the purpose of analyzing how people talk about what they have experienced and later how they recall these experiences. This 6-minute film covering a series of natural events can be

easily understood. After watching, viewers should be able to recall what they saw and describe how events in the film related to their daily life.

When attending the speaking test of *Pear Film*, the subjects were invited to watch the film individually and answered the following three questions orally:

- (1) After watching the Pear Film, please tell me what happened in the film?
- (2) Please explain what you consider to be the meaning behind the film.
- (3) Please relate an aspect of the film (any aspect) to something in your own life or something you have observed in real life.

The subjects were encouraged to speak as much as possible as there was no time limit. Later, the subjects' oral descriptions were transcribed, and the quantity of the lexical collocations was counted and recorded.

### **3.3.3 PhonePass spoken English test**

The PhonePass spoken English test is a standardized test developed by Townshend and Bernstein in 1996 (Validation summary for PhonePass, 2005). It is the only English spoken testing system which combines computerized design, telephone, Internet and professional test contents to measure the test taker's speaking proficiency over the telephone. There are three major characteristics of PhonePass spoken English test. This test is continuously and conveniently available from any telephone on a 24-hour basis. Also, the test is scored automatically by a computer-based system. Lastly, the test scores can be efficiently displayed on the PhonePass website (<http://www.phonepass.com.tw/>) one hour after taking the test. Currently three versions (i.e., 10-minute, 7-minute, and 5-minute) of the PhonePass spoken English tests are available.

The format used in this study was a 7-minute version for each subject. The test consists of 44 items which are divided into four sections (Parts A through D). Each of the four parts presented the test takers with a different task: (A) reading aloud, (B) repeating, (C) answering short questions, and (D) building sentence. In Part A, the subjects read the sentences shown in the test sheet. For example, they read the sentence, *There are advantages and disadvantages to eating alone*. In Part B, the subjects had to repeat what they heard on the telephone. For instance, a voice says, *Leave town on the next train* and the subjects would repeat. In Part C, the subjects must give a simple answer to the questions. For example, a voice says, *Would you get water from a bottle or a newspaper?* and the subjects had to give short answers of *a bottle* or *from a bottle*. In Part D, the subjects had to arrange the scrambled words or phrases into a complete sentence. For

example, a voice says, *was reading.. my mother.. her favorite magazine..* and the subjects had to unscramble them and make one complete sentence of *My mother was reading her favorite magazine.*

ORDINATE Corporation, the publisher, has conducted series of studies to examine the validity and reliability of PhonePass spoken English test. In terms of inter-rater reliability, it reported a positive correlation of 0.73 between PhonePass spoken English test results and paper-pencil TOEFL scores, and a high correlation of 0.94 with human scorers (Validation summary for PhonePass, 2005).

### ***3.4 Data Collection Procedures***

This study consisted of two experimental stages. In the first stage, each subject was required to take the lexical collocation test and then English speaking test of Pear Film individually. In the second stage, the researchers delivered the PhonePass spoken English test sheet to the subjects and explained how to operate the test. All the subjects were asked to take the PhonePass spoken English test one week after completing the first stage of experiment. Some subjects were excluded because they did not complete the PhonePass spoken English test before January 13, 2006 (as this was last class of the academic semester before winter vacation). Altogether, there were 56 subjects at the end of the study.

### ***3.5 Scoring***

#### ***3.5.1 Scoring for lexical collocation test***

The responses on the lexical collocation test were rated by using a 3-point partial-credit scoring method, suggested by Aghbar and Tang (1991). Each item was assigned 0-2 points, totaling 100 points for the 50-item test. The researchers scored the lexical collocation test with the help of *the BBI dictionary of English word combinations* (Benson, Benson, & Ilson, 1997), two online corpora: Simple Search of British National Corpus (<http://sara.natcorp.ox.ac.uk/lookup.html>) and VLC Web Concordancer (<http://vlc.polyu.edu.hk/concordance/WWWConcappE.htm>) and two native speakers who are both experienced EFL teachers. Online corpora were utilized in this study because evidence showed that a corpus could demonstrate how words collocate (Schmitt, 2000; Scholfield, 1997). The two online corpora were used because they were both free of charge. In addition, they were designed to include a wide range of samples from different registers and genres. Although these two corpora are much closely related to British

English, the researchers also asked help of two native speakers from the United States, making efforts to balance the judgment of lexical collocations between British English and American English.

When a response could be found from *the BBI dictionary of English word combinations*, it would receive 2 points. Alternatively, if the response was not found in *the BBI dictionary of English word combinations* but was found in either one of the two online corpora or confirmed by either one of the two native speakers, then the subjects would receive 1 point. This decision was made based on two considerations. First, online corpora may include many possible collocations which are used loosely, representing a wider range of possibility. Second, the native speakers' judgment may merely be limited to the regional feature of their English. Besides, as the purpose of the test was to elicit the subjects' collocational knowledge, rather than judge their grammar or spelling errors, errors of these types were overlooked.

### ***3.5.2 Scoring for Pear Film speaking test***

The scoring of Pear Film speaking test was conducted holistically by two raters who are both experienced English teachers. The IELTS holistic scoring system was adopted to evaluate the subjects' performance based on the two reasons: its popularity in Taiwan and speaking test component.

To ensure inter-rater reliability, a practice session was held. All the materials including IELTS scoring system were introduced, so that the two raters could fully understand the rating scale. After watching the Pear Film and listening to a sample of spoken data, the two raters assigned scores and made recommendations. Their ability to assign scores consistently was practiced before they were exposed to the current research rating task. When starting to score the data of the speaking test of Pear Film, each spoken data was scored by the two raters separately. When the difference in scoring was greater than 2 points, the raters were asked to re-rate the speaking data. Then the average of their ratings was used as the final score for the subject's speaking proficiency in the test of Pear Film.

### ***3.5.3 Tallying the frequency of lexical collocations from Pear Film speaking test***

The spoken data obtained from Pear Film speaking test were transcribed and all the lexical collocations were manually counted. An acceptable lexical collocation was recorded based on the same criteria (i.e., the verification from 2 online corpora or 2 native speakers) adopted in evaluating the lexical collocation test. Since all the subjects' total length of

speeches differed, the total words uttered were different. In order to make comparison at an equal scale, the total words in each subject' spoken data were converted to 400 words<sup>3</sup>.

### **3.5.4 Scoring for PhonePass spoken English test**

The scores for the PhonePass spoken English test were obtained from the official website for PhonePass. The overall scores are reported in the range from 20 to 80. Scores are measured based on the combinations of the four sections on the test. Figure 1 shows one sample score of PhonePass spoken English test.



*Figure 1. A Sample Score of PhonePass Spoken English Test*

After obtaining the scores of the two major speaking tests, the researchers combined them to produce general speaking proficiency (GSP). Because the scoring systems between the two speaking tests were totally different, the scores of the Pear Film speaking test were converted into the measurement of the PhonePass spoken test.

### **3.6 Data Analysis and Statistical Procedures**

All the data were computed by Statistical Package for the Social Science (SPSS) version 12. To understand the general picture of the subjects' knowledge of lexical collocations (KLC), use of lexical collocations (ULC) and their general speaking proficiency (GSP), a

preliminary analysis was first conducted by computing the descriptive statistics of these three variables for the study. In the second stage, Pearson product-moment correlation<sup>4</sup> (Bachman, 2004; Haslam & McGarty, 2003; Schweigert, 1994; Weiss, 2005) was adopted to measure possible correlations of KLC and GSP, ULC and GSP, KLC and ULC respectively and to answer the 3 research questions of the study.

#### **4. Results & Discussion**

##### ***4.1 Analysis of Research Question One***

Research Question One asked: To what extent is the knowledge of lexical collocations (KLC) correlated with the general speaking proficiency (GSP) among the subjects? The Pearson correlation was used to measure correlations between the scores of KLC and GSP among all subjects; in addition, the L1-L5 subtypes in the KLC scores were also further investigated to analyze their correlations with GSP. Table 2 presents the correlations between KLC and GSP.

Table 2 *Pearson Correlation between Knowledge of Lexical Collocations and General Speaking Proficiency (N=56)*

	KLC	KLC-L1	KLC-L2	KLC-L3	KLC-L4	KLC-L5
GSP	$r=.561^{**}$ $p=.000$	$r=.309^{*}$ $p=.020$	$r=.309^{*}$ $p=.021$	$r=.440^{**}$ $p=.001$	$r=.367^{**}$ $p=.005$	$r=.400^{**}$ $p=.001$

*Note.* The asterisks \*\* indicate correlation is significant at the 0.01 level; the asterisk \* indicates correlation is significant at the 0.05 level.

Table 2 shows that the relationship between KLC scores and GSP scores is positively correlated. Statistically significant correlation exists between GSP scores and total KLC scores for the subjects ( $r = 0.561$  at  $p < 0.01$  level). Specifically, the L1-L5 subtypes of KLC scores are also moderately correlated with GSP scores. In brief, there is a moderately positive correlation between GSP scores and KLC scores among the 56 subjects. This could suggest that KLC could be a predictor to measure the subjects' general English proficiency.

In the same manner, the results of previous studies also showed positive relationship between the knowledge of collocations and other language skills, such as writing (Zhang, 1993); language proficiency (Al-Zahrani, 1998; Bonk, 2000), reading (Lien, 2003) and

speaking (Sung, 2003). Thus, it could be concluded that knowledge of collocations is indicative of non-native speakers' language proficiency.

#### 4.2 Analysis of Research Question Two

Research Question Two asked: To what extent is the use of lexical collocations (ULC) correlated with the general speaking proficiency (GSP) among the subjects? The Pearson correlation was employed to measure correlation between the frequency of ULC and the scores of GSP among all subjects; moreover, the L1-L5 subtypes in the frequency of ULC would be also further investigated to determine their correlations with GSP. Table 3 presents the correlations between ULC and GSP.

Table 3 *Pearson Correlation between Use of Lexical Collocations and General Speaking Proficiency (N=56)*

	ULC	ULC-L1	ULC-L2	ULC-L3	ULC-L4	ULC-L5
GSP	$r=.253$ $p=.060$	$r=.259$ $p=.054$	$r=.184$ $p=.175$	$r=.085$ $p=.534$	(a)	$r=-.097$ $p=.478$

Note. The symbol "(a)" indicates that the correlation cannot be computed because one variable is constant<sup>5</sup>.

According to Table 3, most of the variables are weak but positively correlated. The correlation between the GSP scores and the total frequency of ULC is not strong enough to be significant among all subjects. The case between the GSP scores and the individual L1-L5 subtypes is the same. In short, the results show that only a positively weak correlation occurs between the GSP scores and the frequency of ULC. This might suggest that the ULC does not seem to be a strong predictor for measuring the subjects' general English proficiency.

This finding corresponded with Hsu's (2002) study in which he found that there is no strong relationship between the subjects' quantity of collocations and their speaking performance. However, the finding was somewhat different from Sung's (2003) study. Sung reported that there was a moderate correlation between the speaking proficiency and the frequency of collocations used in the speaking test among the non-native speakers residing in the United States. One possible explanation might be due to the different focuses on the categories of lexical collocations. Sung (2003) included more subtypes of lexical collocations (L1-L7 based on Benson et al., 1997) in her recording of students'

spoken data. In the current study, the researchers looked at the subjects' L1-L5 subtypes of lexical collocations which are fewer than Sung's. Another possible explanation might be that the subjects in Sung's study were residents in the United States; they may be more proficient in English.

Still, it is beyond the scope of the current study to draw a concrete conclusion on the relationship between use of collocations and speaking proficiency. Therefore, further larger scale research could investigate the correlation between these two variables.

#### 4.3 Analysis of Research Question Three

Research Question Three asked: To what extent is the knowledge of lexical collocations (KLC) correlated with the use of lexical collocations (ULC) among the subjects? Table 12 presents the correlations between KLC and ULC and the 5 subtypes found in KLC and those from ULC for all the subjects. According to Table 4, only a weak but positive correlation occurs between the KLC scores and the frequency of ULC among the subjects. In other words, the linear correlations between KLC scores and frequency of ULC are not significant. Such similar finding was formerly reported by Zhang (1993) who also found that a weak correlation existed between the subjects' knowledge of collocations and use of collocation from their writing.

Table 4 Pearson Correlation between Knowledge of Lexical Collocations and Use of Lexical Collocations (N=56)

	KLC	KLC-L1	KLC-L2	KLC-L3	KLC-L4	KLC-L5
ULC	r=.035 p=.798					
ULC-L1		r=-.085 p=.533				
ULC-L2			r=.150 p=.269			
ULC-L3				r=.192 p=.157		
ULC-L4					(a)	
ULC-L5						r=.210 p=.120

Note. The symbol "(a)" indicates that the correlation cannot be computed because one variable is constant<sup>6</sup>.

However, the findings of Zhang (1993) and the current study are somewhat different from Sung's study in 2003. In her results, there was a moderately significant correlation between the knowledge of lexical collocations and frequency of lexical collocations in the speaking performance of the non-native speakers in the United States. The relationship between knowledge of collocations and their use of collocations is not obvious and we are

far from drawing any convincing conclusion. Therefore, further investigation could look into this particular relation.

#### 4.4 Summary of Research Findings

The present study intended to examine the possible connections among the three variables: KLC, ULC, and GSP. Table 5 clearly shows the Pearson correlation among the three variables.

Table 5 Pearson Correlation among KLC, ULC, and GSP (N=56)

	GSP	ULC
KLC	$r=.561^{**}$	$r=.035$
	$p=.000$	$p=.798$
ULC	$r=.253$	/
	$p=.060$	

Note. The asterisks \*\* indicate correlation is significant at the 0.01 level.

Based on Table 5, three major findings can be summarized:

- (1) The scores of knowledge of lexical collocations (KLC) showed significant correlations with the subjects' general speaking proficiency (GSP).
- (2) The total frequency of use of lexical collocations (ULC) did not show any significant correlations with the subjects' general speaking proficiency (GSP).
- (3) The subjects' scores of knowledge of lexical collocations (KLC) did not show significant correlations with the subjects' use of lexical collocations (ULC).

Thus, the current study concludes that knowledge of lexical collocations seemed to be a more suitable indicator for determining Taiwanese college EFL learners' speaking proficiency than their Use of lexical collocations. In fact, ULC may not be a good indicator for measuring either the subjects' KLC or their general speaking proficiency. Findings of the current study are summarized and visually illustrated in Figure 2.

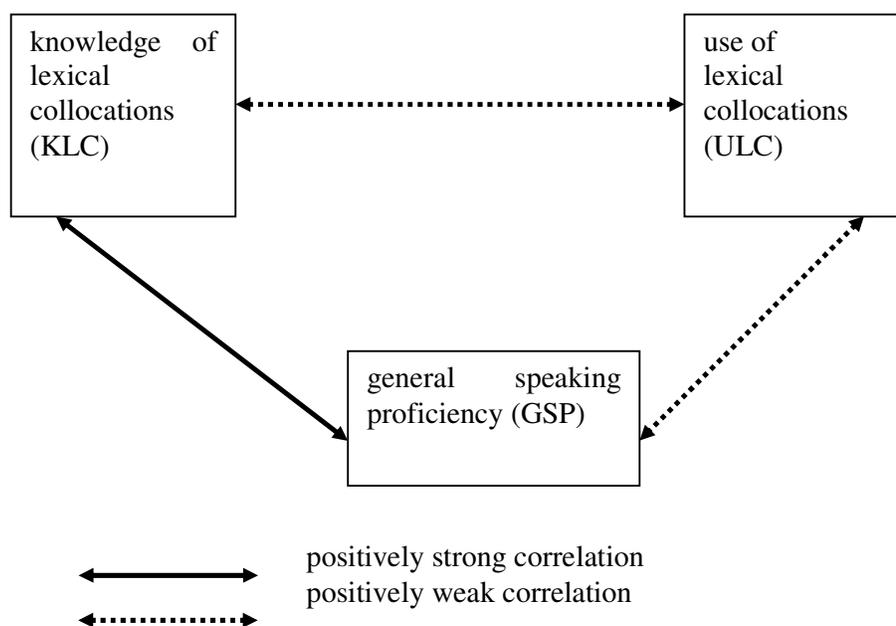


Figure 2. Strength of Correlations among KLC, ULC, and GSP

## 5. Concluding Remarks

### 5.1 Pedagogical Implications

The study has demonstrated that knowledge of lexical collocations is significantly associated with Taiwanese university EFL learners' spoken English. Future bigger scale studies might still need to be carried out in order to examine whether or not direct instruction on lexical collocations will benefit EFL learners' speaking proficiency. Nevertheless, a cautious way to deal with lexical collocations can be the introduction of a lexis-focused syllabus (Willis, 1990; Lewis, 1993, 1997) as a supplementary component to any current existing syllabus. In other words, teachers can at least try to provide EFL learners sources of lexis, such as corpus, concordance, and dictionaries of collocations, ensuring these learners to have direct exposure to chunks of English language and learn to extract and use patterns of lexical collocations effortlessly.

The next pedagogical implication concerns the subtype of lexical collocations for Taiwanese EFL learners. The study showed that the subjects scored lowest in the L4 subtype (adv + adj) in their lexical collocation test. None of the subjects ever produced the L4 subtype of lexical collocations (adv + adj) in their speaking test of Pear Film. In other

words, Taiwanese university EFL students might not have gained much knowledge of the L4 subtype of lexical collocations and still have not mastered well the *adverb* usage in their oral production. Based on such a finding, EFL teachers should especially concentrate on the practice of this certain subtype (adv + adj) in class.

### ***5.2 Limitations of the Study***

The fact that the subjects of the study were not randomly selected is perhaps one obvious limitation. The subjects were limited to the English-major junior students at a national university of science and technology from southern Taiwan. Therefore, while the study reveals interesting findings about the relationship between the Taiwanese EFL university learners' lexical collocational knowledge and their speaking proficiency, these findings may not be easily generalized beyond the subjects of the study.

The reliability of self-design lexical collocation test is another possible concern as well. Since the subjects were only tested once on their knowledge of lexical collocations, the results of their performance may not be reliable enough to truly indicate their real scores. In order to minimize the degree of performance errors, future researchers may also need to consider enhancing instrument reliability.

Another limitation of this study is related to the scope of collocations investigated. This study only examined lexical collocations and did not include grammatical collocations. As a result, the performance of the subjects on the collocational test only reveals their knowledge of lexical collocations, but does not represent their overall collocational knowledge.

### ***5.3 Recommendations for Future Research***

Collocations, an important focus in EFL research, have started to gain increasing attention in the past two decades. As the current study investigated the subjects' knowledge and use of collocations as well as their speaking proficiency, several recommendations may be made for future research.

First, a strong relationship was found between the knowledge of lexical collocations and speaking proficiency among the Taiwanese university EFL learners in the study. However, the subjects' use of lexical collocations did not show significant correlation with either knowledge of lexical collocations or their speaking proficiency in the study. Thus, Taiwanese EFL learners' use of collocations could be further examined. In the current study, the researchers used only one type of speaking test in the study to elicit the

subjects' use of lexical collocations. Future studies could consider using different discourse tasks for eliciting the subjects' spoken data, such as face-to-face conversations, interviews, and speeches. Second, the current study aimed at the subjects' knowledge and use of collocations in relation to their speaking proficiency. Future research could be extended to explore the relationship between knowledge of collocations and other language skills, including listening, reading, and writing. It could provide a better understanding of the connection between Taiwanese EFL learners' collocational knowledge and their general English proficiency. Lastly, the study also found Taiwanese EFL students performed differently on the five major categories of lexical collocations. The students performed the worst on the L4 subtype (adv + adj) both in their knowledge and use of lexical collocations. Therefore, it could be worthwhile to further investigate the patterns of development in the EFL learners' knowledge and use of lexical collocations.

The above are some potential directions this current study can recommend for future research on collocations. In fact, this study has stirred up more questions than it can possibly answer. The role of collocations in second language acquisition and teaching is not yet being fully understood. Much still needs to be done, particularly on the possible relationship between the acquisition of lexical collocations and EFL learners' language skills. It would take a lot of joint efforts from the fields of first and language acquisition, EFL/ESL, and applied linguistics to uncover.

#### **Endnotes:**

1. The present study focuses exclusively on lexical collocations, rather than grammatical collocations. Henceforth, throughout this study *lexical collocations* and *collocations* will be used interchangeably.
2. L. Hsu (2005) refers to Li-chen Hsu; Hsu (2002) is used for Jeng-yih Hsu.
3. Both Zhang (1993) and Hsu (2002) convert their subjects' length of written/spoken data for a possible comparison even though this could be a compromise for accurate data collection.
4. Henceforth, the term *Pearson correlation* will be used in the study.
5. In the spoken data obtained from Per Film test, no L4 subtype lexical collocation was used.
6. The spoken data obtained from Per Film test for ULC on L4 subtype was zero.

#### **References**

- Aghbar, A. A. (1990 October). *Fixed expressions in written texts: Implications for assessing writing sophistication*. Paper presented at a meeting of the English

Association of Pennsylvania State System Universities (ERIC Document Reproduction Service No. 352 808).

- Aghbar, A. A. & Tang, H. (1991). *Partial credit scoring of cloze type items*. Paper presented at 1991 Language Testing Research Colloquium, Education Testing Service, Princeton, NJ, 1-24.
- Al-Zahrani, M. S. (1998). *Knowledge of English lexical collocations among male Saudi college students majoring in English at a Saudi university*. Unpublished doctoral dissertation, Indiana University of Pennsylvania, Pennsylvania.
- Aston, G. (1995). Corpora in language pedagogy: Matching theory and practice. In G. Cook & B. Seidlhofer (Eds.), *Principle and practice in applied linguistics: Studies in honour of H. G. Widdowson* (pp. 257-270). Oxford: Oxford University Press.
- Bachman, L. F. (2004). *Statistical analysis for language assessment*. Cambridge: Cambridge University Press.
- Benson, M., Benson, E., & Ilson, R. (1997). *The BBI dictionary of English word combinations*. Amsterdam: John Benjamins.
- Brown, D. (1974). Advanced vocabulary teaching: The problem of collocation. *RELC Journal*, 5(2), 1-11.
- Chafe, W. (Ed.). (1980). *The pear stories: Cognitive, cultural, and linguistic aspects of narrative production*. Volume III in the series. Norwood, NJ: Ablex Publishing Co.
- Channell, J. (1981). Applying semantic theory to vocabulary teaching. *English Language Teaching Journal*, 35, 115-122.
- Chen, P. C. (2002). *A corpus-based study of the collocational errors in the writings of the EFL learners in Taiwan*. Unpublished master thesis, National Taiwan Normal University, Taiwan.
- Ellis, N. C. & Schmidt, R. (1997). Morphology and longer distance dependencies: Laboratory research illuminating the A in SLA. *Studies in Second Language Acquisition*, 19, 145-171.
- Fillmore, J. C. (1979). On fluency. In C. J. Fillmore, D. Kempler, & W. Wang (Eds.), *Individual differences in language ability and behavior* (pp. 85-101). New York: Academic Press.
- Gay, L. R. (1996). *Educational research: Competencies for analysis and application*. Englewood Cliffs, NJ: Merrill.
- Haslam, S. A. & McGarty, C. (2003). *Research methods and statistics in psychology*. London: Sage Publications Ltd.

- Huang, L. S. (2001). Knowledge of English collocations: An analysis of Taiwanese EFL learners. *Texas papers in foreign language education: Selected proceedings from the Texas foreign language education conference 2001*, 6(1), 113-129. (ERIC Document Reproduction Service No. 465 288).
- Hsu, J. Y. (2002). *Development of collocational proficiency in a workshop on English for general business purposes for Taiwanese college students*. Unpublished doctoral dissertation, Indiana University of Pennsylvania, Pennsylvania.
- Hsu, L. C. (2005). *The effect of lexical collocation instruction on Taiwanese college EFL learners' listening comprehension*. Unpublished master thesis, National Kaohsiung First University of Science and Technology, Taiwan.
- Hsueh, S. C. (2003). *An analysis of lexical collocational errors in the English compositions of junior high school EFL students*. Unpublished master thesis, National Kaohsiung Normal University, Taiwan.
- Kjellmer, G. (1991). A mint of phrases. In K. Aijmer & B. Altenberg (Eds.), *English corpus linguistics: Studies in honour of Jan Svartvik* (pp. 111-127). Harlow, Essex: Longman.
- Lewis, M. (1993). *The lexical approach: The state of ELT and a way forward*. London: Language Teaching Publications.
- Lewis, M. (1997). *Implementing the lexical approach: Putting theory into practice*. London: Language Teaching Publications.
- Lewis, M. (2000). *Teaching collocation: Further developments in the lexical approach*. London: Language Teaching Publications.
- Lien, H. Y. (2003). *The effect of collocation instruction on the reading comprehension of Taiwanese college students*. Unpublished doctoral dissertation, Indiana University of Pennsylvania, Pennsylvania.
- Lin, Y. P. (2002). *The effects of collocation instruction on English vocabulary development of junior high school students in Taiwan*. Unpublished master thesis, National Kaohsiung Normal University, Kaohsiung, Taiwan.
- Liu, C. P. (1999a). A study of Chinese culture university freshmen's collocational competence: "Knowledge" as an example. *Hwa Kang Journal of English Language & Literature*, 5, 81-99.
- Liu, C. P. (1999b). An analysis of collocational errors in EFL writings. *The Proceedings of the 8th International Symposium on English Teaching*, (pp. 483-494), Taipei: Crane.

- Liu, C. P. (2000). An empirical study of collocation teaching. *The Proceedings of the 17th International Symposium on English Teaching*, (pp. 165-178), Taipei: Crane.
- McArthur, T. (1992). *The Oxford companion to the English language*. Oxford: Oxford University Press.
- McCarthy, M. J. (1984). A new look at vocabulary in EFL. *Applied Linguistics*, 5(1), 12-22.
- McCarthy, M. (1990). *Vocabulary*. Oxford: Oxford University Press.
- Mertens, D. M. (1997). *Research methods in education and psychology: Integrating diversity with quantitative & qualitative approaches*. California: Sage Publications, Inc.
- Nation, I. S. P. (1990). *Teaching and learning vocabulary*. New York: Newbury House.
- Nation, I. S. P. (2001). *Learning vocabulary in another language*. Cambridge: Cambridge University Press.
- Nattinger, J. (1980). A lexical phrase grammar for ESL. *TESOL Quarterly*, 14(3), 337-344.
- Nattinger, J. & DeCarrico, J. (1992). *Lexical phrases and language teaching*. Oxford: Oxford University Press.
- Pawley, A. & Syder, F. H. (1983). Two puzzles for linguistic theory: Nativelike selection and nativelike fluency. In J. C. Richards & R. W. Schmidt (Eds.), *Language and communication* (pp. 191-226). New York: Longman.
- Richards, J. C. & Rogers, T. S. (2001). *Approaches and methods in language teaching* (2nd ed.). Cambridge: Cambridge University Press.
- Schmitt, N. (2000). *Vocabulary in language teaching*. Cambridge: Cambridge University Press.
- Scholfield, P. (1997). Vocabulary reference works in foreign language learning. In N. Schmitt & M. McCarthy (Eds.), *Vocabulary: Description, acquisition, and pedagogy* (pp. 279-302). Cambridge: Cambridge University Press.
- Schweigert, W. A. (1994). *Research methods & statistics for psychology*. Pacific Grove, California: Cole Publishing Company.
- Sung, J. (2003). *English lexical collocations and their relation to spoken fluency of adult non-native speakers*. Unpublished doctoral dissertation, Indiana University of Pennsylvania, Pennsylvania.
- Tang, Y. T. (2004). *A study of the collocation errors in the oral and written production of the college students in Taiwan*. Unpublished master thesis, National Taiwan Normal University, Taiwan.

- Tseng, F. P. (2002). *A study of the effects of collocation instruction on the collocational competence of junior high school students in Taiwan*. Unpublished master thesis, National Taiwan Normal University, Taiwan.
- Validation summary for PhonePass (2005). ORDINATE Corporation Technical Report.
- Weiss, N. A. (2005). *Introductory statistics*. Boston: Pearson Education, Inc.
- Willis, D. (1990). *The lexical syllabus: A new approach to language teaching*. London: Harper Collins Publishers.
- Wu, L. H. (2005). *A study of English verb-noun collocational knowledge of technological university English majors in Taiwan*. Unpublished master thesis, National Kaohsiung First University of Science and Technology, Taiwan.
- Wu, W. S. (1996). Lexical collocations: One way to make passive vocabulary active. *The Proceedings of the 11th Conference on English Teaching and Learning in the Republic of China*, (pp. 461-480). Taipei: Crane.
- Zhang, X. (1993). *English collocations and their effect on the writing of native and non-native college freshmen*. Unpublished doctoral dissertation, Indiana University of Pennsylvania, Pennsylvania.

### **Websites**

- The Pear Film*. Retrieved October 20, 2005, from  
<http://www.linguistics.ucsb.edu/faculty/chafe/index.htm>
- PhonePass score search*. Retrieved January 14, 2006, from  
<http://www.phonepass.com.tw/>
- Simple Search of British National Corpus*. Retrieved from January 20, 2006 to February 17, 2006, from <http://sara.natcorp.ox.ac.uk/lookup.html>
- VLC Web Concordancer*. Retrieved from January 20, 2006 to February 17, 2006, from  
<http://vlc.polyu.edu.hk/concordance/WWWConcappE.htm>



## **The Study of English Learners' Synthesizing Process While Reading**

**Lu-Fang Lin**

*National Taiwan Ocean University*

### **Bio Data:**

Lu-Fang Lin received a Ph. D. degree from the Department of Curriculum and Instruction, University of Victoria, B.C., Canada. She is currently an assistant professor in the Foreign Language Teaching and Research Center, National Taiwan Ocean University, Taiwan. She is involved in research into English teaching in the EFL context and English reading comprehending instruction.

### **Abstract**

This research investigated how English learners could retell two types of passages with culturally familiar and unfamiliar topics. The oral retelling was used to measure the fourteen teenage participants' reading comprehension of twelve English passages with Chinese and non-Chinese topics. The participants' retelling protocols were categorized into ways of synthesizing information. The research firstly examined whether there was difference of English learners' ways of synthesizing information between the two types of passages. The results showed that there was no significant difference between the participants' approach for the Chinese and non-Chinese topic passages. The participants appear to have displayed awareness of the macrostructure in a text and then combined some micro-propositions in the text to make a synthesized statement over culturally familiar and unfamiliar topic passages. Secondly, synthesizing information was classified by function and by strategy to explain how English learners utilized the synthesizing process to comprehend an English passage on Chinese versus non-Chinese topics. Thirdly, the process of how the participants utilized prior knowledge to produce synthesizing information was explained. In the last section of the paper, the researcher presented some recommendations for classroom practice in an effort to help language teachers apply the results of the study to the actual instructional contexts.

**Key Words:** English learners; synthesizing information; culturally familiar and unfamiliar topics; cross-cultural prior knowledge

## **Introduction**

A growing number of researchers in the field of reading comprehension have agreed with the concept that the reader's prior knowledge can affect the degree of text comprehension. Several earlier reading comprehension studies have confirmed that the knowledge a person has is understood to have a potential influence on what he or she will learn and read. Anderson, Reynolds, Schallert, and Goetz (1977) believed that "every act of comprehension involves one's knowledge of the world" (p.369). "A reader's knowledge determines to a large extent the meaning that he derives from a text" (Kintsch and van Dijk, 1978, p.371). Fagan (1987) prefaced his book with the following remarks about the essential role of knowledge in learning, "learning is the freedom and challenge to play with knowledge – to experience it, to reflect on it, to evaluate it, and to change it" (p. iii). Furthermore, a large number of empirical studies have demonstrated that the impact of prior knowledge on reading comprehension is significant (Afflerbach, 1990; Dochy, 1992; McKeown, Beck, Sinatra, & Loxterman, 1992). The lack of it might have an impact on text "reconstruction" in which the reader's comprehension of the text was minimal (Bernhardt, 1990). More recently, prior knowledge has been further proved to have great effects on the reader's retrieval of textual information (Caillies, Denhire, and Kintsch; 2002; Vacca, R. T., & Vacca, J. L., 2002); and can be a significant predictor of the student's achievement in the area of content reading (Yenilmez, Sungur, Takkaya, 2006).

With such an important role, prior knowledge is viewed as a key resource in the meaning construction of the reading process. The constructive orientation of the recent reading comprehension research has highlighted comprehension as a constructive process. With this aspect, reading can be regarded as "the act of constructing meaning while transacting with text" (Ruddell, 2005, p.30). A construction process in which "a textbase is constructed from the linguistic input as well as from the comprehender's knowledge base, with an integration phase, in which this textbase is integrated into a coherent whole" (Kintsch, 1988, p.53). The constructive process transacts with various sources, like readers' prior knowledge accumulated from previous life experience, the information conveyed in the text, and immediate social interaction and communication (Cambourne, 2002; Gee; 2001; Van Den Broek & Kremer, 2000). Before detailing the research design, however, the paper first turns to what the pertinent cognitive science tells us about the reader's cognitive

process involved in reading (Fagan, 1987; Kintsch, 1998), and the studies of prior knowledge in the context of second language (L2) reading comprehension.

### **Cognitive Reading Processes**

In the field of cognitive science, reading can be viewed as a literacy process inextricably connected with cognition (Ruddell, 2005). By cognitive reading processes, Fitzgerald (1995) explained that this refers to “any internal or mental aspects of reading” (p.146). The internal cognitive operations the reader engages in can be labeled variously in terms of different reading task demands and different levels of cognitive behavior. For example, as Fagan (1987) proposed, these processes included attending, analyzing, associating, predicting, inferring, synthesizing, generalizing, and monitoring and these processes might operate on various sizes of text units. Fagan (1987) explained that the operation of these processes depended on the reader purpose. However, all cognitive processes require knowledge (Kintsch, 1998). Prior knowledge will then be added as a factor influencing the operation of these cognitive processes.

### **Macrostructure Formation during Comprehension**

In a technical sense, Kintsch (2004) conceptualized comprehension as another paradigm for cognition. Relying on the support from several experimental studies, Kintsch (1998) claimed that macrostructure formation occurred as an integral part of comprehension (Kintsch & van Dijk, 1978; Guindon & Kintsch, 1984; Lorch, Lorch, & Mathews, 1985). According to Kintsch (1998), during the process of comprehension, a reader can select a macroproposition and delete several micropropositions. Thus, in forming a generalization, several micropropositions can be replaced by an appropriate macro proposition (Kintsch, 1998). For Kintsch, the formation of macroproposition is the process of reduction.

The present study applied Kintsch’s model of macrostructure formation to the field of L2 reading comprehension to examine whether L2 readers can as well produce such kind of macrostructures by reducing and replacing the micropropositions in the text. In addition to the reducing and replacing function of macrostructures, the current study further provides additional explanation about the functions of macrostructures in L2 reading comprehension.

### **Effectiveness of Prior Knowledge in L2 Reading Comprehension**

In the field of L2 reading comprehension research, the effectiveness of prior knowledge has also been investigated. In previous studies where the distinction between cultural specific prior knowledge and cross-cultural prior knowledge has been used as a key variable, the effects of cultural specific prior knowledge and global knowledge of the world still compete with each other. For example, research on the effects of content schemata held the perspective that L2 readers' culturally specific schemata might cause reading difficulty. Therefore, comprehension of a culturally unfamiliar text was more difficult than comprehension of a culturally familiar text (Anderson, 1994; Carrell, 1987; Pang, Muaka, Bernhardt, & Kamil, 2003; Steffensen, Joag-dev, & Anderson, 1979). On the other hand, readers' comprehension of text could be attributed to cross-cultural prior knowledge, which was not culturally bound but a global knowledge of the world. Some parts of this type of knowledge in some studies could be termed as subject knowledge or content knowledge which might as well, to some degree, facilitate L2 students' reading comprehension (Brantmeier, 2005, Bügel & Buunk, 1996; Hammadou, 2000; Young & Oxford, 1997).

So, although the importance of cross-cultural prior knowledge and the function of a synthesizing process (that is, the formation of macrostructures) have been understood, no empirical research in the Asian L2 context has yet explored the possible impact of these factors. Given the stakes involved, the current study examined whether differences would appear within specific disciplines in the process of L2 reading comprehension. In this study, the cognitive process the researcher dealt with was a synthesizing process, adopted from Fagan's term for the generative cognitive process. Synthesizing involved readers' "awareness of the whole" (Fagan, 1987, p.65). During the process of reading, readers tended to "search for a unity or integration within the information presented" (Fagan, 1987, p.65). The process of synthesizing was supposed to collapse specific information into different meaningful units and then integrate the individual units into one.

### **Restraints and Conflicts in Previous Research**

From reviewed literature on the role of prior knowledge in L2 reading comprehension, the research on synthesizing comprehension process, that is, the process of the reader's using prior knowledge to synthesize the information in the text, especially the cultural specific text, still needs more attention. In the past decades, most researchers

have tried to combine other variables to examine the effects of prior knowledge, such as text types, test types, the familiarity of topics, the participant's interests, gender or age. The cognitive process variable, the synthesizing process, has not been examined closely. Besides, Cohen, Rosenbaum-Cohen, Ferrara, & Fine's (1988) study that has investigated this issue showed the inconsistent results with those of Kintsch's (1998) macrostructure formation assumption. Cohen et al. (1988) used second-year students with different academic backgrounds at Hebrew University as participants. Cohen and his colleagues (1988) found that unlike the native English speakers, the nonnative English speakers did not attune to recognizing the conjunctive words, like *however*, *also*, *finally* and *thus*. Thus, they concluded that the nonnatives had more trouble synthesizing the information at the intra- and inter sentential levels as well as across paragraphs than the natives did. With such conflicting opinions, this current study examined English learners' synthesizing cognitive process to aid our understanding of EFL/ESL readers' comprehending the written materials.

### **Methodology Elaboration**

In this study, the researcher elaborated on the methodology used in previous research by Cohen et al. (1988) and in previous research on the issue of macrostructure. The research hoped to further provide factual accounts of FL/ESL/EFL readers' synthesizing process.

In the above experiments for measuring macrostructure comprehension (Kintsch & van Dijk, 1978; Guindon & Kintsch, 1984; Lorch, Lorch, & Mathews, 1985), these researchers focused on the recognition task to study the speed and accuracy with which reading times for topic and detail sentences were calculated, and words from topic and detail sentences were recognized. Beyond the recognition priming, the current study used the retelling technique to examine how L2 readers form macrostructures. As used to analyze readers' retellings, synthesized information at intra- and inter sentential levels might "[come] from more than one part of the passage" (Alberta Education, 1986, p. 44) and included synthesis of single words, clauses, phrases, or sentences. For a higher level of synthesizing information, the reader might reconstruct the author's words and ideas and produce synthesizing information across paragraphs.

The Cohen et al study (1988) used just one passage to explore the subject's

synthesizing process, but this is not enough. Also, to show the reader's dynamic development of reading process, the present study increased the number of the topics to prolong the period of data collection. The Cohen et al. study (1988) focused on the multifaceted nature of interactive L2 reading in English for science and technology; thus, the topic of text in their study was related to the participants' academic backgrounds. To generate a concept of English learners' general English reading, the present study added more topics that did not demand discipline-specific information.

### **Research Purpose and Research Questions**

In accordance with previous research's advocacy of the important role of the reader's prior knowledge in comprehension, the purpose of the research is first to examine the effects of prior knowledge on L2 readers' synthesizing process of the text with cultural specific topics (Chinese topics and non-Chinese topics). The study is also to explore how English learners apply their prior knowledge to comprehend English passages with Chinese and non-Chinese topics.

Three research questions were formulated to guide this study:

1. Is there a difference between English learners' synthesizing information while retelling passages with Chinese versus non-Chinese topics?
2. How do English learners utilize the synthesizing process to comprehend an English passage on Chinese versus non-Chinese topics?
3. How do English learners use prior knowledge to produce synthesizing information?

### **Methodology**

#### **Participants**

The participants in the study were from a senior high school in Taipei, Taiwan. Taipei is the capital of Taiwan and has a population of about two million. The Taiwanese senior high school generally includes Grade 10 to Grade 12. In Taiwan, Mandarin is the official language though other dialects are also spoken. The most commonly used foreign language is English. English is instructed as a subject in Taiwanese school curricula where Mandarin is the language of instruction.

In Taiwan there were few studies using high school students as participants and no standardized reading test was used to demonstrate high school students' English general reading ability. In this study, the researcher considered the students' cultural

background and made an adapted Informal Reading Inventory (IRI). The articles were separately adopted from the inventories by Bader (1998), Flynt and Cooter (1999), Leslie and Caldwell (2001), and Swearingen and Allen (2000). After each student was given an English reading test through the IRI, 14 Grade 11 senior high school students were selected from the volunteer pool of 97 to join the study. According to the results of the IRI, their English reading proficiency level was at the grade seven instructional level. The rationale for using this level of students as participants was that according to teachers' comments on this group of participants' general English ability, their English academic achievements were at the top ten from the highest scores in their class and they were able to and would be better able to express their own opinions. In order to collect sufficient verbal data, the researcher used this level of students as a beginning point to examine high school students' reading comprehension.

### **The Procedure of the Study**

The study began with a retelling practice session. The participants had not had the experience of doing retelling. In order to ensure that all participants would have the necessary abilities to retell the passage in Mandarin, a practice session was offered prior to eliciting the retelling protocols used for analysis. Once the performances of all participants in the retelling practice session were estimated to be satisfactory, each participant then joined the retelling meetings. The meeting was individual and each participant read and retold one passage in a meeting. After each retelling meeting, the researcher had an immediate interview with each participant to confirm some vague description in his/her retellings.

### **The Retelling Assessment Technique**

The researcher conducted a descriptive study to explore English learners' retelling response by using the following framework adapted from the Diagnosis Reading Program (DRP) by Alberta Education Student Evaluation (1986). This program was designed to provide teachers in Alberta, Canada with a systematic approach to observing and interpreting students' strengths and weaknesses in reading. To be noticeable, the task of retelling conducted in this study is slightly different from that of recall. The retelling technique encourages participants to retell the story in their own words. With such perspective, participants may be encouraged to restate the essential part of the original text, relate what they knew about the content of the text

and to reconstruct the information they have just read without looking at the passage again (McCormick & Cooper, 1991). The recall process usually involves having subjects recall, either in written or oral model, everything that they can remember of the text after his/her reading of the text by not looking back at the assigned reading material. Under such situation, “free recall tests may measure comprehension, but they also depend very heavily on memory capacity” (Bügel & Buunk, 1996, p. 18). Also, the scoring of the written recall protocols put much emphasis on checking the spelling and calculating the numbers of the correct ideas recalled from the text. To improve the recall task done in previous research (Bernhardt, 1990, 1991; Brantmeier, 2005; Carrell, 1984, 1987; Heinz, 2004; Regina, 1998), the instruction wording in this study was changed to the following, “After your reading, you will retell the content of the passage to a friend who has never heard or read it before”. In this study, the participant retold the passage orally but not in a written way.

### **Materials**

In this study, twelve passages were used as reading materials for the retellings. Six passages have topics on Chinese culture, describing some ancient Chinese historical events, figures, and customs. They include *Chinese Farming* (CF1), *Chinese New Year* (CNY3), *Dr. Sun Yat-Sen* (SYS5), *The Great Wall* (GW7), *The History of Tea* (HT9), and *Cooking and Eating* (CE11). The Chinese passages were selected from the following sources: *China: Our Pacific neighbour* (Evans & Yu, 1992), *Children of the world: China*, (Talan & Sherwood, 1988), and *Ancient China* (Sabin, 1985). The other six passages have topics on non-Chinese culture including Canadian and European historical events, peoples, and customs. They are *River of Salmon* (RS2), *Railway across Canada* (RC4), *First Peoples in Canada* (FPC6), *Easter* (EAS8), *Fishing in Canada* (FC10), and *Ways of Sending a Message* (WSM12). The Canadian passages were selected from, *Connections Canada* (Francis, 2000) and *Young Students Encyclopedia* (Blashfield, 1973). Owing to the participants being at a seventh-grade reading level, the readability level of the passages was at seventh-grade -- as determined with the assistance of the Flesch-Kincaid Grade Level Index in Microsoft Word. The statistical results of the 12 passages received from the Index are demonstrated in Table 1. This index is a computerized program used to determine the difficulty of books or reading material. By doing this, the participants at this level are supposed to be able to comprehend the text fitting their reading abilities.

Table 1 - Flesch-Kincaid Grade Level Indexes of the passages

Chinese passages (Flesch-Kincaid Grade Level Index)	Non-Chinese passages (Flesch-Kincaid Grade Level Index)
Chinese Farming (7.1)	River of Salmon (6.7)
Chinese New Year (7.0)	Railway in Canada (7.4)
Dr. Sun-Yat-Sen (7.2)	First Peoples (7.0)
Great Wall (7.2)	Easter (7.0)
History of Tea (6.8)	Fishing in Canada (7.2)
Cooking and Eating (7.4)	Sending the Message (7.2)

### **Scoring and Labeling the Participants' Retelling Protocols**

After the participant finished retelling, the researcher transcribed the recorded retelling in Mandarin and further translated it into English. Then the researcher adopted the DRP procedure for judging the students' retellings and divided the participants' retellings into smaller meaningful independent units called thought units. A thought unit is a group of words representing a syntactically grammatical and meaningful unit of information represented in a text or retold by the participants. For example, one simple sentence is regarded as one unit for it conveys a piece of meaningful information independently.

After the participant's retelling was divided into thought units, the parsed retelling information was further screened and labeled into synthesizing information (S). The thought units of this category were then summed up (see the sample of how to label the synthesizing retelling information in Appendix). To ensure the credibility of the analysis, the researcher asked another PH. D. student who was a native English speaker as a second rater.

The inter-rater agreement is established by randomly selecting ten of the 168 (14 participants x 12 topics) retelling protocols. It is a mean score of ten percentages of agreement. "The percentage of agreement equals the number of agreements divided by the total number of agreements and disagreements multiplied by 100:

"Percentage of agreement = [agreements/(agreements + disagreements)] x 100" (Martella, Nelson, Marchand-Martella, 1999, p.84). For example, the percentage of inter-rater agreement of synthesizing unit parsing is 91 %. The calculation would have been as follows: For one retelling protocol, 9 agreements/[9 agreements + 1 disagreements] with the result of .90 multiplied by 100. Similarly, the rest of the nine percentages were calculated in this way, summed up and divided by 10; the averaged inter-rater agreement for parsing the participant's retelling into synthesizing units

(91%) was thus achieved. 91% can be a high level of inter-rater agreement between the researcher and the second rater.

**Results and Discussion**  
***Synthesizing Information***

The statistic analysis in Tables 2 and 3 provided the answer to the first research question. Since the observations are paired in questions as shown in Tables 2 and 3, according to the result of the paired *t* test (the *t* value is equal 0.83 with 13 degrees of freedom, which is smaller than the two-tail critical value, 2.16), there is no significant difference in the participants' synthesizing information when retelling the passages with (on) Chinese and those with non-Chinese topics. The possible reason can be that the participants may have had the competence of an awareness of the macrostructures and then combine some information in the text to make a synthesized statement over the passage on culturally familiar and unfamiliar topic.

Table 2 - Means and Standard Deviations for Synthesizing Information

	Chinese						Means
	CF1	CNY3	SYS5	GW7	HT9	CE11	
Victoria	1	5	6	1	2	2	2.83
Amy	3	7	6	3	2	3	4.00
Lucy	2	1	3	2	3	4	2.50
Kathy	1	7	4	1	2	2	2.83
Jane	2	6	7	3	0	4	3.67
Sylvia	2	6	4	3	2	2	3.17
Miffy	1	4	2	2	1	3	2.17
Jim	1	2	4	3	1	1	2.00
George	3	3	4	1	0	1	2.00
Jeff	2	6	3	3	0	2	2.67
Brian	1	5	4	5	4	5	4.00
Dick	1	2	6	3	5	3	3.33
Tom	0	6	4	2	1	4	2.83
Kevin	1	5	3	2	2	1	2.33

Table 3 - Means and Standard Deviations for Synthesizing Information

	Non-Chinese						Means
	RS2	RC4	FPC6	EAS8	FCA10	WSM12	
Victoria	6	3	4	4	1	4	3.67
Amy	3	4	3	2	2	3	2.83
Lucy	3	1	1	2	1	3	1.83
Kathy	2	2	4	2	1	2	2.17
Jane	5	3	4	4	1	5	3.67
Sylvia	3	2	2	1	2	3	2.17
Miffy	6	1	1	3	2	4	2.83
Jim	3	2	1	3	1	0	1.67
George	4	2	5	3	2	2	3.00
Jeff	4	2	1	3	1	2	2.17
Brian	3	0	3	4	5	4	3.17
Dick	4	4	3	5	3	3	3.67
Tom	4	4	4	3	4	4	3.83
Kevin	2	2	0	1	0	2	1.17

The detailed interpretation of the possible reasons was described in the following section and at the same the interpretation was the respond to the second research question. Synthesizing information in the participant's retellings was classified by function (integrated, reconstructing, and reducing synthesizes), by strategy (synthesizes for opening a talk, filling the gap). Each classification was illustrated with the participant's retellings. The third research question was answered with the application of cross-cultural knowledge.

#### ***A process of integration***

As for what Fagan (1987) stated, the synthesizing retellings showed that the reader tended to seek for an integration among pieces of written information. In the current study, most participants generalized ideas from several sentences and produced a larger gist. That is, one synthesizing statement extracted the words directly from two or three sentences and interweaved another new statement. Such kind of synthesizing information from the participants' retelling for the first paragraph of the passage *Easter* was used to illustrate this process. For a clear explanation, the first paragraph of the original passage was presented and each sentence from the first one was

numbered from 1 (see Table 4). The participant's retellings were also marked with numbers that were the sentence numbers in the original passage. The data in Table 4 shows that most of their synthesizing information is at inter-sentential level connecting several parts from sentences in the paragraph.

Table 4 - The Original Passage and Participants' Retellings

---

First Paragraph of the Passage *Easter*:

- (1) Easter is always celebrated on a Sunday in early spring.
- (2) But it comes on a different date each year.
- (3) This is because it always takes place on the first Sunday after the first full moon after the spring equinox (March 21).
- (4) So Easter can come anytime from March 22 to April 25.

Participants' Retellings:

Easter probably comes in spring from March to April. (Victoria)

(1) + (4)

The first Sunday in spring each year. (Jane)

(3) + (1) + (2)

It comes on a different date each year, probably from March 22 to

(2) + (4)

April 25. (Miffy)

Easter comes on different dates every year because it is the first

(2) + (4) (3)

Sunday after the first full moon. (Jeff)

+ (3)

---

### ***A process of reconstruction***

Kintsch (1998) mentioned that in forming a generalization, several micropropositions could be replaced by an appropriate macroproposition. In this study, the participant reconstructed the meaning of the text by using his/her own words rather than the author's words. Such kind of synthesizing information usually conveys the essential meaning presented in sentences. For example, the participant Tom produced a statement for the first paragraph of the passage, *Easter*, "This is about the time when Easter comes each year."

Fagan (1987) proposed this type of synthesis involving the embedding of essential information from different syntactic units. An example of such kind of synthesis is cited from Fagan's (1987) book (p.66):

Text: Mr. Smith owns a German Shepherd. It is black. One day it broke its leash. It ran down the street. It attacked a small boy.

Synthesis: One day Mr. Smith's black German Shepherd broke its leash, ran down the street and attacked a small boy.

Fagan (1987) explained that it could not be easy for a reader to undergo such process of synthesizing at this level in an impromptu situation. "The greater the number of structures from which information must be embedded, the more difficult the task" (Frase, 1970, cited from Fagan, 1987, p.66). This finding suggests that for English learners at the age of 16 with Grade 7 English reading abilities, a higher level of synthesis can be achieved.

### ***A process of deletion***

As Kintsch (1998) pointed out, macrostructure could be formed by the deletion of insignificant details. In this study, the participant retold a generalized statement by reducing some minor details, especially those with unfamiliar vocabulary. The examples withdrawn from Brian's and Dick's retellings of the Easter passage are as follows:

It probably comes after March 21, the first full moon, then probably between at the end of March and at the end of April, one Sunday is Easter.  
(Brian)

Then Easter usually comes on a different each year because it takes place on the first Sunday after the first full moon after March 21. (Dick)

From the examples, the participants summarized a segment of the text about the fact when Easter comes every year. In the interview, the participants expressed that they did not use much of their life experience to understand this part because they did not have this holiday in Taiwan. They stated that they had no idea about the words *the spring equinox*, so they skipped retelling the segment with unknown vocabulary in it, and thus made a generalized statement for the paragraph with their general knowledge of date sequence. This situation can be explained by Kintsch's (1988) model of text recall.

[A] subject selects a path through the network that forms the mental representation of the text. Those nodes that are not on that path are automatically deleted when recalling the text (cited from Kintsch, 1993, p. 196).

***A pragmatic strategy of opening a talk and filling up the gap***

After reading a passage, the participant was to retell the passage. In the beginning section of most retelling meetings, there was a period of silence. Most participants usually prefaced their retelling with a short over-generalization for the whole passage, such as:

This article talks about Easter. (Brian)

This is about First peoples in Canada. (Jim)

Dr. Sun-Yat Sen is a great man. This article is about him. (Jane)

As shown in the above, the underlined words were the topics of the passages. Most participants directly used the topic to produced such synthesizes. These broad over-generalizations were still categorized as synthesizing information in this study although they were a much different synthesis that did not exactly summarize main ideas from the original text. Such over-generalizations could be a strategy they used to opening their talk.

Besides, this format could be a strategy to fill up the blank during their retelling or also when they had trouble understanding the content. For example, when facing the trouble of retelling the third paragraph of the passage *Easter*, some participants with high motivation would try to solve the silent, embarrassing moment and made some response. Their resolutions were to quickly squeeze some vague generalizations, such as the following retelling:

The whole week, people do these things. (Tom)

From Monday to Friday, there are lots of activities. (Kevin)

Christians did some activities in a week. (Victoria)

In the interview, the researcher asked Tom what these things were then. He answered that he had no idea about that. Another participant Victoria explained that she could only recognize the words, *Christ, week, Thursday, Friday*, so she combined

these words and made such response. However, she had no exact idea about these activities before the holiday, Easter.

### ***Utilization of cross-cultural prior knowledge***

In this study, the participant expressed that they did not have much of prior knowledge about non-Chinese topics, so they mostly could not retell many detailed contents of the passage but they could retell synthesizing information.

For Taiwanese students in this study, their cross-cultural prior knowledge may include knowledge shared between Taiwan, Canada and some European countries. For example, in the interview, most participants indicated that they did not have much prior knowledge about the first peoples in Canada but they knew what aboriginal people were. The following synthesizing information from the participants' retelling protocols demonstrated Taiwanese students' applying cross-cultural prior knowledge to retell the passage *First Peoples in Canada*:

The aboriginal people make a living by hunting in the woods, and fishing in the lake. (Amy)

The aboriginal people use the wood to build a house. (Tom)

The aboriginal people make clothes from animal skins or fur. (Jeff)

One possible reason for this is that there are several groups of aboriginal people residing in Taiwan. With the government's advocacy of protecting the culture of Taiwanese aborigines and the study of local culture in the course, *Social Science*, the students had been educated about aboriginal cultures and learn general knowledge of how aboriginal peoples make a living. Some concepts included in the passage *First Peoples in Canada* were related to Taiwanese students' cross-cultural prior knowledge about the aboriginal peoples in Taiwan. As a result, Taiwanese students could make use of their cross-cultural prior knowledge in this area to embed a synthesized statement from this passage. This finding suggests that cross-cultural knowledge assists the reader to search for potential relevant information in memory and thus retell synthesizing information for the passages on Chinese and non-Chinese topics.

### **Summary**

After discussing different phenomena of the participants' retelling synthesizing information, general findings can be summed up to show that the production of

synthesizing information may primarily depend on the participant's prior knowledge. If the participant lacks culturally specific knowledge about the text, he/she may further rely on his/her cross-cultural prior knowledge and thus synthesize information in an ambiguous, generalized way. In this study, the passages with non-Chinese topics may include several messages related to culturally specific prior knowledge and also, in some part, to cross-cultural prior knowledge. When retelling the passages with non-Chinese topics, synthesizing information may occur as often as in passages with Chinese topics. Therefore, there is no difference found in synthesizing information between the two types of passages. Consistent with Bernhardt's opinion (1990), this study also finds that if the participant has neither culturally specific knowledge nor cross-cultural prior knowledge, most of the message in the text cannot be synthesized.

### **Recommendations**

In this study, the general finding is that the participants may have the competence of an awareness of the whole and then combine some information in the text to make a synthesized statement for culturally familiar and unfamiliar topics. It can be noted that the retold synthesizing information is directly related to the participants' ability to capture some information in the text and reconstruct the relations between the information. Given the participant's process of synthesizing in reading comprehension, some practical recommendations for classroom practice were suggested.

First, Kitsch (1998) stated that "a well-organized macrostructure is crucial for understanding and remembering a text". In this study, most participants could provide a synthesized topic statement in the beginning of their retellings. This result recommends that the instructor notify the students the text with a topic or a paragraph with a topic sentence. The topic may provide a direction for the reader to retrieve his/her prior knowledge. This finding can also suggest that the instructor may lead the students to read a passage without a topic first and ask them to assign a topic for the passage they have just read.

Second, the study finds that the synthesizing information integrates main ideas from several sentences. The teaching activity can be that after learners finish reading a passage, the instructor asks them to figure out the essential parts in the passage and use the following patterns to lead them to describe the generalized concepts of the segments of a text, such as, "The couple of sentences is about yyy," "This paragraph

is about xxx,” or “The whole passage is about XXX.” These patterns are designed to train English learners to integrate ideas from fewer and smaller proposition in the text to form larger proposition.

Third, the result of this study shows that the participants still can do well in synthesizing the information from the English passages on non-Chinese topics. The result recommends that except the familiar topics, the teacher can lead the student to read a passage on unfamiliar topic to produce synthesizing retellings as long as the readability of the passage fits English learners’ English reading ability. Confronting the passage on unfamiliar topics, students can be activated to identify main ideas from details and further delete minor details to generate macrostructural statements for the passage. By doing this, students can be trained to focus on macropropositions in the passage rather than on each single word.

Fourth, the general task of synthesizing information may include the ability to give summary statements, which adequately reflect the essence of small parts in a text. The summarizing procedure essentially involves a series of deletions and generalizations (Farstrup & Samuels, 2002). Summary writing can be suggested. In this study, the researcher recommended another way of summary writing. Teachers may also focus on the summary retelling activity for collapsing a whole paragraph into smaller meaningful chunks, pointing out essential features in each chunk, and then asking learners to integrate the essential features in larger synthesizing statements. The teacher may encourage the policy that the number of the sentence statements is the fewer, the better in terms of making use of all marked essential parts. To encourage learners to produce sufficient retellings, sentence grammatical structure is not highly emphasized here. The following is an example of the teaching material for activating synthesizing process provided from the researcher’s classroom instruction:

Text: Some things around us live. . . . Things that live need air. Things that live need food. Things that live need water. Things that live move and grow. Animals are living things. Plants are living things (Leslie & Caldwell, 2001, p. 155).

Synthesis: Things that live around us need air, food and water. Some living things can move and grow, like animals and plants.

Fifth, the results of this study illuminate the importance of English learners’ cross-cultural knowledge. Prior to this study, most Taiwanese senior high school

students to some degree have studied several subjects such as Math, History, Geography, Science and Chinese Language Arts in school. Some of them can be academically proficient and have learned some knowledge in these subjects. To some degree, the subject knowledge is cross-culture. However, the subject knowledge is in Mandarin. The instructor may try to transfer their Mandarin prior knowledge into English prior knowledge while reading a new English passage. The instructor encourages the reader to make use of his/her cross-cultural knowledge to comprehend the text.

### **Conclusion**

In this study, this group of teenagers has proved that they have the ability to retell synthesizing information over familiar and unfamiliar topic passages with the assistance of their prior knowledge. The general result from this study was consistent with Kintsch's (1998) claim that macrostructure formation occurred as an integral part of comprehension. Moreover, the results of the study provide further evidence in the field of prior knowledge studies to ensure the essential impact of the cross-cultural knowledge (Brantmeier, 2005; Hammadou, 2000). More than that, the findings of the study suggest that the reader's cross-cultural knowledge can facilitate English learners to operate a synthesizing process.

In contrast, the result of the study was not in agreement with Cohen et al's (1988) conclusion that non-natives had more trouble synthesizing the information at the intra- and inter sentential levels as well as across paragraphs than natives. The participants in the Cohen et al. study were second-year university students from four different departments in the Hebrew University. Cohen et al (1988) pointed out that they were fairly good readers in English. The result of their research demonstrated the situation when the proficient older English learners read in specialized English. Noticeably, this current study addressed the situation when 16-year-old high school English learners whose English reading proficiency level was grade seven do general English reading. The different result from this study and the classification of synthesizing information can add new knowledge to the field of English learners' cognitive reading process. In the near future, the researcher will include other groups of English learners with different levels of English reading abilities to further examine the nonnatives' synthesizing process via culturally

specific and cross-cultural topic passages.

## **References**

- Afflerbach, P. P. (1990). The influence of prior knowledge on expert readers' main idea construction strategies. *Reading Research Quarterly, 25*, 31-46.
- Alberta Education (1986). *Diagnostic Reading Program*. Alberta, Edmonton: Minister of Education.
- Anderson, R. C. (1994). Role of the reader's schema in comprehension, learning, and memory. In R. B. Ruddell, M. R. Ruddell, & H. Singer (Eds.), *Theoretical models and processes of reading* (4th ed., pp. 469-482). Newark, Delaware: International Reading Association.
- Anderson, R. C., Reynolds, R.E., Schallert, D. L., & Goetz, E. T. (1977). Frameworks for comprehending discourse. *American Educational Research Journal, 14*, 367-381.
- Bader, L. A. (1998). *Bader Reading and Language Inventory*. NJ: Prentice-Hall, Inc.
- Bernhardt, E. B. (1990). A model of L2 text comprehension: The recall of literary text by learners of German. In A. Labarca & L. M. Mailey (Eds.), *Issues in L2: Theory as practice/practice as theory* (pp.21-43). Norwood, NJ: Ablex.
- Bernhardt, E. B. (1991). *Reading development in a second-language*. Norwood, NJ: Ablex.
- Blashfield, J. E. (Ed.). (1973). Easter. In *Young students encyclopedia* (Vol.7, p. 955). New York: Funks & Wagnalls.
- Brantmeier, C. (2005). Effects of reader's knowledge, text type, and test type on L1 and L2 reading comprehension in Spanish. *The Modern Language Journal, 89*(1), 37-53.
- Bügel, K., & Buunk, B. P. (1996). Sex differences in foreign language text comprehension: The role of interests and prior knowledge, *The Modern Language Journal, 80*(1), 15-31.
- Cailies, S., & Denhière, G., & Kintsch, W. (2002). The effect of prior knowledge on understanding from text: Evidence from primed recognition, *European Journal of Cognitive Psychology, 14*(2), 267-286.
- Cambourne, B. (2002). Holistic, integrated approaches to reading and language arts

- instruction: The constructive framework for an instructional theory. In A. E. Farstrup & S. J. Samuels (Eds.), *What research has to say about reading instruction* (pp. 25-47). Newark, DE: International Reading Association.
- Carrell, P. L. (1984). Schema theory and ESL reading: Classroom implications and applications. *The Modern Language Journal*, 68(4), 332-343.
- Carrell, P. L. (1987). Content and formal schemata in ESL reading. *TESOL Quarterly*, 21(3), 461-481.
- Cohen, H. G., Rosenbaum-Cohen, P. R., Ferrara, J., & Fine, J. (1988). Reading English for specialized purposes: discourse analysis and the use of student informants. In P. L. Carrell, J. Devine, & D. E. Eskey (Eds.), *Interactive approaches to second language reading* (pp. 152-167). Cambridge: Cambridge University Press.
- Dochy, F. J. R. C. (1992). *Assessment of prior knowledge as a determinant for future learning*. London: Jessica Kingsley Publications.
- Evans, B., & Yu, A. (1992). *China: Our pacific neighbour*. Edmonton: Reidmore Books.
- Fagan, W. T. (1987). *The learning and teaching of reading: An interactive process model*. Edmonton: University of Alberta.
- Farstrup, A. E., & Samuels, S. J. (2002). *What research has to say about reading instruction*. Newark, DE: International Reading Association.
- Flynt, E. S., & Cooter, R. B. (1999). *Englis-Espanol Reading Inventory for the Classroom*. NJ: Prentice-Hall, Inc.
- Fitzgerald, J. (1995). English-as-a-second-language learners' cognitive reading processes: A review of research in the United States. *Review of Educational Research*, 65(2), 145-190.
- Francis, D. (2000). *Connections Canada*. Canada: Oxford University Press.
- Gee, J. P. (2001). Reading as situation language: A sociocultural perspective. *Journal of Adolescent and Adult Literacy*, 44(8), 714-725.
- Guindon, R., & Kintsch, W. (1984). Priming macropropositions: Evidence for the primacy of macropropositions in memory. *Journal of Verbal Learning and Verbal Behavior*, 23, 501-18.
- Hammadou, J. (2000). The impact of analogy and content knowledge on reading comprehension: What helps, what hurts. *The Modern Language Journal*, 84, 38-50.
- Heinz, P. J. (2004). Towards enhanced second language reading comprehension

- assessment: Computerized versus manual scoring of written recall protocols. *Reading in a Foreign Language, 16*(2).
- Kintsch, W. (1988). The role of knowledge in discourse comprehension: A construction-integration model. *Psychological Review, 95*(2), 163-182.
- Kintsch, W. (1993). Information Accretion and reduction in text processing: Inferences. *Discourse Processes, 16*, 193-202.
- Kintsch, W. (1998). Modeling comprehension processes: The construction-integration model. In W. Kintsch, *Comprehension: A paradigm for cognition* (pp. 91-120). Cambridge: Cambridge University Press.
- Kintsch, W. (2004). The construction-Integration model of text comprehension and its implications for instruction. In R. B. Ruddell, & N. J. Unrau (Eds.), *Theoretical models and processes of reading* (pp. 1270-1328). Newark, DE: International Reading Association.
- Kintsch, W., & van Dijk, T. (1978). Toward a model of text comprehension and production. *Psychological Review, 85*, 363-394.
- Leslie, L. & Caldwell, J. (2001). QRI-3. New York: Longman.
- Lorch, R. F., Lorch, E. P., Mathews, P. D. (1985). On-line processing of the topic structure of a text. *Journal of memory and language, 24*, 350-62.
- Martella, R. C., Nelson, J. R., & Marchand-Martella, N. E. (1999). *Research method: Learning to become a critical research consumer*. MA: Allyn and Bacon.
- McCormick, S., & Cooper, J. O. (1991). Can SQ3R facilitate secondary learning disabled students' literal comprehension of expository text? Three experiments. *Reading Psychology: An International Quarterly, 12*, 239-271.
- McKeown, M. G., Beck, I. L., Sinatra, G. M., & Loxterman, J. A. (1992). The contribution of prior knowledge and coherent text to comprehension, *Reading Research Quarterly, 27*, 79-93.
- Pang, E. S., Muaka, A., Bernhardt, E. B., & Kamil, M. (2003). Teaching reading: Educational practices series-12. UNESCO: the International Academy of Education.
- Regina, R. (1998). *Reading and recall in L1 and L2: A sociocultural approach*. Stamford, CT: Ablex.
- Ruddell, M. R. (2005). *Teaching content reading and writing*. NJ: John Wiley & Sons.
- Sabin, L. (1985). *Ancient China*. Mahwah, NJ: Troll Associates.

- Steffensen, M. S., Joag-dev, C., & Anderson, R. C. (1979). A cross-cultural perspective on reading comprehension. *Reading Research Quarterly, 15*, 10-29.
- Swearingen, R. & Allen, D. (2000). *Classroom assessment of reading processes*. Boston: Houghton Mifflin Company.
- Talan, S., & Sherwood, R. (Eds.). (1988). *Children of the world: China*. Wisconsin, USA: Gareth Stevens, Inc.
- Vacca, R. T., & Vacca, J. L. (2002). *Content area reading* (7<sup>th</sup> ed.). Boston: Allyn & Bacon.
- Van Den Broek, P., & Kremer, K. (2000). The mind in action: What it means to comprehend during reading. In B. M. Taylor, M. F. Graves, & P. Van Den Broek (Eds.), *Reading for meaning: Fostering comprehension in middle grades* (pp. 1-31). Newark, DE: International Reading Association.
- Yenilmez, A., Sungur, S., & Takkaya, C. (2006). Students' achievement in relation to reasoning ability, prior knowledge and gender. *Research in Science & Technological Education, 24*(1), 129-138.
- Young, D. J., & Oxford, R. (1997). A gender-related analysis of strategies used to process input in the native language and a foreign language. *Applied Language Learning, 8*, 43-73.

**Appendix**

*Sample for Categorizing the Synthesizing Retelling Information.*

This is about the story of Dr. Sun Yat-Sen. (S) / [He was a], He came out to save China when the government in China was in the weakest condition. (S) / He wanted to change the government in China, [the emperor, that is, to abolish the monarchy / and used the western republic government system. (S) / [Then] he caused a revolution / and needed money so he asked for the support from the people overseas / and he had been to Canada for three times to raise money. / [Then he,] He in 1911, finally succeeded, [succeeded] in the revolution / and set up a new republic government. (S) / It is a pity that he did not live long enough to lead the new Chinese government. / [Then] the result is that the warlord, Yuan Shih-Kai succeeded him and turned to be a leader to lead the whole China. / [Then] warlords are those who own armies and control them. / [Then] warlords themselves did not understand Dr. Sun-Yat's idealization. (S) / [Then] they destroyed Chinese people's lives / and invaded their land. / People did not have enough food. / The people had little hope for the future/

where S = 6

Note: Parentheses were used around mazes. A maze consists of "irrelevant information", such as noises (*er, um, or uh*), self-corrections (*then she asked her tech, teacher*), repetitions of words (*when, when . . . and he, and then he*), and personal comments about the passage (*That's all, I don't remember any more, or I like that story*) (Alberta Education, 1986, p.42).



## **An Economical Approach towards Interaction in the L2 Classroom: A Task-based Learning Experiment**

**Nilton Hitotuzi**

*Secretaria de Educação do Estado do Amazonas (SEDUC-AM)\**

*Fundação de Amparo à Pesquisa do Estado do Amazonas (FAPEAM)*

### **Bio Data:**

Nilton Hitotuzi holds an MA in TEFL/TESL from the University of Birmingham and a BA in Liberal Arts from the Federal University of Amazonas. He has taught General English and Composition at the Federal University of Amazonas, and General English at the State University of Amazonas. Currently he is on sabbatical from the Amazonas State Department of Education, studying for a doctorate in TEFL at the Federal University of Bahia. He is interested in teaching methodologies, formal pronunciation teaching, process drama, and classroom research.

### **Abstract**

This paper reports on a piece of classroom research involving a group of Liberal Arts/TEFL undergraduates from the Federal University of Amazonas. Informed by Jane Willis's framework for task-based language learning and Michael Breen's insights into the involvement of learners in the evaluation of learning task cycles, a unit of study was designed and implemented to experiment with clustered tasks as a means of maintaining peer-peer oral/aural interaction in the classroom at substantial levels. While the results indicate that Breen's suggestion is effective in keeping learners engaged in meaningful interactions in the classroom for an extended period of time, it is still only intuitively established that Willis's framework is an adequate way of dealing with the focus-on-form versus focus-on-use dilemma in the second/foreign language classroom. A key assumption underlying the experiment is that the longer learners use the target language to communicate in the classroom the more their interlanguage is enhanced. Furthermore, it is suggested that this analytical approach can be an alternative to the task-repetition approach proposed by Martin Bygate.

**Keywords:** analysis of learning tasks, clustered tasks, meaningful interaction, enhancement of interlanguage

\* Funding for this project was provided by SEDUC-AM and FAPEAM.

[Teaching is] the purposeful creation of situations from which motivated learners should not be able to escape without learning or developing.

Cowan, 1998, p. 112

## **1. Introduction**

In a previous paper, I reported the results of a study as an attempt to probe for the actual teacher talking time versus student talking time in my foreign/second language (L2) classrooms. And, one of the groups investigated was attending an undergraduate Liberal Arts-TEFL course at the Federal University of Amazonas (UFAM). The results revealed that 85 per cent of the total of classroom talk in the sessions recorded with this intermediate-level group was done by the students (Hitotuzi, 2005, p. 103). Nonetheless, behind these figures there is painstaking and time-consuming preparation of a considerable amount of small activities handed over to the students in order to keep them talking.

Working on a large variety of short activities may seem interesting from the point of view of the students; but it is quite an overload of work for the teacher who, depending on his or her teaching context, may have to cope with several numerous groups in different educational institutions. Thus, no sooner had I confirmed empirically that, in terms of talking time, there was a comfortable level of learner empowerment in that intermediate group than my classroom-management style began posing this somewhat challenging problem. It was obvious to me though that whatever the solution to this deadlock, it had to do with parsimony. That is, the number of small activities had to be reduced without compromising student talking time. Grounded in this hypothesis, I decide to design and implement a tentative unit of work capitalising basically on Willis's (1996) framework for task-based learning and Breen's (1989) evaluation of learning task cycles.

Based on the unit work plan, the students were required to go through a series of integrated micro task cycles that constituted the building blocks of a macro task cycle. The macro task, in turn, consisted of the evaluation of the activities comprising the whole of the micro-task frameworks. Nevertheless, while some specific learning aims of the activities conducted through this task-within-a-task model were outlined in terms of grammar, lexis and pronunciation, there were no expectations as to what the group would exactly learn from them, since, as many language-oriented theorists and practitioners suggest, the process of learning an L2 does not seem to be linear (e.g. in the same order as the teacher presents it in the classroom), or cumulative (Corder, 1967; Selinker, 1972; Rutherford, 1987; Ellis, 2003 and others). It was expected, however, that to some extent the unit would contribute to the development of the students' target-language system (Breen, 1989). This could be observed through the comparison between their performances in writing, reading, speaking and

listening prior to the experiment and those thereafter.

If, on the one hand, the apparent purpose of the experiment reported here was to check whether or not clustered tasks would prove effective in keeping the participants engaged in face-to-face meaningful oral/aural interactions for an extended period of time; on the other, its immanent aim seemed to be in line with those of the designed unit of work, and the tenets of the task-based learning (TBL) approach (Willis, 1996; Skehan, in Willis & Willis, 1996; Ellis, 2003; Leaver & Willis, 2004; Nunan, 2004; Van den Branden, 2006; Willis & Willis, in press). The choice of this approach for the experiment was based on the literature reviewed in this paper, which suggests that a task-based learning approach towards teaching an L2 is likely to contribute to the development of learner *interlanguage*, which ultimately is the language with which people communicate in an alternative language (Selinker, 1972).

## **2. In defence of a task-based learning approach towards L2 teaching**

In this section, on the basis of the latest research developments on task-based language learning and teaching, a case will be made in favour of the adoption of this approach in the L2 classroom. But first, for the sake of contextualisation, it seems important to briefly trace back the roots of TBL.

### **2.1. The emergence of TBL in the L2 classroom**

The literature presents three major predecessors of TBL within the field of language teaching: (1) the Grammar-Translation Method (GTM); (2) the Audiolingual Method (ALM); and (3) Communicative Language Teaching (CLT). During many centuries, The GTM was the common methodological tool used in the teaching of Latin and Greek in Europe (Rivers, 1981). According to Bowen, Madsen and Hilferty (1985), until the advent of the World War II, this method was widely applied to the process of learning and teaching modern languages around the world. Typically, the GTM capitalised on the teaching of grammar and translation practice, being focussed mostly on reading and writing (Richards & Rogers, 2001, cited in Flowerdew & Millier, 2005).

Howatt (1984) argues that the Reform Movement in the second half of the nineteenth century arose from language-oriented professionals' dissatisfaction with the GTM. As a result of this movement, a number of methods ensued, such as Lambert Sauveur's Natural Method, the Direct Method (popularised by Maximilian Berlitz), and the Audiolingual Method. The major priority of these methods was the speaking and listening skills. The ALM, which outlasted the other offspring of the Reform Movement, was heavily relied on by L2 teachers between the 1950's and the 1970's. Broadly speaking, the formula of this method was quite

simple: intensive rote learning plus avoidance of error equalled assimilation and accurate L2 use in any context (Leaver & Willis, 2004).

Nonetheless, for its behaviouristic nature, the ALM would soon suffer severe criticisms from scholars such as David Ausubel (Hitotuzi, 2006) and Noam Chomsky. The latter, for instance, argued that the child is equipped with an innate basic rule system which accounts for his or her potential to learn any given language to which he or she may be exposed:

To learn a language, then, the child must have a method for devising an appropriate grammar, given primary linguistic data. As a precondition for language learning, he must possess, first, a linguistic theory that specifies the form of the grammar of a possible human language, and, second, a strategy for selecting a grammar of the appropriate form that is compatible with the primary linguistic data.

(Chomsky, 1965, p. 25)

The proposal of a genetic universal *mould* in which any language could be *cast* was undoubtedly a serious argument against behaviourism. This foundation of the ALM was further undermined by research into second language acquisition. Corder (1967), for instance, found evidence that errors can be instrumental in the developmental process of L2 learning. Moreover, according to Leaver and Willis (2004, p. 5), an investigation conducted by Pienemann (1988) indicated that the order in which learners acquire linguistic items may not correlate with the one in which these are presented to them. It seems learners have an “inbuilt internal syllabus” which regulates the intakes of the target language to which they are exposed.

CLT is the teaching model from which TBL stems directly (Leaver & Willis, 2004). It began taking shape in the early 1970’s, as a reaction to focus-on-form language teaching methods of the time. As such, CLT has mustered insight from a number of fields of knowledge. The notions of competence and performance, for instance, are associated with Chomsky’s (1965) Transformational-Generative-Grammar theory. Furthermore, from the stand of Anthropology and Sociolinguistics, Hymes’s disagreement with Chomsky on the boundaries of *competence* led to a redefinition of this concept, which, from his perspective, should comprise language use (performance) as well. Thus, focusing on language in actual performance, Hymes devised an interdisciplinary (Hayes Jacobs, 1989) model of communicative competence which was summarized by Canale and Swain (1980, p. 16, quoted in Neves, 1993) as:

[...] the integration of grammatical (what is formally possible), psycholinguistics (what is feasible in terms of human information processing), sociocultural (what is the social meaning or value of a given utterance), and probabilistic (what actually occurs) systems of competence.

The seminal posthumous work of the philosopher of language John Langshaw Austin, *How to Do Things With Words* (Austin, 1962), is another contribution to the development of CLT. Later, one of Austin's disciples, John Searle, fully developed a speech-act theory, which is credited to his master (Searle, 1970). This theory was welcomed by the 'artisans' of the communicative approach mostly for its focus on the functional aspects of language, and its argument that meanings derive less from grammatical form than from rules of interpretation prevalent in a given social context. Following suit, in *Learning how to mean*, published in 1975, the English linguist Michael Alexander Kirkwood Halliday presented a typology of language functions, which was added to the theory of communicative competence.

CLT has also received important contributions from the field of psycholinguistics. Through his input+1 theory hypothesis, for instance, Krashen (1982, 1985) suggests that exposure to authentic language is fundamental for language acquisition.

It is important to point out, though, that regarding the purpose of this paper one could not do justice here to the various scholars, from different fields of knowledge, whose works have played a major role in the development of Communicative Language Teaching. It seems to suffice to say that from an interdisciplinary *invisible* movement CLT emerged, a version of which, known as task-based language learning, began to materialise some twenty years ago. On the issue of paradigm shifting, Hermans (1999) argues for the existence of an invisible college which mostly unnoticed establishes or changes theory paradigms. And it appears that Prabhu's Communicational Teaching Project in Bangalore (Prabhu, 1987) was a major milestone in the process of "changing winds and shifting the sands" (the phrase is from Albert Marckwardt, 1972, p. 5, quoted in Brown, 2000, p. 13) towards this new language-teaching paradigm (Leaver & Willis, 2004; Van den Branden, 2006). In reality, the results of this project seemed to have indicated that TBL might represent a promising alternative to existing methods of the 1980's, as suggested by Tarone and Yule (1989, p.102):

Reports from the Bangalore Project indicate that a syllabus organized around problem-solving tasks and feedback can effectively accomplish, and in many respects improve on, what a traditional linguistic syllabus provides.

But before presenting any further evidence as a justification for applying TBL in the L2 classroom, it is deemed of necessity to introduce the definition of 'task' upon which the investigation reported in this paper is based.

## **2.2. Defining a communicative task**

As language-oriented scholars' interest in task-based approaches to language teaching have increased from the last quarter of the last to the beginning of the present century, so has the literature on TBL where one can find a variety of ways in which a task can be defined. Nonetheless, after close examination of a number of task definitions, it was decided that the one proposed by Rod Ellis appears to be more adequately in line with the kinds of activities developed in the study reported here:

A task is a workplan that requires learners to process language pragmatically in order to achieve an outcome that can be evaluated in terms of whether the correct or appropriate propositional content has been conveyed. To this end, it requires them to give primary attention to meaning and to make use of their own linguistic resources, although the design of the task may predispose them to choose particular forms. A task is intended to result in language use that bears a resemblance, direct or indirect, to the way language is used in the real world. Like other language activities, a task can engage productive or receptive, and oral or written skills, and also various cognitive processes.

(Ellis, 2003, p. 16)

In other words, an educational task should resemble an outside-world task insofar as it requires interaction among participants and the application of all abilities and cognitive processes involved in actual language use. And this is what I had in mind when designing the unit for the experiment.

### **2.3. Support from second-foreign language research findings**

Since the implementation of the Bangalore Project, considerable amounts of research findings have provided reasonably firm grounds for the adoption of a task-based approach in the L2 classroom, in various guises, to meet specific requirements of different classroom contexts (Bygate, Skehan & Swain, 2001). Such flexibility of the model seems to account for the variety of task definitions recurrent in the works of some scholars, such as Ellis (2003), Nunan (2004), Leaver and Willis (2004) and Van den Branden (2006). Thus, in this section I shall present a summary of findings deriving from tasks which may not conform entirely to the task definition proposed by Ellis (2003).

Amongst the many language-oriented researchers who have informed TBL with their empirical investigations is Cathcart (1986, cited in Chaudron, 1988) who, after observing eight Spanish-speaking kindergarten children in various activities for a year, pointed out that 'An increase in utterance length or complexity was found [...] in those peer-peer interactions, involving tasks with a joint goal (as in a joint block-building activity)' (Chaudron, 1988, p. 98).

The results of Rulon and McCreary's (1986, cited in Chaudron, 1988, p. 108) 'comparison between teacher-fronted and group work negotiation for meaning' also seem to endorse the reliability of TBL. The point they make is that through group work focussed on

meaning, interaction is promoted and, eventually, L2 learning ensues (see also Johnson, 1983; Gaies, 1983b; Pica & Doughty, 1985; Duff, 1986, all cited in Chaudron, 1988).

On the aspect of meaning negotiation during the interaction event, Long (1990), one of the pioneers of TBL according to Van den Branden (2006), identified a larger volume of negotiation in tasks in which the participants need to respond to a common problem. Pica, Kanagy and Falodun (1993) also suggest the possibility of an increase in the volume of negotiation from specific interactive activities.

Studies based on experiments with tasks requiring justifications indicate that these generate highly complex utterances. By the same token, results from comparisons between interactive and monologic tasks showed that the former produces much more precision and complexity, whereas the latter generates more fluency (Foster & Skehan, 1996; Skehan & Foster, 1997, 1999). In several studies conducted by Foster and Skehan (1996, 1999), Foster (1997, 1999), Mehnert (1998) and Ortega (1999), it was verified that task planning produces positive influence on these two aspects of learner performance.

Experiments with task repetition have also demonstrated the positive results of TBL in the L2 classroom. Grounded in several studies, Bygate (1996, 1999, 2001) establishes that task repetition contributes substantially to the enhancement of the development of L2 learners. He argues that it produces more propositional density and syntactic quality insofar as the tasks are repeated. Following suit, Lynch and Maclean (2001) report that, after conducting a number of tasks, an ESP group of medical specialists expressed the perception of their improvement in the target language as a result of undergoing the process of task repetition.

Lochana and Deb's (2006) project in a school run by the Basaveshwara Education Society in India is yet another evidence in support of a task-based approach to language teaching and learning. They developed an experiment in which non-task-based textbook activities were converted into task-based ones in order to test two hypotheses: (1) 'Task-based teaching enhances the language proficiency of learners'; and (2) 'Tasks encourage learners to participate more in the learning processes' (Lochana & Deb, 2006, p. 149). Their findings suggest that TBL is beneficial to learners not only in terms of proficiency enhancement but also motivationwise. Similar results were also obtained by Rocha (2005) and Gutiérrez (2005).

Reports of research findings such as these are likely to encourage teachers to comfortably apply TBL to their classrooms, inasmuch as it seems to fulfil fundamental conditions for learning a second language, namely exposure (or input), meaningful use, motivation and language analyses, as Willis (in Willis & Willis, 1996) makes the point.

#### 2.4. The ‘backbone’ of the unit of work

It appears that the great challenge facing teachers and researchers in the 1970s, and in years to come, was learning how to cope with *form* and *use* in the L2 classroom (Morrow, 1981). For one thing, the kind of approach which overemphasised the teaching of form would be likely to provide learners with the ability to produce well-formed sentences and yet lacking in, for instance, communicative and sociolinguistic competences (Tarone & Yule, 1989). For another thing, the approach focussed solely on use would probably yield ‘ungrammatical’ L2 communicators (Stern, 1992; Tarone & Yule, 1989).

Along the years, defenders of TBL have also addressed this question, perhaps as an attempt to swerve the radicalism of focus on meaning (or use) in the initial stage of this new language-teaching paradigm. From the late 1980’s a number of theorists and practitioners began admitting of tasks focussed on *form* as a preparation for later focus on *use* (Breen, 1989; Tarone & Yule, 1989; Widdowson, 1990). Recently, this approach has been classified as “task-supported language teaching” as opposed to “task-based language teaching” (Ellis, 2003, p. 27 and elsewhere).

Current discussions on task-based learning have also contributed to striking a balance between these two binary features of the L2 classroom, as evinced by the most recent addition to the TBL literature (e.g. SKEHAN, 2003; Ellis, 2003, 2005 and 2006; Nunan 2004; Nunn, 2006; Van den Branden 2006, just to mention a few). Nevertheless, as much as this array of laudable insights may represent an expansion of investigations into the field of language teaching and learning, in my view, Jane Willis’ seminal work, *A Framework for Task-Based Learning*, published in 1996, still provides the groundwork for an adequate way of addressing the dilemma of form versus use in the L2 classroom (Dave Willis and Jane Willis’s (in press) *Doing Task-Based Learning* seems to be a detailed explanation and expansion of the ideas proposed in that work). Willis divides her TBL framework into three major stages: *pre-task*, *task*, and *language focus* (Table 1).



Table 1: Components of the task-based learning framework (adapted from Willis, 1996, p. 38).

Close scrutiny of Willis's framework seems to indicate that its last stage is the key component to achieving a desirable balance between use and form. Nonetheless, it is important to point out that the teacher's conception of language learning is crucial to establishing such equilibrium. Apparently, teachers who hold a holistic view of language learning may comfortably work with TBL, and may be happy with applying consciousness-raising (C-R) activities to their students during the language-focus stage. The simple fact that they offer learners the opportunity to reflect on the possibilities and mechanisms of the language seems to make C-R activities a useful instrument in preparing students to use the target language effectively (Rutherford, 1987; Willis, 1990; Nunan, 1991; Widdowson, 1990; Willis & Willis, 1996; Lewis, 1993).

For the design of the unit of work tested on the group of undergraduates from UFAM, in conjunction with Jane Willis's ideas, I have also capitalised on Michael Breen's argument concerning the involvement of L2 learners in the evaluation of learning task cycles as a means for target-language development and use. Although my focus was on lengthening student talking time while reducing the number of tasks carried out in the classroom, I felt comfortable with electing these two TBL-oriented authors as the "backbone" of the experiment for the valuable insights that they provide. Besides, it appears that, irrespective of

any specific remedial purposes for which a task-based approach may be used, tasks are likely to provide L2 development, as Breen (1989, p. 192) makes the point:

Every task is, in essence, a means for learning. Its purpose is to provide an opportunity for language learners to move from their present state of knowing and capability towards a new aspect of knowledge and specific use of skills and abilities.

But it is Breen (1989, p. 192) himself who signals caution in terms of expectations as to learning outcomes deriving exclusively from a task-based approach. He argues that ‘a task cannot predict what a learner will actually contribute to it or learn from it.’ This appears to be a relevant point, but, although establishing precise results through *workplans* (to use Breen’s terms) might seem unrealistic, there was a tentative attempt at defining the learning objectives of the unit. Thus, by the end of the experiment the participants may display enhancement in terms of:

- awareness of some uses of perfect tenses;
- listening comprehension;
- production of more coherent e cohesive oral and written texts;
- pronunciation;
- spelling;
- vocabulary repertoire.
- understanding of the rationale and principles informing TBL;

As can be seen, these learning goals cover both receptive and productive skills, the development of which can be promoted by tasks, according to Leaver and Willis (2004).

### **3. Method**

#### **3.1. Participants**

The experiment was conducted on a group of thirteen Brazilian students in their second year at the Federal University of Amazonas studying for a First Degree in Liberal Arts/TEFL. The participants (4 males and 9 females) attending the 60-hour IHE104 – English IV course held twice a week in 100-minute sessions were in the 20-35-age range; all but one had part-time jobs. Moreover, from the working group eight students had teaching posts as teachers of English themselves.

In the term that the study was conducted, part B of the intermediate level of the *True to Life* series (Gairns & Redman, 1996a and b) was adopted as the core coursebook for IHE104 – English IV, as part A had been used experimentally in the previous term. Thus a selection of activities both from the Student’s Book and the Workbook of the series were used in the experiment.

### **3.2. Materials**

**3.2.1. Sources of language input (other than the teacher's and the learners' input):** A 6-minute extract from a coverage of May Day demonstrations on *BBC News* (Sambrook, 2001); the book *Basic Law for Road Protestors* (Gray, 1996); a tape-recorded *Speak Up* (Gould, 1988) interview with Nadine Gordimer (both the tapescript in the magazine and the audiotape accompanying it were used in the experiment), and Peter Roach's *English Phonetics and Phonology* (Roach, 1991).

**3.2.2. Recording equipment:** An NV-VJ60PN Panasonic video camera and one 90-minute VHSC TC-30 Panasonic videotape were used for the recording of the lesson involving the task informing the answer to the initial question. In addition, a DW-003 CASIO stopwatch was used to time peer-peer interaction during this final task, the timing of which was done when I viewed the 64-minute video recording in private (Table 4). Moreover, for the recording of (1) an enactment of Nadine Gordimer interview, and (2) an original monologue on types of protests, the participants were given thirteen 60-minute HF Sony audiocassettes. The equipment for these two recordings was provided by the participants themselves.

In passing, at least two variables seem to have affected these two micro-task outcomes. For one thing, the words on the 13-minute audio recording of the *Speak Up* interview with Nadine Gordimer were somewhat 'muffled'. For this reason, the participants complained they had difficulty in listening to the recording without the help of the script. For another, some of the cassette recorders used by the participants were of poor quality. It seems nothing could be done about this, insofar as I was unable to provide adequate recording equipment.

**3.2.3. Worksheets:** Five different kinds of C-R activities were devised: two on perfect tenses (Worksheets 1 and 4), two on pronunciation (Worksheets 2 and 3), and another one (Worksheet 5) containing some of the students' utterances during the videotaped lesson.

### **3.3. Procedure**

#### **3.3.1. Phases of the macro-task framework**

The unit was structured in a peculiar fashion. A macro-task framework was designed into which data from four micro-task frameworks were fed in order to be analysed by the participants in its task-cycle stage. As demonstrated in Table 2, the four micro tasks functioned as a pre-text for the macro task to materialize (Cecily O'Neill uses the expression 'pre-text' as an umbrella term to cover any 'text' that can provide 'occasions for initiating dramatic action' (O'Neill, 1995, p.19).

THE MACRO-TASK FRAMEWORK

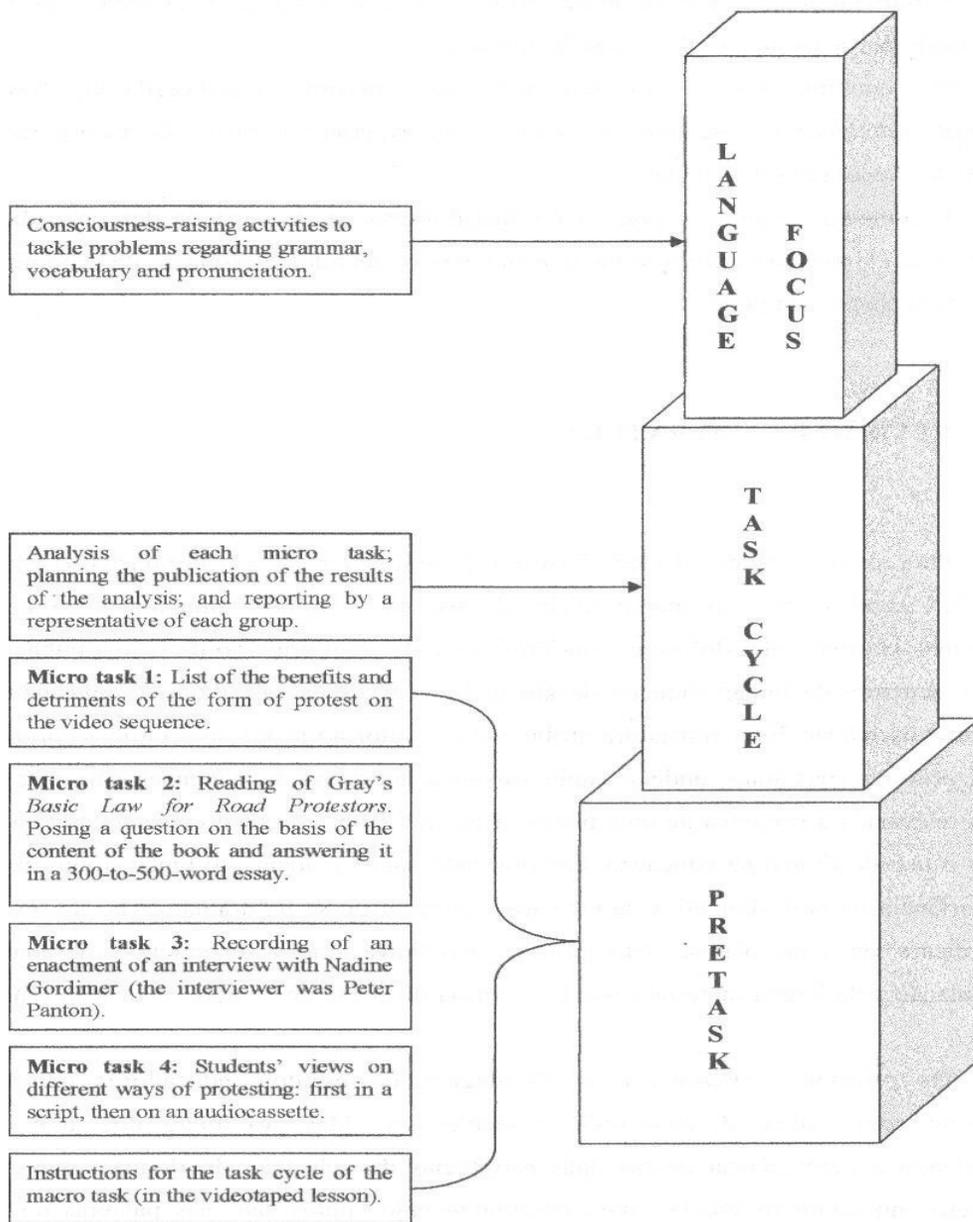


Table 2: The planned unit of work

**The pre-task:** After being briefed about the whole cycle of activities comprising the unit of work, the group first watched a video sequence on May Day demonstrations, discussing it in small groups afterwards. They then read Gray's book (14,292 words), and were requested to write an essay on protests (Appendix I). Subsequently, the group listened to and read the tapescript of the interview with Nadine Gordimer; here, in pairs, the students tape-recorded

the enactment of the dialogue between the interviewer (Peter Panton) and the interviewee. Next, they also tape-recorded the monologue expressing their views on different kinds of protests. Finally, in a videotaped session, they reflected on the whole cycle of TBL activities, and prepared a report on its strong and weak points, presenting possible solutions for the flaws.

**The task cycle:** (a) The task – in small groups, the students analysed each one of the phases of the micro-task frameworks within the planned unit, detecting positive and/or negative aspects of the phases, and of the whole process as well as possible solutions for the weak points. b) The planning – each group prepared a report about the strong and weak points of the entire cycle of activities that they had carried out so far, presenting some possible solutions for the flaws. (c) The report – a spokesperson elected within each group presented their group's report in front of the class. The whole sequence of this phase of the macro-task framework was captured on videotape. As demonstrated here, the outcome (i.e., the end-product, or the 'successful completion of the task', as Leaver and Willis (2004, p.13) put it) of the macro task was evinced by the participants' completion of the evaluation of the activities within the frameworks of the micro tasks, and their suggestions presented to the entire class.

**The language focus:** After analysing the video in private, I prepared the following C-R activities. Small groups were given a list of some of their assertions captured on tape, and were requested to identify possible errors, justify their views, and think of more appropriate ways of rephrasing the sentences or chunks that they thought were incorrect (Appendix II: Worksheet 5: Spot the error). The whole class viewed the videotaped lesson, and again responding to my request, attempted to identify problems concerning pronunciation. I then called their attention to pronunciation problems overlooked during the 'spot-the-error' activity. The activities carried out by the participants at this stage are labelled as focused tasks by Ellis (2003).

**3.3.2. Summary of the videotaped lesson:** This lesson was conducted in a 64-minute session, and all 13 participants attended it.

**The pre-task:** This stage corresponded to the debriefing on the activities that the participants had engaged in within the frameworks of the four micro tasks, plus the instructions to the macro task (Table 2). First, in small groups, the students made a detailed list of those activities. Next, the students were given my original list of

activities (Table 3); they compared lists and made comments. Then, slips of paper containing all the steps of the task cycle were distributed around the class. After reading the instructions for the task cycle, a volunteer stood up and explained what each small group was expected to do – this technique can provide reassurance of instruction comprehension. Finally, I followed the volunteer’s accounts and gave feedback.

<p><b>Framework 1 – Video sequence</b></p>	<ul style="list-style-type: none"> <li>⇒ Exploratory activity involving new vocabulary study and general comprehension</li> <li>⇒ Task involving discussion of the topic, and presentation of oral report</li> <li>⇒ C-R activities highlighting the past perfect tense</li> </ul>
<p><b>Framework 2 – Book: <i>Basic Law for Road Protestors</i> by Peter Gray</b></p>	<ul style="list-style-type: none"> <li>⇒ Reading comprehension</li> <li>⇒ Writing activity on the basis of the book in two phases:                             <ul style="list-style-type: none"> <li>a. Writing/analysis of the first draft</li> <li>b. Editing/analysis of the essay</li> </ul> </li> </ul>
<p><b>Framework 3 – Interview with Nadine Gordimer (printed and audiotaped)</b></p>	<ul style="list-style-type: none"> <li>⇒ Recording of the enactment of the interview</li> <li>⇒ C-R activities on pronunciation involving the IPA</li> <li>⇒ C-R activity on word classes, involving strong forms and weak forms, informed by Peter Roach’s <i>English Phonetics and Phonology</i> against the recording and the tapescript of the interview</li> <li>⇒ C-R activities on perfect tenses</li> </ul>
<p><b>Framework 4 – Writing and Speaking activities</b></p>	<ul style="list-style-type: none"> <li>⇒ Analysis, in groups, of each one of the three previous phases</li> <li>⇒ Writing: scripted comments on different types of protests</li> <li>⇒ Recording of the script</li> <li>⇒ Teacher’s feedback on these activities</li> </ul>
<p><b>Framework 5 – Analysis of four micro-task frameworks</b></p> <p><b>Note: as the participants were engaged in this framework when they were asked to make their lists, it was not mentioned by them in detail.</b></p>	<ul style="list-style-type: none"> <li>⇒ Debriefing on the activities involving the micro tasks</li> <li>⇒ Instructions for the micro task</li> <li>⇒ Analysis of the of activities carried out within the frameworks of the micro tasks</li> <li>⇒ Identification of positive and negative aspects involving the whole of the macro-task framework</li> <li>⇒ Suggestions aiming at the improvement of future cycles of activities based on the task-based-learning approach</li> <li>⇒ C-R activities on relevant morphosyntactic and suprasegmental information that were necessary but were either missing or incorrectly supplied</li> </ul>

Table 3: My original list of activities for the whole cycle of TBL activities

Because they are part and parcel the macro-task framework, both the task cycle and the language focus of the videotaped lesson coincide with those within it.

### 3.4. Analysis

In order to find out whether or not the participants would in fact be engaged extensively in peer-peer meaningful interaction in the classroom as a result of their engagement in the

macro task, the total amounts of talking time in the classroom captured in the videotaped lesson was timed. Moreover, to reckon the amounts of time the students spent interacting with each other, the length of classroom talk (LCT) was divided into teacher talking time (TTT), student talking time (STT), and periods that neither the teacher nor the students spoke (henceforth, silence length (SL)), as it is demonstrated in Table 4 (For a similar approach in a different study, see also Hitotuzi, 2005.).

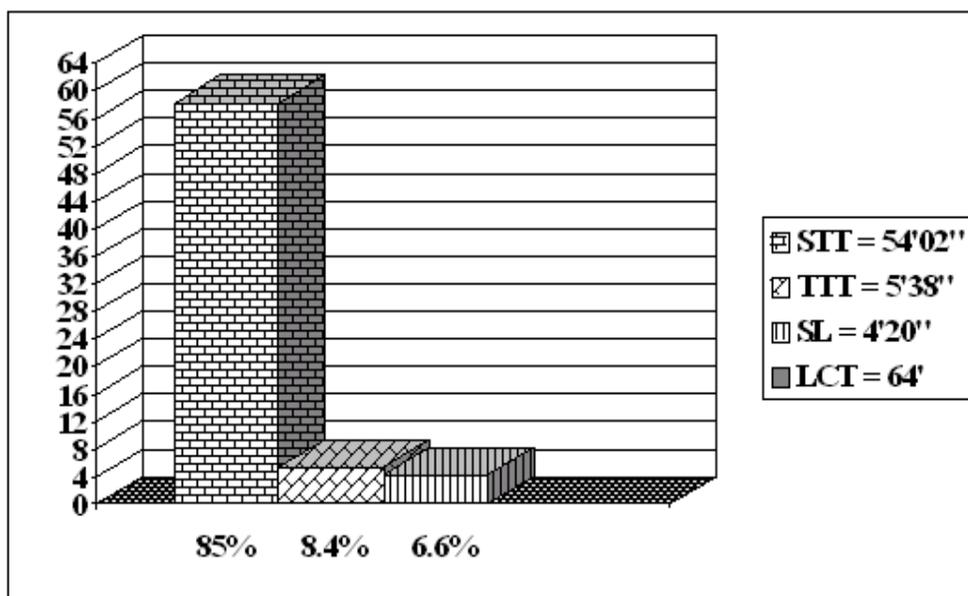


Tabela 4: Videotaped lesson

## 4. Results

### 4.1. On the outcome of the macro task

The graphic in Table 4 indicates that 85 per cent of LCT was spent on STT, which is above my expectations for intermediate-to-advanced STT in a 64-minute session (e.g. 45 to 70 per cent). This seems to show that a few integrated small learning tasks, used as a pre-text, can provide scaffolding for extended face-to-face meaningful interaction amid learners engaged in task-cycle analyses. Furthermore, these results also point to the fact that such an approach can be more economical in terms of time allocated for preparation of lessons aiming at the provision of ‘fodder’ for massive student talk than one involving a variety of small tasks. At any rate, because the bulk of the macro task consisted essentially of discussions, it appears that these were an important variable accounting for the high proportion of STT. According to Skehan (2003, p. 5) a discussion task provides ‘facilitation for extended turns’, and allows learners to reach ‘the greater depth of interaction’.

Another important dimension of this task-within-a-task (TWAT) model seems to be its

implications for the development of the learner as a whole. Besides the opportunity with which they provide learners to use the target language purposefully, they can be an important enhancer of learner autonomy. With the participants of the experiment, for instance, the approach was particularly useful in this respect, because it placed them on both sides of the classroom. Through the analysis of task cycles, the participants became more aware of the objectives of the lessons and important aspects of the teaching methodology, amongst other aspects relevant for their teaching qualification.

Planning is yet another important characteristic of the TWAT model. The fact that this metacognitive learning strategy (Oxford, 2006) seems to allow learners to work on tasks more comfortably and confidently may account for its beneficial influence on fluency and complexity, as demonstrated in a number of studies (Foster & Skehan, 1996, 1999; Foster, 1997, 1999; Mehnert, 1998; Ortega, 1999).

Lastly, but equally important, is the fact that the figures from the videotaped lesson evince that the macro-task cycle provided interactional data that are consistent with its workplan. This may not be the case with most task-based planned activities. Seedhouse seems to concur to that. He claims that ‘The vast majority of studies do not provide task-in-process interactional transcript data that may be compared with the task-as-workplan’ (Seedhouse, 2005, p. 548). Nonetheless, one should be quick to point out that, in order to support his argument in “Task” as Research Construct, Seedhouse capitalises only on the task-supported-learning (Ellis, 2003) view of task definition (which is not the one adopted in the study reported here), despite his claims otherwise: ‘These conceptions of task-as-workplan and task-in-process, then, apply to any and all activities that are planned and occur in second language (L2) classrooms, whether or not they conform to definitions of ‘task’ in the TBL literature’ (Seedhouse, 2005, p. 535).

#### **4.2. Feedback on skill development**

Retrospectively, the cycles of TBL activities described here seem to have contributed to integrating the learners’ target-language skills. This was verified by simple comparison between the participants’ performances in writing, reading, speaking and listening prior to the experiment and those during and after it. Although *thought* was not assessed in the experiment, it is likely that it has been developed throughout the stages of the investigation, inasmuch as the cycles of activities involved a thinking process (Ur, 1981, p. 13).

**4.2.1. Writing:** It was observed that the approach adopted towards the writing task that the participants were asked to carry out as a result of their reflections on the theme underlying the

tasks they had done so far, yielded positive results. It appears that the use of acronyms to pinpoint errors in the group's first drafts, and the opportunity that the students were given to edit them may have contributed to the drastic decrease in the number of errors normally found in their essays prior to the study. The effectiveness of the approach can be amply exemplified by the comparison between the learners' first drafts and the edited essays (Appendix Ia and Ib -for lack of space, only two pairs of the participants' essays are appended to this paper).

**4.2.2. Reading:** Basically, the feedback on this skill was derived from the essay on Peter Gray's book plus the discussion in the classroom involving the video sequence, Nadine Gordimer's interview, and the book proper. It was observed, however, that some of the learners avoided referring to the content of the book in crucial occasions, perhaps for the same reason that other L2 learners strategically avoid using complex-structure forms (Tarone & Yule, 1989; Schachter, 1974, cited in Richards, 1985). At least two plausible hypotheses can be proposed in light of such avoidance strategy adopted by the participants: (1) either they had difficulties in understanding the content of the book, or (2) they did not appreciate it. In hindsight, to solve the problem of comprehension, a glossary of the relevant technical jargon in the book should have been provided, as one of the participants pointed out in the analysis stage (see participants' quotes below). Probably allowing the learners to choose a book of their liking would have invalidated the second hypothesis.

**4.2.3. Speaking:** While not following Willis's TBL framework *ipsis litteris*, the group had already been working interactively in the classroom for the last two terms. Thus, perhaps for this reason, most of them spoke with fluency compatible with the level of the course. Nonetheless, it was observed, during the implementation of the unit, that the learners were more focussed on the discussions than before. Apparently this is accounted for by the necessity to produce a tangible outcome found in Willis's framework. Additionally, perhaps on account of their knowledge of the purpose of the set of activities, the participants tried to 'show off' their English. However, many are likely to agree that the halo effect (Brown, 1988) in such a context cannot be seen as a threat to the validity of this kind of experiment, since ultimately it is aimed at the enhancement of the participants' performance in the target language. Ellis (2005) provides support for the beneficial effects of massive L2 use in terms of discourse skills.

**4.2.4. Listening:** Due to its brevity, any improvement in the listening comprehension skills of the participants as a result of the experiment was unlikely to be measurable. However, it is

possible that the administration of TBL activities for long periods of time can have a positive influence on L2 learners' listening abilities. And this may be the case because a genuine TBL model capitalises heavily on authentic input (Nunan, 2004), which seems to be the crux of language skills enhancement in general. In fact, it appears that exposure to concocted texts cannot harness L2 learners to cope adequately with the demands of language tasks in the outside world (Van den Branden, 2006).

The perception that a task-based approach to language learning can be productive in terms of listening comprehension development is also shared by the chief actors of the learning process. An example of that is the investigation on the effects of TBL on learning outcomes in the ESP classroom conducted by Kavaliauskienė (2005). Responding to a questionnaire on this topic, ESP learners elected listening as one of the skills which were positively influenced as result of their engagement in TBL activities.

#### **4.3. The participants' comments on the cycles of TBL activities**

The participants had had genuine opportunity for meaningful communication in and outside the classroom throughout some of the phases of the cycles of TBL activities. This was especially the case with the task in which they were requested to analyse the micro-task frameworks as a whole, and present suggestions for improving future TBL activities. Notwithstanding all the intimidation a video camera can cause (Allwright & Barley, 1991), the students spent over 50 minutes (precisely 54'02") engaged in genuine oral/aural interaction in the target language in the attempt to convey their viewpoints on each and every stage of the planned unit. Thus, contrary to Nunan's (2004, p.14) suspicion of 'rhetoric' not matching 'reality', the experiment has revealed how effective a task-based learning approach can be in terms of creating the necessity for real communication in the classroom.

Here are some of the participants' comments on the cycles of TBL activities, which seem to demonstrate the students' purposefulness and engagement in the discussions:

*Nadine Gordimer's interview was nice because (...) it gave us the opportunity to see how native speakers express themselves.*

*We learned a lot of things from this video passage.*

*We learned some expressions... idiomatic expressions.*

*We could test our listening skill.*

*It was a kind of activity that really got us involved.*

*It also increased our knowledge about British culture and their law [system].*

*We had the opportunity to be in touch with different levels of registers.*

*We improved our speech and pronunciation when we recorded and listened to ourselves.  
We developed the four skills.*

*The difficulties were not exactly a negative point.*

*We should have more contact with the video... use video passages more often.*

*The first time we saw the video it was hard to understand everything.*

*I couldn't understand what I was supposed to do.*

*The book was very boring.*

*One thing I consider very important is the composition, you know.*

*We had contact with natural and real-time English.*

*A negative point about the book is that there were many technical words. I think we should have a kind of glossary [on] a separate [sheet].*

*We didn't have a special microphone, a special tape recorder, and things like that.*

*I think we should have a special laboratory, a special palace to do [the recording].*

*We didn't have special [equipment] to record the tape.*

*We had to listen to [the interview], and later had to record it. After [recording it], we could listen to our own voice and sy... so I have improved [on] this, I'm good [at] that (...).*

*When I heard my voice I was so surprised [at] myself because... I... I... It was funny because er... I said: 'It's my voice! I don't believe!' And I [was] surprised, especially because I'm improving now. Yes... I think this experience was great for me.*

*About the composition that we [wrote] er linking the three activities... we had er a little bit of difficulty to to that because, in our opinion, it was hard to link the three because they were three different kinds of speech.*

## **5. Conclusions**

In conclusion, despite the limited scope of the experiment, the results as well as the entire process reported in this paper seem to indicate that the main objective of the unit of work was successfully achieved; and that the major factor accounting for it was the set of integrated micro tasks and its subsequent analysis by the participants. For one thing, each phase within the four micro-task frameworks fed important cumulative data into the task cycle of the

macro-task framework, which was a key element informing the results of the experiment. For another thing, as the participants experienced each component of Willis's framework, they grew more familiar with the TBL model, and, as a corollary, their performance improved significantly in the subsequent micro task cycles and in the macro task cycle. The principle behind this finding is echoed in the following assertion by Rod Ellis (See also Foster & Skehan 1996; Skehan & Foster 1997):

When learners know what they are going to talk or write about they have more processing space available for formulating the language needed to express their ideas with the result that the quantity of the output will be enhanced and also fluency and complexity (Ellis, 2006, p. 23).

The main task of the macro-task framework described here can be categorised as both a pedagogic task and a target task, following Nunan's (2004) task classification. At the same time that it was conducted in the classroom and had an educational aim, it was a real-world task to the extent that those were real TEFL undergraduates engaged in analysing methodological approaches to language teaching and learning.

On the face of it, at least four assumptions emerge from the experiment. Firstly, micro tasks within a macro task may lend themselves as a useful tool for familiarising learners with the task-based learning process, which in turn may enhance their proficiency in the target language (Bygate, 1996). Secondly, micro cycles of TBL activities as an element in the pre-task stage of a macro cycle may provide data for genuine communication to take place in the classroom in an economical fashion. Thirdly, integrated workplans can be designed to cater for learners' needs in terms of all five skills, namely listening, speaking, reading, writing and thinking. Lastly, a TWAT model can provide enough 'fodder' for meaningful peer-peer interactions over extended periods of time. Additionally, there is also room in this approach for work on lexis and grammar through C-R activities, which are most likely to play a pivotal role towards striking the complex balance between form and use in the L2 classroom, as suggested by a number of theorists and practitioners (Leaver & Willis, 2004).

These assumptions seem to suggest positive implications for the L2 classroom. One such is that the accomplishment of a task-cycle evaluation involving the learners might be an important component in the developmental process of proficiency in the target language (Breen, 1989). Additionally, this approach might be a less uniform route towards achieving the results of the task-repetition approach proposed by Bygate (1996, 1999, 2001), which may not be welcomed in certain classroom contexts (Plough & Gass, 1993). Seemingly, another beneficial implication is rendered by the ability to develop all the five skills, and to tackle structural and lexical problems through activities that raise learners' awareness of the

target language.

While it might be argued that the assumptions outlined here are forgone conclusions, validating them in the classroom may not bear so clear-cut predictability. The prime example of this is the usual lack of homogeneity in the L2 classroom in terms of learner level of proficiency. Consequently, the notion of small groups of learners involved in processing a meaningful final task may be easily dispelled by huge gaps in their levels of proficiency. There is more to the TBL model than collaborative work in small groups though. Learners' preconceptions of how languages are learnt can be equally dismissive of such predictability (Stern, 1992; Johnson, 1989). Even if half the group has understood and accepted the rationale and principles of the approach, the other half willing, say, to focus exclusively on form might be likely to hinder unimpeded meaningful peer-peer interaction in the course of the cycles of activities. Finally, providing learners with inadequate kinds of input is yet another factor working against this 'obviousness'. All of the previous caveats being dismissed, still the foregoing assumptions would be unrealistic had the learners been unable, or found it too hard, to digest the material to which they were exposed (Krashen, 1982, 1985).

### **Further research**

The present study has raised a number of questions that may merit further attention, three of which will be listed here:

1. Considering Gardner's (1991, 1993, 2001, 2006) suggestions on multiple intelligences and learning styles, how would this task-within-a-task approach benefit learners who, for instance, prefer to keep aloof in the classroom?
2. Other than analysing the phases in cycles of pedagogic activities with the purpose of identifying positive and negative aspects, as well as presenting alternative strategies to tackle possible flaws, what other *outcomes* (if any) can the learners aim at as they carry out a macro task? Perhaps a study on the development (if feasible) of a typology of possible macro-task outcomes would be desirable at this point.
3. Would a TWAT model be adequate for beginner-level learners?

### **References**

- Allwright, D. & Bailey, M. (1991). *Focus on the language classroom. An introduction to classroom research for language teachers*. Cambridge: Cambridge University Press.
- Austin, J. L. (1962). *How to do things with words*. Oxford: Oxford University Press.
- Bowen, J. D. Madsen, H. & Hilferty, A. (1985). *TESOL techniques and procedures*. Rowley, MA: Newbury House Publishers.

- Breen, M. (1989). The evaluation cycle for language learning tasks. In R. K. Johnson (Ed.), *The second language curriculum*. Cambridge: Cambridge University Press.
- Brown, H. D. (2000). *Principles of language learning and teaching* (4th ed.). White Plains, NY: Pearson Longman.
- Brown, J. D. (1988). *Understanding research in second language learning: A teacher's guide to statistics and research design*. Cambridge: Cambridge University Press.
- Bygate, M. (1996). Effects of task repetition: Appraising the developing language of learners. In J. Willis & D. Willis (Eds.), *Challenge and change in language teaching* (pp. 136-146). Oxford: Heinemann.
- Bygate, M. (1999). Task as context for the framing, reframing, and unframing of language. *System*, 27, 33–48.
- Bygate, M. (2001). Effects of task repetition on the structure and control of oral language. In M. Bygate, P. Skehan & M. Swain (Eds.), *Researching pedagogic tasks: Second language learning, teaching, and testing* (pp. 23-48). London: Longman.
- Bygate, M., Skehan, P. & Swain, M. (Eds.) (2001). *Researching pedagogic tasks: Second language learning, teaching, and testing*. London: Longman.
- Chaudron, C. (1988). *Second language classrooms: Research on teaching and learning*. Cambridge: Cambridge University Press.
- Chomsky, N. (1965). *Aspects of the theory of syntax*. Massachusetts: The MIT Press.
- Corder, S. P. (1967). The significance of learner's errors. *International Review of Applied Linguistics*, 5(4), 162-170.
- Cowan, J. (1999). *On becoming an innovative university teacher: Reflection in action*. Buckingham: Open University Press.
- Ellis, R. (2003). *Task-based language learning and teaching*. Oxford: Oxford University Press.
- Ellis, R. (2005). Principles of instructed language learning. *System*, 33(2), 209-224.
- Ellis, Rod. (2006). The methodology of task-based teaching. *Asian EFL Journal*, 8(3), 19-43.
- Flowerdew, J. & Miller, L. (2005). *Second language listening: Theory and practice*. Cambridge: Cambridge University Press.
- Foster, P. & Skehan, P. (1996). The influence of planning on performance in task-based learning. *Studies in Second Language Acquisition*, 18(3), 299–324.
- Foster, P. & Skehan, P. (1999). The effect of source of planning and focus of planning on task-based performance. *Language Teaching Research*, 3(3), 185–214.

- Gardner, H. (1991). *Intelligence in seven steps. New Horizons for Learning*. Retrieved March 22, 2007 from [http://www.newhorizons.org/future/Creating\\_the\\_Future/crfut\\_gardner.html](http://www.newhorizons.org/future/Creating_the_Future/crfut_gardner.html)
- Gardner, H. (1993). *Frames of mind: The theory of multiple intelligences*. New York: HarperCollins Publishers.
- Gardner, H. (2001, March). *An education for the future: The foundation of science and values*. Paper presented at the Royal Symposium Convened by Her Majesty, Queen Beatrix Amsterdam, Netherlands. Retrieved March 06, 2002 from <http://www.pz.harvard.edu/WhatsNew/Amsterdam.htm>
- Gardner, H. (2006). *Multiple Intelligences: New horizons*. New York: Basic Books.
- Gairns, R. & Redman, S. (1996a). *True to life: Intermediate: Personal study workbook*. Cambridge: Cambridge University Press.
- Gairns, R. & Redman, S. (1996b). *True to life: Intermediate: Student's book*. Cambridge: Cambridge University Press.
- Gould, B. (Ed.) (1988). The speak up interview – Nadine Gordimer. *Speak Up – A revista que fala inglês com você, 11*.
- Gray, P. (1996). *Basic law for road protestors* (2nd ed.). Retrieved December 08, 2001 from <http://www.urban75.com/Action/protest.html>
- Gutiérrez, D. G. (2005). Developing oral skills through communicative and interactive tasks. *Profile: Issues in Teachers' Professional Development, 6*, 83-96.
- Halliday, M. A. K. (1975). *Learning how to mean*. London: Edward Arnold.
- Hayes Jacobs, H. (Ed.) (1989). *Interdisciplinary curriculum: design and implementation*. Alexandria, VA: Association for Supervision and Curriculum Development.
- Hermans, T. (1999). *Translation in systems*. London: St Jerome.
- Hitotuzi, N. (2005). Teacher talking time in the EFL classroom. *Profile: Issues in Teachers' Professional Development, 6*, 97-106.
- Hitotuzi, N. (2006). The learner's mother tongue in the L2 learning-teaching symbiosis. *PROFILE: Issues in Teachers' Professional Development, 7*, 161-171.
- Johnson, R. K. (Ed.) (1989). *The second language curriculum*. Cambridge: Cambridge University Press.
- Kavaliauskienė, G. (2005). Task-based learning and learning outcomes in the ESP classroom. *Studies about Languages, 7*, 1-6.

- Krashen, S. D. (1982). *Principles and practice in second language acquisition*. Oxford: Pergamon.
- Krashen, S. D. (1985). *The input hypothesis: Issues and implications*. London: Longman.
- Leaver, B. L. & Willis, J. R. (Eds.) (2004). *Task-based instruction in foreign language education: Practice and programs*. Washington, D. C.: Georgetown University Press.
- Lewis, M. (1993). *The lexical approach*. Hove: Language Teaching Publications.
- Lochana, M. & Deb, G. (2006). Task based teaching: learning English without tears. *Asian EFL Journal*, 8(3), 140-164.
- Long, M. H. (1990). Task, group, and task-group interactions. In A. Sarinee (Ed.) (1990). *Language teaching methodology for the nineties* (pp. 31-50). Singapore: SEAMEO Regional Language Centre (Anthology Series 24)
- Lynch, T. & Maclean, J. (2001). 'A case of exercising': Effects of immediate task repetition on learners' performance. In M. Bygate, P. Skehan & M. Swain (Eds.), *Researching pedagogic tasks: Second language learning, teaching, and testing* (pp. 141-162). London: Longman.
- Mehnert, U. (1998). The effects of different lengths of time for planning on second language performance. *Studies in Second Language Acquisition*, 20, 83–108.
- Morrow, K. (1981). Principles of communicative methodology. In K. Johnson & K. Morrow (Eds.) *Communication in the classroom* (pp. 59-66). Essex: Longman.
- Neves, M. de S. (1993). *Communicative teaching of English as a foreign language: An approach haunted by traditional and structural myths*. Unpublished MA Dissertation, Universidade Federal de Minas Gerais, Belo Horizonte, MG.
- Nunan, D. (1991). *Language teaching methodology: A textbook for teachers*. Hemel Hempstead: Prentice Hall.
- Nunan, D. (2004). *Task-based language teaching: A comprehensively revised edition of designing tasks for the communicative classroom*. Cambridge: Cambridge University Press (Cambridge Language Teaching Library).
- Nunn, R. (2006). Designing holistic units for task-based learning. *Asian EFL Journal*, 8(3), 69-93.
- O'Neill, C. (1995). *Drama worlds: A framework for process drama*. Portsmouth, NH: Heinemann.
- Ortega, L. (1999). Planning and focus on form in L2 oral performance. *Studies in Second Language Acquisition*, 21, 109–148.

- Oxford, R. L. (2006). Task-based language teaching and learning: An overview. *Asian EFL Journal*, 8(3).
- Pica, T., Kanagy, R., & Falodun, J. (1993). Choosing and using communication tasks for second language instruction. In G. Crookes & S. Gass (Eds.). *Tasks and language learning: Integrating theory and practice* (pp. 9-34). Cleve-don, England: Multilingual Matters.
- Pienemann, M. (1988). Determining the influence of instruction on L2 speech processing. *AILA Review*, 40-67.
- Plough, I. & Gass, S. M. (1993). Interlocutor and task familiarity: Effect on interactional structure. In G. Crookes & S. Gass (Eds.). *Tasks and language learning: Integrating theory and practice* (pp. 35-56). Cleve-don, Avon: Multilingual Matters.
- Prabhu, N.S. (1987). *Second language pedagogy: A perspective*. Oxford: Oxford University Press.
- Richards, J. C. (1985). *The context of language teaching*. London: Cambridge University Press.
- Rivers, W. M. (1981). *Teaching foreign language skills* (2nd ed.). Chicago: University of Chicago Press.
- Roach, P. (1991). *English phonetics and phonology: A practical course*. London: Cambridge University Press.
- Rocha, Y. F. (2005). Promoting oral interaction in large groups through task-based learning. *Profile: Issues in Teachers' Professional Development*, 6, 73-81.
- Rutherford, W. E. (1987). *Second language grammar: Learning and teaching*. New York: Longman.
- Sambrook, R. (Director). (2001, May 1). *BBC News* [Television broadcast]. United Kingdom: The British Broadcasting Corporation.
- Searle, J. R. (1970). *Speech acts: An essay in the philosophy of language*. Cambridge: Cambridge University Press.
- Seedhouse, P. (2005). "Task" as research construct. *Language Learning*, 55(3), 533-570.
- Selinker, L. (1972). Interlanguage, *IRAL*, 10(3), 209-231.
- Skehan, P. (1996). Second language acquisition research and task-based instruction. In J. Willis & D. Willis (Eds.). *Challenge and change in language teaching* (pp. 17-30). Oxford: Heinemann.
- Skehan, P. & Foster, P. (1997). The influence of planning and post-task activities on accuracy

- and complexity in task based learning. *Language Teaching Research*, 1(3), 185-211.
- Skehan, P. & Foster, P. (1999). The influence of task structure and processing conditions on narrative retellings. *Language Learning*, 49(1), 93–120.
- Skehan, P. (2003). Task-based instruction. *The Language Teacher*, 36, 1–14.
- Stern, H. H. (1992). *Issues and options in language teaching*. Oxford: Oxford University Press.
- Tarone, E. & Yule, G. (1989). *Focus on the language learner*. Oxford: Oxford University Press.
- Ur, P. (1981). *Discussions that work*. Cambridge: Cambridge University Press.
- Van den Branden, K. (Ed.) (2006). *Task-based language education: From theory to practice*. Cambridge: Cambridge University Press. (The Cambridge applied linguistics series).
- Widdowson, H. G. (1990). *Aspects of language teaching*. Oxford: Oxford University Press.
- Willis, J. (1996). *A framework for task-based learning*. Essex: Longman.
- Willis, D. (1990). *The lexical syllabus*. London: Collins Cobuild.
- Willis J. & Willis D. (Eds.) (1996). *Challenge and change in language teaching*. Oxford: Heinemann.
- Willis, D. & Willis, J. (in press). *Doing task-based teaching*. Oxford: Oxford University Press.

Appendix I

Ia – Sample of the first draft of the participants' essays

(1)

***Would you go out on the streets for standing up for your rights?***

*One thing which is not usual today is to claim for your rights. Long time ago a great deal of people had no idea about what they should do in order to have your rights meet. But there are many reasons that make someone to do it, chiefly nowadays where most of people live in a misering condition. The want of money, which is closely related with the poverty of people, shows us the real picture of our society. If we give a general outlook on what is going on in all over the world we'll find out many other things that really lead people to claim for their rights.*

*Why there are many thieves hanging around the city? The first answer would be "Well they have no jobs", or "they don't want to get anything done", or whatever. But that's their way to demand what they want, as the government doesn't give any opportunity to them, or we could say it's the system which runs their lives and they don't have any conscience about it. We can consider a way of protest, a little bit strong, of course, but that's a way anyway. And there are even more reasons that suddenly raises from the deep of your core and makes you think hard and twice about how you are living lately. If you are really satisfied with the things you deal every single day. That's a kind of claiming that is there, patiently waiting for your move. That's a sort of thing that demands gumption in order to claim. We all must stand up for our rights in any circumstance, just because we need to be ourselves, just because we need a change all the time, just to feel that something in the world makes any sense.*

*This is what people are really doing when they go on strike, when they demand a better wage. Everything has to do with the rights of human being, because people want to lead a better life and with that to build up a better future.*

**(2)**  
**Is there a Road Protesting without violence?**

*Talk about this "question" is not an easy task, but it's very interesting because there are some important points and situations for discussing as the organization, the leadership and the road scheme.*

*About the organization the main points to note are that the organizers should think very carefully about the proposal and also have confidence to attracting a larger number of people. Furthermore, make plans, divide activities among people, make contacts and others.*

*Therefore, the leadership is an important person who will conduct the group. This person should have sufficient self-control not to be provoked and sometime, should be very strict in order to hold the demonstration. It's a very difficult position but it's necessary.*

*Then, in the road scheme, the group decides which strategy or methodology the protesting will follow. In this case, all the aspects and points are analyses and as a result, they build a map with some important considerations to prevent serious public disorder, damage to property, traffic congestion and blocking shopping streets.*

*Thus, if the protesters decide to follow these procedures, all actions and demonstrations will be non-violent and nobody will get hurt. In practice, the facts don't happen like this. In the majority of the time, there is violence in different kind of road protesting and sometime, a small number of people go to prison or die.*

*To conclude these points, I would say that, on all situations people should have respect each other and protesting without violence, is a good way of respect.*

Ib – Sample of the final version (edited version) of the participants' essays

(1)

*you should stand up for your rights.*

*Long time ago a great deal of people had no idea about what they should do in order to have their rights met. But it seems that things have not improved so far, as people are not used to claim for their rights today. However, there are many reasons that can make someone fight for his rights, chiefly nowadays as most people live in a miserable condition. The want of money, which is closely related with people's poverty, shows us the real picture of our society. If we look hard and twice to what is going on all over the world we'll find out many other things that really lead people to claim for their rights.*

*Why are there many thieves hanging around the city? The first answer would be "Well they have no jobs", or "they don't want to get anything done", or whatever. But that's their way to demand what they want, as the government doesn't give them any opportunity, or we could say it's the system that runs their lives and they are not aware of their real role in the society. We can say that it is a strong complaint on which they show they do not really agree with the situation they are going through. And there are even more reasons that suddenly might come from the deep of your core and make you think hard and twice about how you lead your life lately. If you are really satisfied with the things you deal with every single day, you'll probably won't have any reasons to fight the good fight, but when you least expect, there comes something, a kind of troublesome which will always be there patiently waiting for your move. That's a sort of thing that demands gumption in order to make any change in the journey through your life. We all must stand up for our rights in any circumstances just because we need to be ourselves, just because we need a change all the time, just to feel that something in the world makes any sense.*

*This is what people are really doing when they go on strike, when they demand a better wage. Everything has to do with the rights of human nature, because people want to lead a better life and with that to build up a better future.*

(2)

***Is there Protesting in the street without violence?***

*Talking about this "question" is not an easy task, but it's very interesting because there are some important points for discuss such as: the organization, the leadership and the road scheme.*

*About the organization the main points to note are that the organizers should think very carefully about the proposal for attracting a larger number of people. Furthermore, they should make plans, divide activities among people and be in contacts with share. Therefore, the leader is an important person who will conduct the group. This person should have sufficient self-control not to be provoked and sometimes, should be very strict but to be respected in order to hold the demonstration. It's a very difficult position but a person could deal with this situation.*

*Then, in the road scheme, the group decides which strategy or methodology the protesting will follow. In this case, all the aspects and points have to be analysed and as a result, the members of the group have to build a map with some important considerations to prevent serious public disorder, which means damage to property, traffic congestion and blocking shopping streets.*

*Thus, if the protesters decide to follow these procedures, all actions and demonstrations will be non-violent and nobody will get hurt. Actually, the facts don't happen like this. In the majority of the cases, violence exists in different kinds of protesting of the streets and often, as a consequence of this fact, a small number of people go to prison or even die.*

*To conclude these points, I would say that, all situations depend upon resignation, tolerance, patience and free will, then people should have respect for each other. Protesting without violence is a good way to be successful.*

**Appendix II: Worksheet 5**

**Spot the error**

**Correct the errors (if any) in each of these sentences. Be prepared to justify your answers.**

1. 'We could test our listening skill real English.'
2. 'It would be good if the quality of the next is much better.'
3. 'It gave us the opportunity to see how native speakers express themselves in a kind of situation like this.'
4. 'It also increased our knowledge about British culture and their law.'
5. 'We should have more contact with the video.'
6. 'He showed us how to improve and how to learn with our own mistakes.'
7. 'The book uses some uncommon words.'
8. 'I surprised with myself.'
9. 'I was so surprised with myself.'
10. 'We did the composition using a first draft and a second draft.'
11. 'After record, we could listen to our own voice.'
12. 'Another positive point about the interview is that we improved our pronunciation.'



## **English-teaching in Elementary Schools in Japan: A Review of a Current Government Survey**

**Daniela Nikolova**

*Department of English Teaching, Graduate School of Education  
International Christian University, Tokyo, Japan*

### **Bio Data:**

Daniela Nikolova graduated from Sofia University with an MA in Japanese linguistics (1998). She is a holder of an MA degree in Education from Tokyo Gakugei University (2004) and a Ph.D. degree in Education from the International Christian University in Tokyo (2007). Between 2003-2006, she taught English at various public elementary schools in Tokyo as an instructor with Japan Association of Teaching English through Images. She is currently working at the Embassy of the Republic of Bulgaria in Tokyo as a Culture, Press and Education Counselor.

### **Abstract**

In this article, I present a critical overview of the current situation in English-Teaching in public elementary schools in Japan whilst making some productive suggestions based upon contemporary research and social trends. I review briefly the history of English language education in Japan – previous and current methods of teaching and approaches, and I point out the challenges that the most recent English Language programs face in current school curricula. The conclusions I give are based on the results seen in the ‘Period of Integrated Studies’ classes as well as responses to questions in a substantial nationwide questionnaire conducted by the Ministry of Education, Culture, Sports, Science and Technology. This survey was in Japanese public elementary schools; with responses from students, homeroom teachers and parents. I also discuss possible solutions to the problems that current English Activities programs in elementary schools encounter and make suggestions for radical changes in the English curriculum.

### **Introduction: The Tragic Sense**

Japan has a relatively long history of English language education. Its level of linguistic comprehension, however, cannot compare to other Asian countries where English has been introduced into the Elementary School Curriculum like Singapore, the Philippines or South Korea.

The English language has become a very important part of every 21<sup>st</sup> century person’s life. English speaking skills have become a common requirement for jobs, and are considered to be a ‘ladder’ to a successful career. Meanwhile, Japanese words are coined every day from

English and in some cases, mostly in specialized areas people need to know English to understand some of the newly-emergent terms. A great part of the research in this field has been conducted by Japanese native scholars. Thus, studies, theory and criticism may have lacked a multi-perspective dimension. There is the ritual connotation that English is a 'difficult language' and no matter what solutions to the learning problems have been suggested, tragic feelings are continually expressed how Japanese cannot cross the psychological barrier of learning a foreign language.

Recently (2002), the Ministry of Education, Culture, Sports, Science and Technology (MEXT) in Japan decided to introduce English language education in elementary schools. However, the curriculum does not recommend a compulsory number of lessons per school year/school age, etc. There is insufficient research on the current system and there are no criteria for evaluation of the on-going program.

MEXT conducts annual surveys in all public elementary schools in Japan to gather information to fully comprehend and stay updated on the problems which schools, teachers and parents are facing; however, it seems that no one uses the data to try to improve the situation. Ironically, MEXT also lacks sufficient specialists in English studies who can use English fluently. The Ministry seeks help from foreigners but this is limited to translation or other linguistic problems, not related to curriculum development and planning, or goal-setting for criteria of the English language programs like in other Asian countries' (Singapore, the Philippines, South Korea, Hong Kong) education systems.

This paper consists of three main sections: A brief history of English education in Japan, analysis of the English activities at public elementary schools and conclusions which address future tasks and possible approaches. The first section outlines the history of English education in Japan. The second section reviews a government survey from 2004 about English activities at elementary schools and presents an adaptation of the data. The third section is an attempt to point out some ways to improve the existing English education programs at elementary schools as well as give suggestions to the curricula developers.

### **A brief history of English language education in Japan**

This section consists of 3 subsections. In the first subsection I give the background of the English language education in Japan; in the second subsection I present a brief history and in the third subsection I focus on the public elementary schools in Japan and the existing English language programs.

### *Background*

English language education in Japan is a major topic of discussion since many Japanese people consider it their weak point and they try to improve their skills through every possible means. There are numerous textbooks, CDs, DVDs, books on how to more effectively teach and (self-) study English. A great part of the research done on foreign language learning and curricula development discusses why Japanese cannot learn English with the ease that other nations do. Seminars and congresses are organized to clarify how different English is to Japanese and what the biggest challenges to a Japanese native learner are. Articles criticizing the language education programs used in the past are giving directions, or outlining policies for improving and establishing a more effective curriculum.

In Japan, there are more people who have six years of English at school than those who have not. However, the learning outcome has been markedly inadequate. People have not developed proficiency in English as a language of international communication. After many years of English language education they are not able to converse on simple/everyday topics. However, even though some of them have no speaking nor listening skills, some learners can, to some extent, write adequate texts as well as read; yet, in terms of speaking, sometimes a single greeting can be too hard to achieve.

### *A Brief history of English language education in Japan*

English was first taught in Japan following the brief stop in Nagasaki of a British ship in 1808. In order to be able to communicate with the sailors from the ship, the Tokugawa government ordered the interpreters to add the study of English to their study of Dutch, French, and Russian languages (Buruma, 2003; Omura, 1978).

In 1853, Japan was forced to open its frontiers by the Americans, and the study of English began in earnest. Two major methods emerged early in the history of English teaching in Japan. One emphasized correct pronunciation as well as meaning, and the other emphasized meaning, not pronunciation or syntax. It was observed that students who learned by the second method seemed to have better comprehension, while students who learned by the first method were compared to “reading machines” with correct pronunciation but with little understanding of what they were reading (Omura, 1978).

In 1871, the Ministry of Education was established and 19 years later a system of language teaching was introduced in the country. The study of English became compulsory in the middle and higher secondary schools. Foreign language study was closely associated with Westernization. Native English speakers, mostly Christian missionaries, taught English.

English-medium classes were held at the institutions of higher learning, and most textbooks came from the U.S.A.

In the early 1900's, Japanese textbooks and teachers replaced the American textbooks and teachers. Many of the students sent previously to western countries returned to Japan and many academic books were translated into Japanese. English was no longer the means of access to knowledge of the western world (Imura, 2003). It became primarily a subject of study, mainly learned in order to be able to read written texts, instead of a means of communication (Kitao & Kitao, 1982).

In 1921, Harold E. Palmer, an English linguist and specialist in TEFL, was invited to Japan as an advisor to the Ministry of Education. He advocated the Oral and Direct methods, which were applied in various parts of Japan, and although good results were obtained those methods never came into general use since they required a command of English far beyond that of most Japanese teachers of English (Kitao & Kitao, 1982).

Before and during World War II, English was proclaimed the "enemy language", and the study of English was discouraged. However, this policy was quickly reversed after the war. The school system was reorganized to create six years of elementary school, three years of junior high school, three years of high school, and four years of college. Therefore, the time spent learning English was extended from three to six years as more and more students chose to go to senior high after junior high. In 1991, over 94% of the nation's 15-year-olds went to senior high (Japan Education Yearbook Publication Committee, 1985).

Back in 1970s, a Congressional representative named Wataru Hiraizumi, questioned the validity of the English language education in Japan at that time, and pressed the need for a more practical approach. It was immediately rebuffed by Shoichi Watanabe from Sophia University in Tokyo who advocated the need to teach a foreign language as an intellectual endeavor, and the famous debate continued on for several years, with neither side giving in.

However, the situation started to change with the advent of the Communicative Approach or Communicative Language Teaching, which in many ways answered the needs of contemporary globalization. In the early 1990s, MEXT announced the beginning of communicative course of study for foreign languages introduced in the high school curriculum.

Nowadays, foreign languages are offered as electives in the reorganized junior high and high schools, and the most common foreign language chosen is English. Since April 2002, English Conversation classes, aiming for 'International Understanding' are introduced at all public elementary schools in Japan (MEXT Guidelines, 1998).

*Public elementary schools and English language education*

In response to a heightened sense of crisis over Japan's lack of facilities with English in the internet age, Prime Minister Kenzo Obuchi released a report in January 2000 on "Japan's Goals for the 21<sup>st</sup> Century" that proposed making English the nation's official second language. "In order to achieve a world class excellence, all Japanese should acquire a working knowledge of English", the report said. It called for reorganizing the English classes according to level of achievement, not grade, improving teacher training, and contracting language schools to teach English (The Guardian, Feb. 23, 2000). To address these issues, a special committee was set up by the Minister of Education in January, 2000.

The increased internationalization of the global society as well as the constantly amplifying capacities of the Internet Network and the Multimedia led to the spread of English as a *lingua franca* in the world. It is no wonder that debates rage around the intensive study of English from an early age or its acceptance as a second officially used language.

In the meantime, the Japan Association for the Promotion of Foreign Language Association (*zenkoku gaikokugo kyoiku shinko kyokai*) was established and among its goals stated in the project plan for year 2004-2005 were the following:

- To promote understanding of the importance of foreign language education and to increase foreign language learners' motivation
- To take part in the campaign started by the Ministry of Education, Culture, Sports, Science and Technology 'The Promotion of Committee Activities for Children's Education'
- To conduct surveys on the actual conditions of language teaching institutions
- To conduct surveys on the needs of learners at the language teaching institutions

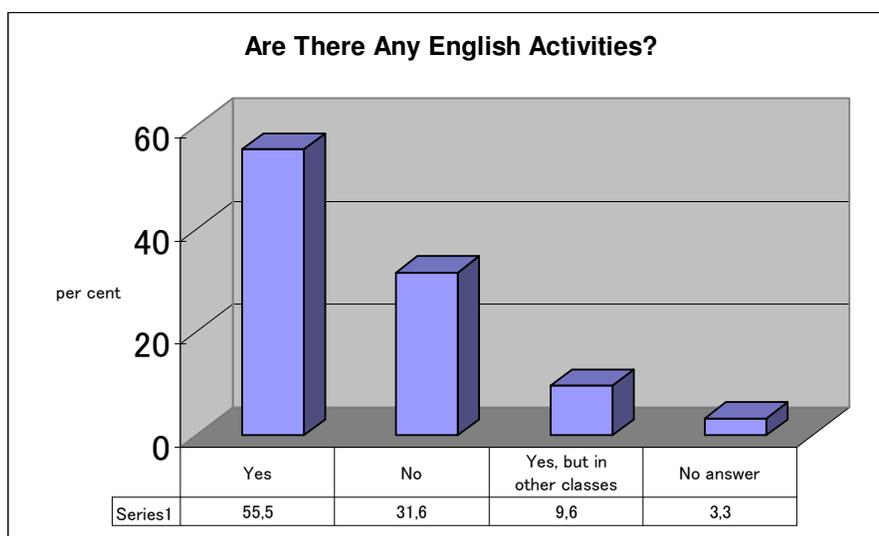
Four years passed since English was introduced as a foreign language taught in the Period of Integrated Studies classes at elementary schools. The current situation, reflected in the results from a survey obtained by Mainichi Shinbun (Mainichi Shinbunsha, 2005), is that the older the age of children, the fewer books they read. 'Entertaining' materials were not included in the category of 'books'. However, it is a fact that the older the children, the higher the number of comic books they read.

According to a survey from the National Language Institute (Shimamura, 1999), Japanese children hate writing composition. This might be interpreted as an expression of less creativity and imagination that can be observed in children nowadays. Currently, there is no reading and writing included in the English language teaching curriculum. A reason for this might be the fact that Japanese people try to overcome previous tendencies when the main

focus of teaching English was on reading and writing, and even after long years of study few people could actually speak English. Thus, MEXT tries to abolish and avoid everything that recalls the unsuccessful methods of foreign language education in the past, instead of reevaluating the ratio of different skills which should be taught at schools.

### **Analysis of the English activities at public elementary schools**

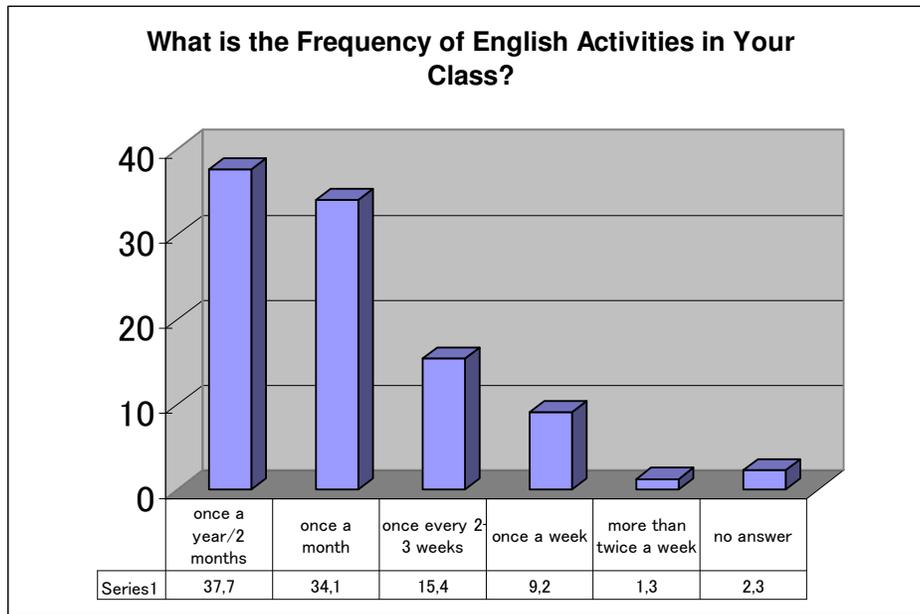
This section is a review of the current situation as presented in a nationwide public elementary schools survey of MEXT, 2004. It is a practice for every local Board of Education to assign a certain number of English lessons per school year to the public schools in their ward or village. However, that number is not uniform throughout the country. According to the questionnaire from June, 2004 conducted by MEXT in 230 elementary schools, 31.6% of the home room teachers questioned answered that English is not taught in their class, let alone their school. (See graph 1)



Adapted from: *An Attitude Survey on English Language Education in Elementary Schools*, Ministry of Education, Culture, Sports, Science and Technology, (June, 2004)

### **Graph 1 - Is English Taught in Your Class? (Teachers)**

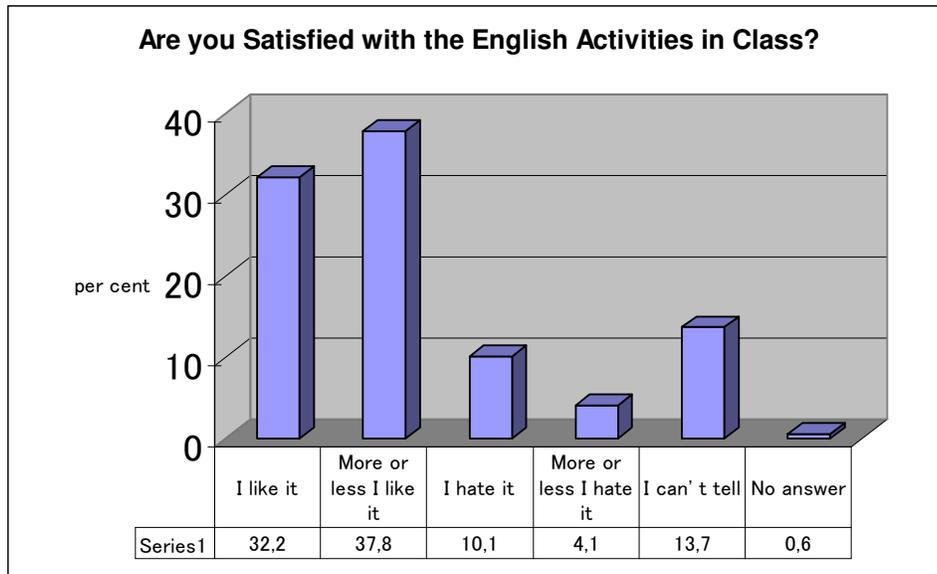
There is a discrepancy in the number of classes taught in different schools. The highest number (37.7%) is for 'once a year/ every two months', followed by 34.1% 'once a month' and 15.4% 'once every 2-3weeks'. (see graph 2)



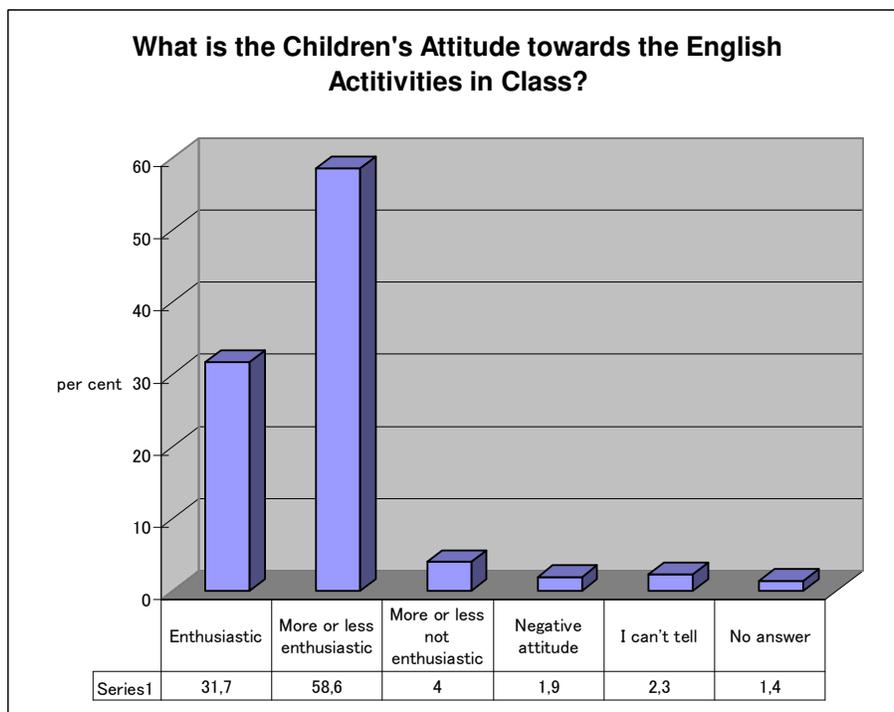
**Graph 2 - Frequency of English Activities at Elementary Schools (Teachers)**

The number of classes typically depends on the budget, which the Board of Education can work out for each school year. Each year, an English Study Research school is designated in each ward or area, where English is taught more frequently than the other schools. At the end of the school year, or throughout the year, the designated school gives a presentation on how to teach English to elementary school students or shows what students have learned throughout the year.

In general, children (here only results from 6<sup>th</sup> graders are considered, since their intellectual level is higher than the rest of the students questioned) are satisfied with the English language classes they take and their homeroom teachers confirm the statement with their assessment that the interest from students is quite high. (Children: 71.5% satisfied versus 14.3% who say they hate English) (Homeroom teachers: 90.3% see positive attitude, while only 5.9% see a negative attitude towards English) (see graph 3.1 and graph 3.2)

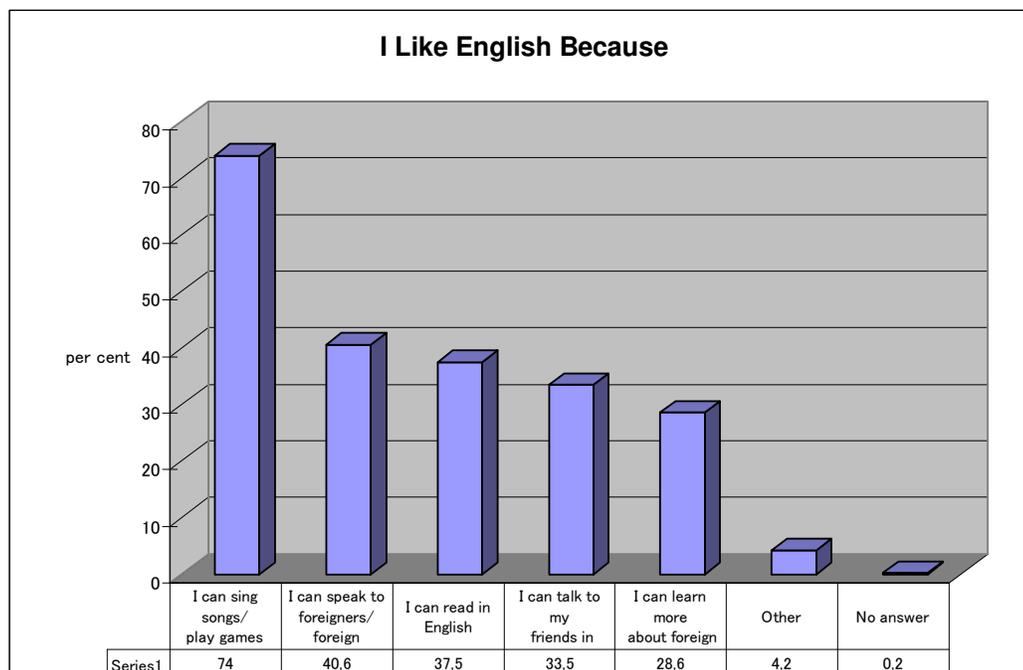


**Graph 3.1 - How Satisfied are you with The English Activities in Class? (6<sup>th</sup> graders)**



**Graph 3.2 - What is the Children's Attitude towards the English Activities in Class? (Teachers)**

6<sup>th</sup> graders enjoy English because: they can sing songs and play games in English (74%), they can speak to foreigners or a foreign teacher (40.6%), they can read in English (44.6%). (see graph 4)



**Graph 4 - Why do you Like English? (6<sup>th</sup> graders)**

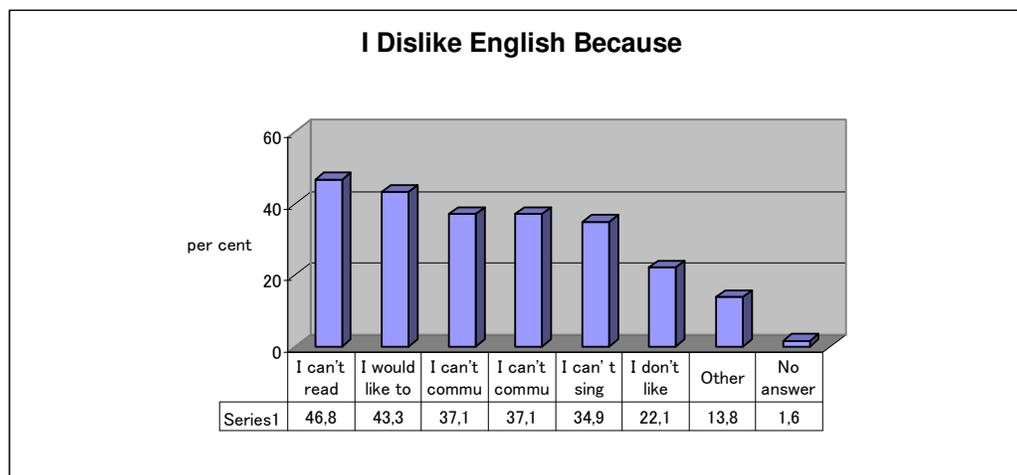
However, if we carefully read some of MEXT Guidelines (2005), we notice that the “Aims of Elementary School “English Activities” are:

to foster interest and desire - not to teach a language.

It further says:

Communication takes place primarily through verbal and written means. Nevertheless, it is too much to ask elementary school students to communicate using both means, and doing so may cause them to develop a dislike for English.

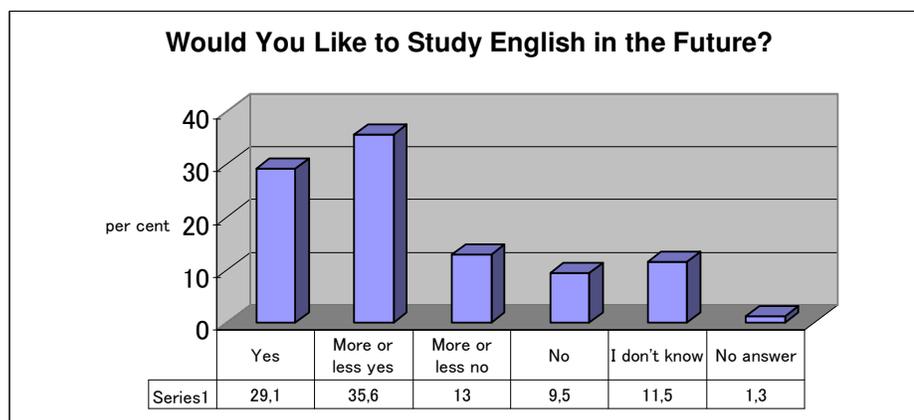
On the other hand, 6<sup>th</sup> graders ‘hate English’ because: they cannot read properly (46.8%), in the time allotted for the Period for Integrated Studies they would like to study other subjects (43.3%), and they cannot communicate properly with friends (37.1%) and with foreigners (37.1%) (see graph 5)



**Graph 5 - Why do you Dislike English? (6<sup>th</sup> graders)**

Building reading and even writing skills is considered inappropriate, just ‘too much’ for that age group (MEXT Guidelines, 2005) and MEXT aims to eradicate these skills from the curriculum. However, as we notice in the above graphs, the students actually have no resistance to the reading itself. Rather, they want to learn English because they will learn how to read English texts (reason mentioned in the top #3). Those who dislike English are, to some extent, related to the reading issue: they cannot read well, or they simply have no confidence in their reading skills. Note that it is not because reading is too high an aim for their intellectual level. This can be explained with bad teachings how to read English texts and inefficient approaches towards this matter.

Fortunately, there is a palpable future in the English teaching in elementary schools in Japan: the majority of 6<sup>th</sup> graders have answered that they would like to study English in the future (64.7%), versus less than one-third of the respondents who would not like to study English (22.5%). (see graph 6)

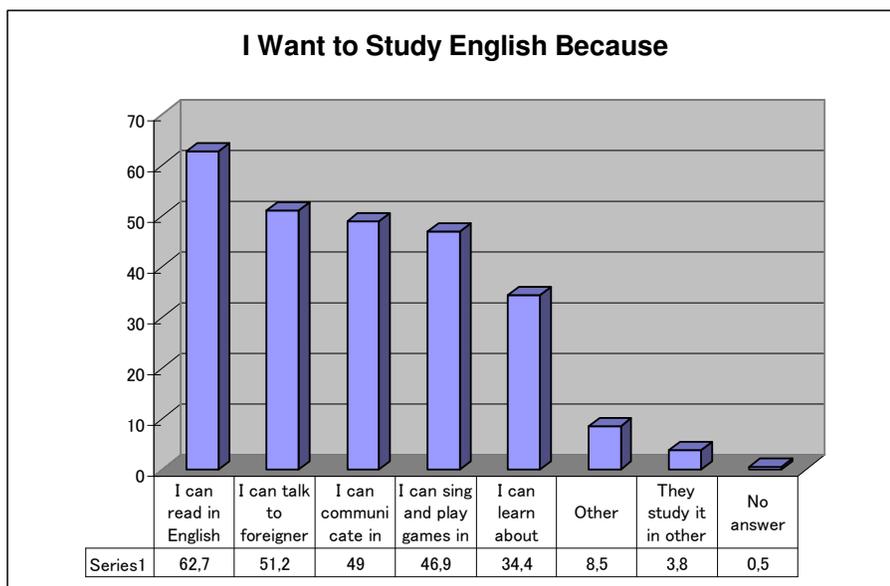


**Graph 6 - Would you Like to Study English in the Future? (6<sup>th</sup> graders)**

The top three reasons why the majority of children “want to study English in the future” are as follows: (see graph 7)

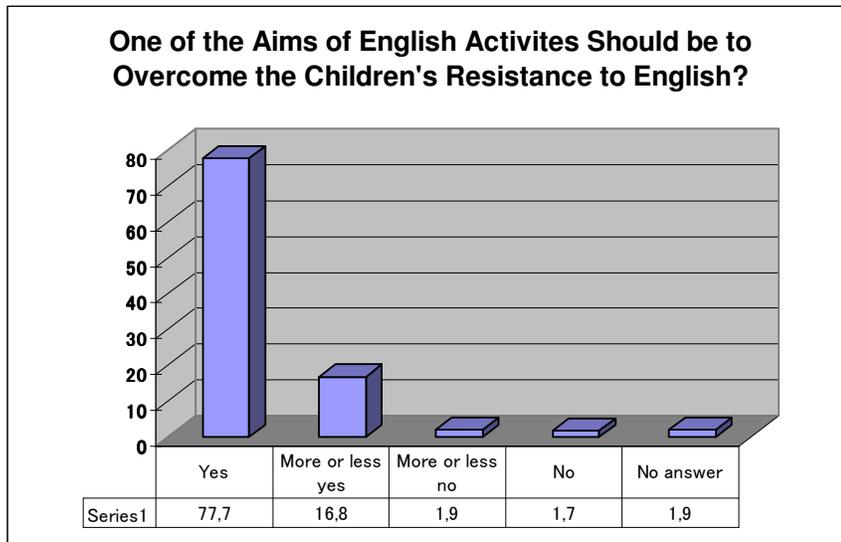
1. I can learn to read English (62.7%),
2. I can communicate with foreigners in English (51.2%),
3. I can communicate with friends in English (49%)

It seems that MEXT need not worry about implementing reading in the English-teaching curriculum or cultivating “a proper attitude towards using the language for communication” – children wish to read in English and communicate with foreigners (in English), these are the top 2 reasons why they want to study the language.

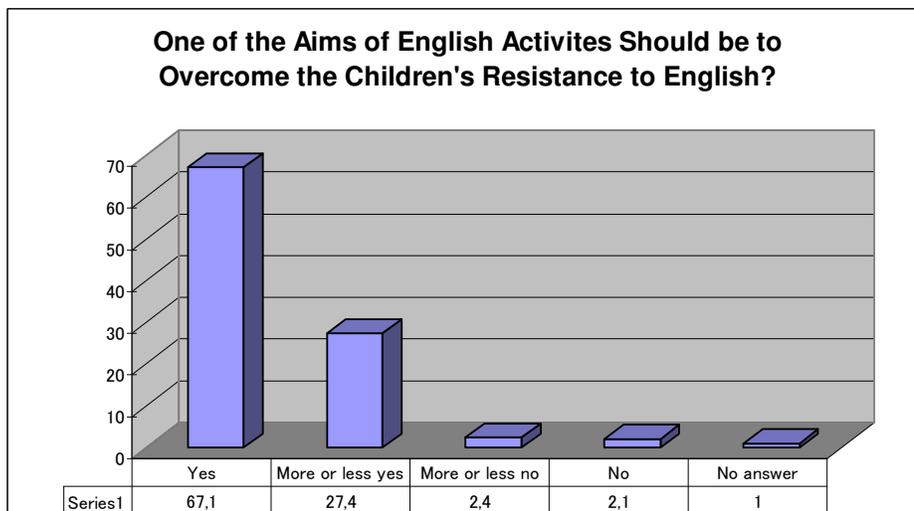


**Graph 7 - Why do you Want to Study English? (6<sup>th</sup> graders)**

The following is a question only for parents and homeroom teachers, concerning the children’s resistance to English. (see graph 8.1 and 8.2) When parents and teachers have such attitude towards a scholastic matter, they subconsciously ‘implant’ the idea of difficulty in the children themselves, thus they form an invisible language barrier in their minds. This is one of the most difficult factors that educators in Japan need to overcome or work on ways to diminish its effect on the overall foreign language education in Japan. However, this cannot be achieved through abolishing reading (and writing) in the classroom.



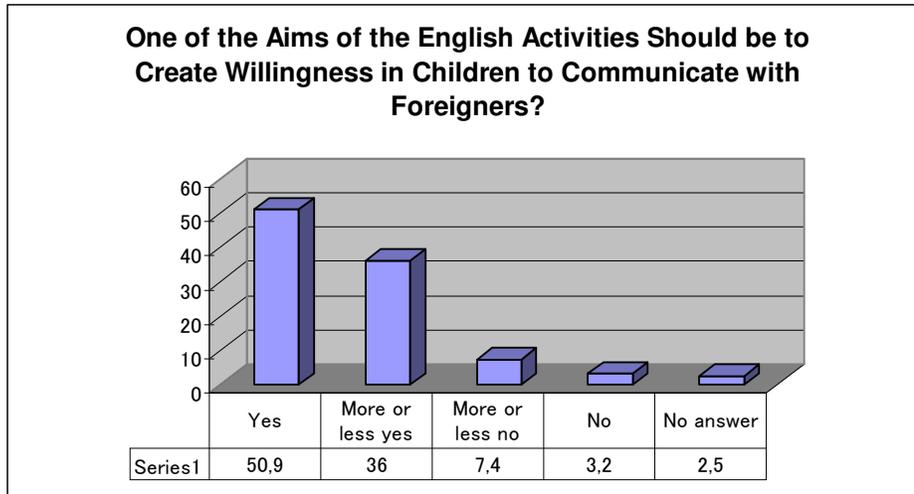
**Graph 8.1 - Aims of English Activities at Elementary Schools: to help children overcome their resistance to English (parents of 6<sup>th</sup> graders)**



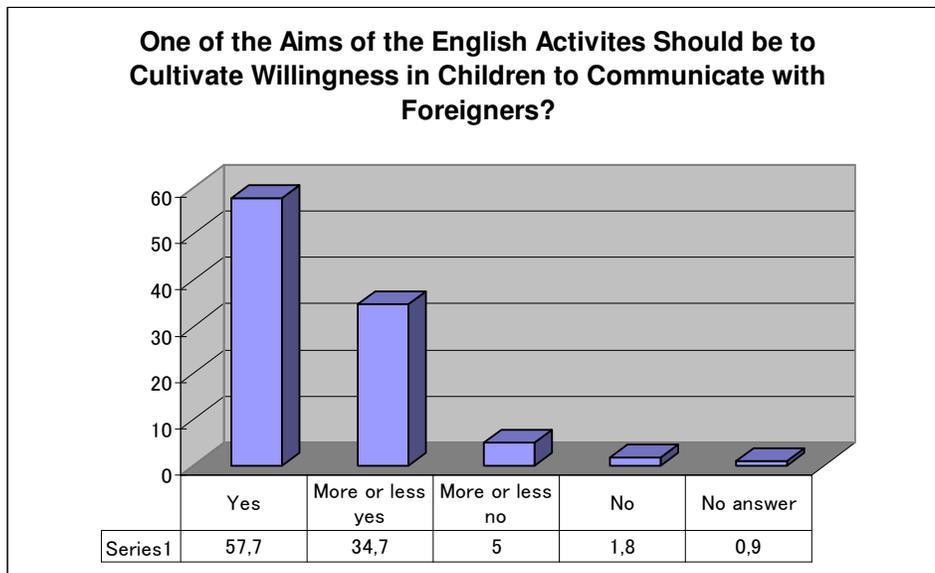
**Graph 8.2 - Aims of English Activities at Elementary Schools: to help children overcome their resistance to English (teachers)**

Another kind of answer (see graph 9.1 and 9.2) may lead us think that there is also a cultural problem in addition to the curriculum or educational problems. The fact is that parents and teachers claim that the aim for the future English education at elementary schools is to set such an attitude, and also that children will be willing to communicate with foreigners, and not stray from foreigners. Again if we refer to graph 7, we can notice that children have already showed willingness to communicate with foreigners – that is why they want to study English. Parents and teachers are afraid for their children even though they

don't need to; teachers of English these days should not let the cultural walls enclosing the parents, or their fear or lack of confidence to speak English, interfere with the education of the elementary schoolchildren. In other words, raising such issues seems unsustainable and groundless.



**Graph 9.1 - Aims of English Activities at Elementary Schools: to cultivate a proper attitude and willingness in children to communicate with foreigners (parents of 6<sup>th</sup> graders)**



**Graph 9.2 - Aims of English Activities at Elementary Schools: to cultivate a proper attitude and willingness in children to communicate with foreigners (teachers of 6<sup>th</sup> graders)**

Nowadays, popular culture affects children's life and language more than their parents or teachers (Nikolova, 2005). Partial representations of popular culture can be seen in the elementary school curriculum. TV (visual teaching) and comic books become part of the

teaching materials for elementary school children. Not only in the English language classes, but also Period of Integrated Studies and Social Studies do teachers resort to using visual materials. Such kind of materials, even not considered educational, might help children accept language as a means of communication or a tool with which they can learn many new things, a truth their parents could not understand when they were little.

### **Conclusions: Future tasks and possible approaches**

In conclusion, I suggest here that, in reflection of emerging social trends and published research, MEXT should reconsider its approach towards foreign (sic. English) language teaching in Japan. The Ministry has started its attempts for collaboration between native and non-native teachers in the EFL teaching in elementary schools, as seen in Hong Kong and other Asian countries (Carless, 2006). What is more important is that MEXT should wholeheartedly incorporate the cultural specificities of the contemporary lifestyle and tradition in Japanese society into its pedagogical policy. These trends are quickly absorbed by elementary school children, and especially 6<sup>th</sup> graders, a transitional age group which is the most susceptible to absorbing linguistic features of a foreign language and forming a proper cultural attitude towards that language. This same cultural attitude, teachers try to create and cultivate desperately in all schools in Japan (MEXT, 2005). We may enumerate our recommendations.

Firstly, MEXT should consider a uniform number of English Activities lessons throughout the country. This is one of the factors for achieving some of the results set in the English Activities guidelines.

Secondly, MEXT should consider approving a certain number of English textbooks for all elementary schools in Japan (especially 6<sup>th</sup> grade) written either in English or Japanese. There is a need for more elaborate research on the positive effects of the printed forms of popular culture on children's education, especially the reading aspect of language acquisition, as well as conducting experimental studies on those forms as a means of teaching difficult topics through 'entertaining and non-serious' methods. Reading gives a choice to 6<sup>th</sup> graders to decide what they need to learn. This provides them with a sense of autonomy and a chance to reflect upon their progress, two very important factors to children's learning process (Scott et al, 2004). The need to create curricula including reading and writing arises naturally; implementing comics' linguistic features and their role as a language facilitator might help; however further analyses are needed.

Thirdly, MEXT should consider assigning English Activities classes to younger age

homeroom teachers, who do not ‘fear’ English and who do not need any special attitude training and have no resistance to English or communicating with foreigners.

Lastly, if the introduction of reading and writing skills in the current curricula cannot be made immediate, MEXT should consider such at first for the English studies Research schools only, thus providing a ground for further analyses of the specifics of reading and writing in English, the cultural ambience at elementary schools and the feelings of teachers, parents and students.

It is very important that the Ministry of Education, Culture, Sports, Science and Technology in Japan focuses on addressing the needs of students (especially 6<sup>th</sup> graders) not only to learn a new language, but also a new culture and ways of socializing into that culture. If the new culture is too ‘foreign’ to children, and if socializing in that culture seems unnatural to children, it may be a good idea to use some intrinsic methods of teaching it that could help build reading skills in 6<sup>th</sup> graders, as well as nurture confidence in using any foreign language as a means to acquire universal knowledge about the world. The *bona fide* language classroom is, in fact, located both within and beyond the classroom walls: a geography lesson that is frequently overlooked.

English Activities programs are now being introduced by the Ministry of Education, Culture, Sports, Science and Technology in the school curricula of public elementary schools. This remains controversial in various aspects: the wider social consensus, implementation, teacher support, curriculum and syllabus. It has been argued in this paper that MEXT – which assumes responsibility for educational steerage and policy - should pay close attention to the cultural ambience of children. This includes reading habits, tendencies and natural likings for symbols or codes (i.e. the English alphabet). It is argued also that pupils and their teachers adopt a more rigorously ‘international’ stance in order to develop a constructive and forward-looking attitude to the matter of foreign language learning and teaching.

## **References**

- A Survey on Reading Habits, 50<sup>th</sup> Survey on Readings at School Age. (2005). *Tokyo: Mainichi Shinbunsha*, 70-73.
- Buruma, I. (2003). *Inventing Japan 1853-1964*. New York: Modern Library.
- Carless, D. (2006). Collaborative EFL teaching in primary schools. *ELT Journal*, 60(4), 328-335
- Fujimoto-Adamson, N. (2006), “Globalization and history of English education in Japan”, *Asian EFL Journal*, 8(3), Retrieved March 19, 2007 from:

[http://www.asian-efl-journal.com/Sept\\_06\\_nfa.php](http://www.asian-efl-journal.com/Sept_06_nfa.php)

- Imura, M. (2003) *English education in Japan for the past 200 years*. Tokyo: Taishukan Shoten.
- Japan Association for the Protection of Children (1984). *White paper on children*. Tokyo: Sodo Bunka.
- Kitao, K. (1983, February 17). Common misconceptions about “reading English”. *The Daily Yomiuri*, p. 7.
- Kitao, K., & Kitao, S. K. (1982, September 16) *College reading textbooks do not meet needs*, *The Daily Yomiuri*, p. 7.
- Kitao, K., & Miyamoto, H. (1982). Japanese college students' English reading ability—a study based on testing. *Doshisha Studies in English*, 30, 135-165.
- Kitao, K., & Miyamoto, H. (1983). Japanese college students' problems in reading English - Tendencies and changes in errors. *Doshisha Studies in English*, 32, 118-147.
- Ministry of Education, Culture, Sports, Science and Technology. (2004). *An Attitude Survey on English Language Education in Elementary Schools*.
- Ministry of Education, Culture, Sports, Science and Technology (2005). *Practical Handbook for Elementary School English Activities*, 123, Tokyo: Kairyudo Publishing.
- Nikolova, D. (2005). *The Language World of Elementary School Children in Japan*, Unpublished doctoral dissertation, International Christian University, Tokyo
- Nogami, M. (1978). Elementary school: The importance of TEFL for children. In I. Koike (Ed.), *The teaching of English in Japan* (pp. 141- 155). Tokyo: Eichosha
- Omura, K. (1978). History of English teaching in Japan 1 Prewar (before 1945): From the Phaeton Incident up to the Pacific War. In I. Koike (Ed.), *The teaching of English in Japan* (pp. 104-114), Tokyo: Eichosha
- Oura, Y., & Tada, M. (1978). Adult education: Student attitudes toward English language learning at Kyoto English Center. In I. Koike (Ed.), *The teaching of English in Japan* (pp. 156-168). Tokyo: Eichosha.
- Scott, C., Thoghdha, M.N., Smowton, D., Bergner, G. and Williams, M. (2004). Survey: Primary Courses. *ELT Journal*, 58(3), 277-293
- Shimamura, N., (1999). On children's linguistic problems: Disliking Japanese language. *Journal of Sociolinguistics*, 2(1), 21.



## Book Review

### *Assessing Language through Computer Technology*

Carol A. Chapelle and Dan Douglas. Cambridge, U.K.: Cambridge University Press, 2006.

Pp. xi + 138.

Reviewed by Deepti Gupta

Panjab University

Chandigarh, India

In *Assessing Language through Computer Technology*, Chapelle and Douglas examine the complete paradigm of CALT or computer assisted language testing and take the reader on a voyage with the technology thread that runs through the process of assessment starting from 1985.

Divided into six chapters, this handy volume scans most of the CALT related issues in this era. The first chapter is appropriately titled “The Technology Thread.” Running into 20 pages, it describes the various forms in which computers have been used in assessment. Using instances of CALT from across the board, it makes a persuasive case for using computers.

The second chapter goes on to perform a comparative study of assessment with and without the use of computers. The authors take up each test method characteristic beginning with physical and temporal test circumstances and leading up to input and response to examine the whole range to determine the ways in which CALT can affect these characteristics. The argument is that the use of computers does improve the efficacy of tests and examples are taken from all over the world to persuade the reader.

The next chapter deals with the ways CALT has emerged as a threat. While chapter 2 had briefly touched upon two validity issues pertaining to the inferences drawn from CALT test scores and the ways in which these scores can be used for purposes such as certification and admission decisions, this chapter expands these two concerns into six areas that are often brought up as potential threats to the whole practice of CALT. After examining the full range of threats from different test performance to negative consequences, the concluding message is that any mode of testing would face such issues and that the threats related to validity need further discussion, one which is promised in chapter 5.

Chapter 4 delineates the actual implementation of CALT, the authoring tools available and how to use them. Taking as an example the test for students entering a graduate level applied linguistics program, the authors go through all the steps required to implement CALT, the steps of test creation, publishing on the web, student access, test taking, and viewing examiner records using *Respondus* software. Other software tools are also clearly described to give readers ample choice. The chapter goes on to describe various other computer-based tests, taking the trouble to explain how and why they were designed as well as their strengths and weaknesses. It also mentions the tools that are still forthcoming.

Chapter 5 gives an overview of the various historical evaluations of CALT made by different researchers and the detailed evaluation that was promised in chapter 3 now takes place as each concern is given careful analysis, especially the concerns dealing with test validity. The concluding chapter inserts a note of caution because the authors feel that “in second language assessment, despite the significant changes and advances made through the use of technology, the revolution portrayed by Bennett has not yet occurred” (p. 103). Thus, briefly touching upon the earlier chapters, the authors state that they see CALT not so far as an evolution but rather as a revolution: “A revolution may be coming sometime in the future, but in the meantime, . . . the changes brought about by technology intersect in important ways with other areas of applied linguistics” (pp. 106-7). The focus then shifts to the different areas of applied linguistics: how technology affects language ability and use, in what ways it could influence SLA research, and how it could alter the complete paradigm of language teaching. Here, the chapter and the book end on a note of hope that in future CALT would draw upon input from various disciplines and emerge stronger as a result of this interdisciplinary contribution.

The text is definitely an asset to the *Cambridge Language Assessment* series, is sure to open new avenues for applied linguists, encourages thinking out of the box, and should be, as the series editors Aldersen and Bachman conclude in their preface, “required reading for any test developer” (p. xi). With this in mind, although it may be a bit premature and quite daunting for assessors who function in parts of the world where traditional manners of assessment are not yet well executed, *Assessing Language through Computer Technology* can be a very useful resource for those who have these tools at their disposal.



## Book Review

### *Literature and Stylistics for Language Learners: Theory and Practice*

Greg Watson and Sonia Zyngier (Eds.), Basingstoke: Palgrave Macmillan, 2007. Pp. xxiii + 217.

Reviewed by Vander Viana

Catholic University of Rio de Janeiro

Rio de Janeiro, Brazil

According to Carter (2007), it is possible to identify three historical stages in the teaching of English as a foreign language leading up to today. The first one dates from the beginning of the 20<sup>th</sup> century when literary texts were used as models for writing. In the 40's and 60's, literature left the limelight as a more functional approach to language teaching came to the fore. The third period arrived with the communicative approach in the 70's and literary texts regained prominence in TEFL as they were now seen as authentic materials and as ways of helping students deal with more imaginative texts. It is in this scenario that *Literature and Stylistics for Language Learners: Theory and Practice* has been launched.

The book, containing 15 chapters written by scholars representing most continents, has a major pedagogical concern--it focuses on the teaching of stylistics to speakers of English as a first language, as a second language and/or as a foreign language (L1, L2, or FL)--and is divided into five parts. In the first part, "Theoretical Perspectives," the theory which underpins stylistics is offered. Here, Geoff Hall comments on the role of stylistics in the teaching of English as a second language and Peter Stockwell argues in favor of the teaching of literature as such.

The second part, "New Approaches," opens with Joanna Gavins and Jane Hodson's account on their experience of setting up a stylistic course for third-year undergraduates at the University of Sheffield. John McRae goes on to discuss the narrative point of view in three literary works. Afterwards, Rocio Monteiro expands on the relationship between literature and cinema. Finally, Urszula Clark reports on her work with detective fiction.

"Corpus Stylistics" is the title of the third part in which those interested in corpus linguistics can see how its tools may be applied to stylistic analysis. Donald E. Hardy begins by probing the fiction of Flannery O'Connor by means of a text-analytic computer program.

In the next chapter, Bill Louw discusses how collocations may be used in order to identify the semantic prosodies of specific lexical items. In the last chapter, Mick Short, Beatriz Busse, and Patricia Plummer comment on their experience regarding the teaching of *Language and Style*, a web-based course on stylistics.

The fourth part, “Stylistics, Grammar and Discourse,” also presents another study on the work of Flannery O’Connor. Different from Hardy, David L. Gugin, however, focuses on pseudo-cleft sentences and makes a link between their use and the activation of students’ schema in the Persian Gulf. In the chapters that follow, Paul Simpson investigates the use of the Hiberno-English Emphatic Tag, and Judit Zerkowitz explores Gricean maxims.

The last part of the volume, entitled “Awareness and Cognition,” expands on empirical work. David Hanauer begins by investigating the differences between explicit and implicit educational approaches to the teaching of literature to secondary students in Tel Aviv. In the following article, Willie van Peer and Aikaterini Nousi observe whether there are differences in reading and in reading and discussing texts with relation to reducing negative stereotypes. Finally, Sonia Zyngier, Olívia Fialho, and Patrícia Rios reappraise the concept of literary awareness and apply it to a Brazilian context.

This volume certainly offers a vast array of studies from different theoretical orientations which can be easily applied to the language classroom. The only shortcoming seems to be the title of the volume which does not cover all the text has to offer. A much better description, however, is given in the preface to the book where the editors state that the collection may be of interest to teacher trainers, teachers, arts instructors, educational administrators, and lecturers. Nevertheless, on the whole, this volume is a collection which surely contributes to pedagogical stylistics by helping teachers conduct such type of work in a language classroom--be it a L1, L2, or FL environment--and looks ahead into future developments in the area.

## **Reference**

Carter, R. (2007). Literature and language teaching 1986-2006: A review. *International Journal of Applied Linguistics*, 17(1), 3-13.



## Book Review

*Practical English Language Teaching: Speaking*

Kathleen M. Bailey. New York: McGraw-Hill, 2005. Pp. vii + 199.

Reviewed by Aysha Viswamohan and Hannah M. Sanala

Indian Institute of Technology, Madras

Chennai, India

Designed for teachers who “may or may not have had formal training in second and foreign language teaching methodology” (p. vi), Bailey’s *Speaking*, the fifth volume of McGraw Hill’s *Practical English Language Teaching* series, is indeed a valuable resource for both beginning and experienced teachers in the profession who wish to grow as teachers in the area of speaking.

Eminently readable, the text complements the core (and first) volume of the series, *Practical English Language Teaching*, by delving deeper into the teaching of speaking than was possible in the first volume to acquaint experienced teachers with “current theoretical and practical approaches to teaching speaking” (p. vi) and is effectively divided into three sections: Introduction, Teaching Speaking at Different levels, and Key Issues in Speaking and Pronunciation.

The introduction, which comprises chapter 1, lays a foundation and provides the terminology for the rest of the text. After providing an in-depth definition of speaking, the author sets the historical background of the many approaches outlined in the book. This is followed by a methodical yet uncomplicated elucidation of each approach--the grammar translation method, the direct method, audiolingual method, communicative language teaching, and communication strategies--and concludes with counsel on the subject of assessment.

Chapter 1 is followed by three chapters that form the second section of the book, each of which begins by providing goals for the reader to focus on and then, using extracts and examples to illustrate the topics, offers step-by-step guidance on the teaching of speaking: syllabus design, principles of teaching, tasks and materials, teaching, speaking, and assessment, respectively. The chapters also clearly demarcate these subjects as they apply to

three types of learners (beginner, intermediate, and advanced): general topics, pictures and sounds in chapter 2, more complicated activities using logic in chapter 3, and formal use of English in chapter 4. Each chapter also provides activities which are paced and detailed to suit the needs of the particular kind of learners and supplies examples in highlighted boxes to make the ideas explicit.

Chapter 2 introduces the reader to the goals that need to be set for beginner-level students and to the concepts of lower-level language, English as a foreign and second language, and Nunan's formula of memorizing formulaic expressions that need not be broken down into grammatical elements, for example, "A little more slowly please" or "How do you say . . . ?" Here, Bailey explains the importance of providing palpable topics, opportunities for group and pair work, communicative language teaching, and the physical arrangements of the classroom such as the inside-outside circle, tango-seating, and cocktail party. She also outlines speech events and activities suitable for the beginner-level learners: Guided conversation, interviews, information gap, jigsaw activities, scripted dialogues, drama, role-playing, logic puzzles, picture-based activities, and physical actions in speaking lessons. In the process of describing these, she draws attention to some important terminology. This chapter also focuses on the visual component of communication and on the phonetic nuances of the English language.

Chapter 3 addresses teaching intermediate level learners. In this chapter, Bailey, presents her ideas on the important points of syllabus design, principles of teaching, and speaking and assessment of speech, and illustrates several more important terms. The tasks Bailey illustrates as apt for intermediate level learners are role-plays, picture-based activities, logic puzzles, information gap, and jigsaw activities.

Chapter 4, the last chapter in section 2, provides a comprehensive guide to teaching speaking to advanced learners with the following tasks: Conversations and other interactions, information gap, jigsaw activities, picture-based activities, extemporaneous speaking, role-plays, and simulations. The chapter also explains more terms to further aid the novice teacher.

The last chapter makes up section 3 of the book and addresses several key issues: learners' use of first language in the classroom, significance of speaking activities, learning styles, error correction, large classes, multi-level classes and using technology in teaching. Here, the author concentrates on metalanguage, wait time, turn overlaps, turn taking, learning styles, and backward build up. She also introduces the use of pronunciation tests, chat rooms, corpora and concordances, and computer-enabled functions that have much to offer the learning of speaking.

Afterwards, each chapter steers readers to reflect on the various hypotheses employed in the book, directs testing of these hypotheses, furnishes references and related websites, and concludes with a summary. Bailey then wraps up the book with an exhaustive glossary and index.

While certainly a complete text, *Speaking* is not without its shortcomings. A chapter on the current research practices in the area of teaching speaking and an accompanying CD would have been an additional benefit for readers. Nevertheless, ESL/EFL professionals will find that the book serves as an excellent guide for their empowerment and competence building in the area of teaching speaking, an ideal textbook for courses in ELT methodology, teacher training and lesson planning, and a highly recommended read for the practicing ELTician.



## Book Review

*Culturally Contested Pedagogy: Battles of Literacy and Schooling between Mainstream Teachers and Asian Immigrant Parents*

Guofang Li . Albany: State University of New York Press, 2006. Pp. xvi + 265.

Reviewed by Ozgur Yildirim

State University of New York at Buffalo

New York, USA

Guofang Li, in her book *Culturally Contested Pedagogy: Battles of Literacy and Schooling between Mainstream Teachers and Asian Immigrant Parents*, points out that some people in Canada hold the misconception that Asian students are almost always high achievers: They are good at mathematics, get good scores on the GRE and GMAT, and outperform their non-Asian classmates. The purpose of Li's book is to dispel this belief by providing an ethnographic study which shows that one portion of this population, immigrant children, members of one of Canada's fastest growing minority groups, is having a lot of problems in their education and then offer suggestions for both the parents and teachers of these children.

In the first two chapters, Li provides the background of the study and describes the setting and the participants. The study was conducted in Canada with eight Chinese children in grades one through five from high socioeconomic families, their parents, and their Canadian teachers. The ages of the children in the study range from six to twelve--six of the children were born in Canada and two were born in China. The problem the study focuses on is the different perspectives parents and teachers have about education and how these affect the children.

The next portion of the book, the following five chapters, presents the results of the study. Chapter 3 focuses on teacher-parent conflicts. In this chapter, Li outlines how the different views parents and teachers hold about education are the root of most of the problems. An important factor, Li notes, is the battle parents and teachers have about the methods of literacy instruction. That is, Chinese parents prefer traditional teacher centered education whereas Canadian teachers try to use student centered methods in the class. Chinese parents, for example, want their children to spend a lot of time doing assignments and to always be controlled by the teacher. On the other hand, Canadian teachers want the children to learn

gradually through experience and to try to promote learner autonomy in their classrooms. The fourth and fifth chapters go on to describe the literacy conflicts the children face in the midst of these cultural battles. Here, Li details the students' home and school literacy practices and the negative effects of living in separate worlds of literacy. For example, in some of the situations Li describes, children are caught in school-home conflicts where their teachers emphasize a language-experience approach to literacy instruction in which there is a lot of play and drawing whereas at home their parents want them to memorize more and more words. Chapter 6 discusses the meaning of the battles over literacy and culture. It examines the cultural conflicts over literacy instruction and the complexities which contribute to children's learning difficulties. Having a lot of impact on children's in and out of school learning, the conflicts and complexities include well-intentioned parental involvement into literacy education, different interpretations of school policies, and differences in school and home approaches to children's underachievement.

In the final chapter, Li makes suggestions for bridging the literacy-based differences parents and teachers have. One of the ways she discusses to solve conflicts is to help the families become aware of the positive outcomes of the teaching system in Canada. She points out that Chinese parents rarely resist changes that can be beneficial to their children if they understand the changes and believe the alterations are fruitful. For teachers, Li recommends classroom instruction would be more effective and meaningful if teachers would endeavor to understand the different cultural background the families have. For example, teachers might try to learn more about minority beliefs and look for ways to accommodate what they do not believe in. Lastly, one of the major conclusions of the book is that both parties believe they are doing something useful for the children and thus it would be in the best interest of the children if both parties put more effort into understanding each other and negotiate more.

With its in-depth look at the negative effects the cultural conflicts between Chinese families and Canadian educators can have on the literacy development of the children of Asian immigrants in North America and its helpful solutions, the book is an insightful read to both parties and those who wish to help them work together for the benefit of the children.



**Publisher**

**Dr. Paul Robertson**  
Time Taylor International Ltd

**Senior Associate Editor**

**Dr. Roger Nunn**  
*The Petroleum Institute*  
*Abu Dhabi, UAE*

**International Business Relations**

**Scott Grigas**  
*Yongsan University*  
*Korea*

**Associate Editors**

**Darren Lingley**  
*Kochi University*  
*Japan*

**Dr. Esmat Babaii**  
*University for Teacher Education*  
*Iran*

**Dr. Sivakumar Sivasubramaniam**  
*American University of Armenia*  
*Yeravan, Armenia*

**Dr. Aly Anwar Amer**  
*Sultan Qaboos University*  
*College of Education*  
*Sultanate of Oman*

**Jeremy Cross**  
*University of Melbourne*  
*Australia*

**Dr. John Adamson**  
*Shinshu Honan College*  
*Japan*

**Dr. Erlenawati Sawir**  
*Melbourne University*  
*Australia*

**Dr. Eva Bernat**  
*Division of Linguistics & Psychology,*  
*Macquarie University, Australia*

**Dr. Reima Sado Al-Jarf.**  
*College of Languages and Translation,*  
*King Saud University,*  
*Riyadh, Saudi Arabia*

**Dr. Ahmet Acar**  
*Dokuz Eylül University*  
*Turkey*



## Editorial Group

- Dr. Mingsheng Li**  
Department of Communication and Journalism  
Massey University  
Wellington, N.Z.
- Dr. Kota Ohata**  
International Christian University  
Tokyo, Japan
- Professor Chen Yong**  
School of Foreign Languages  
Chongqing University  
China
- John A. Unger, Ph.D.**  
Institute of American Indian Art  
Santa Fe, New Mexico, USA
- Dr Aysha Viswamohan**  
Dept of Humanities & Social Sciences  
Indian Inst. of Technology, Madras
- Dr. Toshiyuki Takagaki**  
Faculty of Art & Culture  
Onomichi University, Japan
- Mr. David Litz**  
United Arab Emirates University
- Dr. Ming Cheung**  
Department of English and Communication  
City University of Hong Kong
- Dr. Mohammad Ali Salmani-Nodoushan**  
University of Zanjan  
Iran
- Dr. Khalid Al Hagsayer**  
Imam University  
Saudia Arabia
- Mr. Tim Thompson**  
Woosong University  
South Korea
- Dr. Tan Bee Tin**  
The University of Auckland,  
New Zealand
- Alison Smith**  
United Arab Emirates University  
UAE
- Naoki Fujimoto-Adamson**  
Tokyo University of Science  
Japan
- Dr. Mustafa Z. Altan**  
Erciyes University, Faculty of Education  
Kayseri, Turkey
- Annie Hong Qin Zhao**  
Education Department  
University of Bath
- Dr. Toshinobu Nagamine**  
Prefectural University of Kumamoto  
Japan
- Dr. Kourosch Lachini**  
Associate Professor in TEFL  
Qatar University  
Qatar
- Alka Singhal**  
M.S. University  
Baroda, India
- Mr. David Brown**  
The University of the Thai Chamber of Commerce, Thailand
- Dr. Michael Thomas**  
Associate Professor  
Nagoya University of Commerce and Business, Japan
- Book Review Editor  
John Baker**  
Tamkang University, Taiwan
- Dr. Nguyen Thi Thuy Minh**  
National University, Hanoi  
Vietnam
- Dr. Hüsni Enginarlar**  
Director, School of Foreign Languages  
Middle East Technical University  
Ankara
- Marcus Otlowski**  
Dept. of International Studies  
Kochi University, Japan
- Mr. Huw Jarvis**  
Salford University  
United Kingdom
- Dr. Chris Alexander**  
International University of Cyprus  
Nicosia, Cyprus
- Dr. Xiuping Li**  
Sunderland University  
UK
- Mr. Roger Cohen**  
University of Guam  
Guam
- Dr. Deepti Gupta**  
Department of English  
Panjab University  
Chandigarh, India
- Mr. Neil Heffernan**  
Hiroshima Shudo University  
Japan
- Chieko Aoyama**  
Shinshu Honan College, Japan
- Dr. Benedict Lin**  
Nanyang Technological University, Singapore
- Dr. Ahmed Shakir El-Kilabi**  
Department of English  
Nizwa College of Education  
Oman
- Mr. Kevin Landry**  
Hongik University  
Korea
- Ian Isemonger**  
Kochi Women's University  
Japan
- David Hall**  
Macquarie University  
Australia
- Stuart D. Warrington**  
Asia University, Tokyo,  
Japan



## Board of Editorial Advisors

**Professor Rod Ellis**

*University of Auckland  
New Zealand*

**Dr. Phyllis Ghim-Lian Chew**

*Nanyang Technological University  
Singapore*

**Dr. Robert David Carless**

*Hong Kong Institute Education  
Hong Kong*

**Professor Terry Shortall**

*Birmingham University  
Birmingham, UK*

**Marc Helgesen**

*Miyagi Gakuin Women's University  
Sendai, Japan*

**Dr. Yasuko Kanno**

*Asst' Professor English  
University of Washington, U.S.A.*

**Dr. David Shaffer**

*English Language Dept  
Chosun University, Korea*

**Dr. Amy Tsui**

*Faculty of Education  
University of Hong Kong*

**Professor Robert Phillipson**

*Faculty of Languages, Communication  
and Cultural Studies  
Copenhagen Business School  
Denmark*

**Professor David Nunan**

*Director & Chair of Applied Linguistics  
The University of Hong Kong*

**Dr. James P. Lantolf**

*Centre for Language Acquisition  
Pennsylvania State University, U.S.A.*

**Dr. Rebecca Oxford**

*University of Maryland  
U.S.A.*

**Dr. Francis Mangubhai**

*University of Southern Queensland  
Australia*

**Dr. Alan Tonkyn**

*Head of Dept. of Applied Linguistics  
The University of Reading, UK*

**Dr. Jeong-Bae Son**

*University of Southern Queensland  
Australia*

**Ms. Cheon Heesook**

*Columbia University, N.Y.  
U.S.A.*

**Professor Claire Kramersch**

*University of California, Berkeley  
U.S.A.*

**Dr. Robert J. Dickey**

*Gyeongju University  
Korea*



## **Guidelines for Submissions**

### **Submissions for the Quarterly Issue**

#### **Submissions guidelines**

The *Asian EFL Journal Quarterly* is a fully peer-reviewed section of the journal, reviewed by a team of experts in EFL from all over the world. The *Asian EFL Journal* welcomes submissions written in different varieties of world Englishes. The reviewers and Associate Editors come from a wide variety of cultural and academic backgrounds and no distinction is made between native and non-native authors. As a basic principle, the *Asian EFL Journal* does not define competence in terms of native ability, but we are a strictly reviewed journal and all our reviewers expect a high level of academic and written competence in whatever variety of English is used by the author. Every effort will be made to accept different rhetorical styles of writing. The *Asian EFL Journal* also makes every effort to support authors who are submitting to an international journal for the first time. While major revisions may be requested, every effort is made to explain to authors how to make the necessary revisions.

Each submission is initially screened by the Senior Associate Editor, before being sent to an Associate Editor who supervises the review. There is no word minimum or maximum.

#### **There are two basic categories of paper:**

\* Full research papers, which report interesting and relevant research. Try to ensure that you point out in your discussion section how your findings have broad relevance internationally and contribute something new to our knowledge of EFL.

\* Non-research papers, providing detailed, contextualized reports of aspects of EFL such as curriculum planning. Very well documented discussions that make an original contribution to the profession will also be accepted for review. We cannot accept literature reviews as papers, unless these are "state of the art" papers that are both comprehensive and expertly drafted by an experienced specialist.

When submitting please specify if your paper is a full research paper or a non-research paper. In the latter case, please write a paragraph explaining the relevance of your paper to our *Asian EFL Journal* readership.

Authors are encouraged to conform with international standards of drafting, but every effort will be made to respect original personal and cultural voices and different rhetorical styles. Papers should still be fully-referenced and should use the APA (5<sup>th</sup> edition) format. Do not include references that are not referred to in the manuscript.

Some pieces submitted to the quarterly issue may be reclassified during the initial screening process. Authors who wish to submit directly to the Teaching Articles section should read the separate guidelines and make this clear in the submission e-mail.

**Referencing:** Please refer to the *Publication Manual of the American Psychological Association* (5<sup>th</sup> ed.) – Contributors are also invited to view the sample PDF guide available on our website and to refer to referencing samples from articles published from 2006. Due to

the increasing number of submissions to the *Asian EFL Journal*, authors not conforming to APA system will have their manuscripts sent back immediately for revision. This delays publication and taxes our editorial process.

**Format for all submissions** (Please read this before submitting your work)

All submissions should be submitted to: [asian\\_efl\\_journal@yahoo.com](mailto:asian_efl_journal@yahoo.com)

- i) The document must be in MS Word format.
- ii) Font must be Times New Roman size 12.  
Section Headings: Times New Roman (Size 12, bold font).  
Spacing: 1.5 between lines.
- iii) 'Smart tags' should be removed.
- iv) Footnotes must not 'pop up' in the document. They must appear at the end of the article. Use the superscript font option when inserting a note rather than the automatic footnote or endnote option.
- iv) Citations - APA style. (See our website PDF guide)  
Use the APA format as found in the Publication Manual of the American Psychological Association (APA), 5th Edition, for headings, citations, reference lists and in text referencing. Extra care should be taken for citing the Internet and must include the date the site was accessed.  
  
About APA Style/format: <http://www.apastyle.org/aboutstyle.html>  
APA Citation Style: <http://www.liu.edu/cwis/CWP/library/workshop/citapa.htm>  
APA Style Workshop: <http://owl.english.purdue.edu/workshops/hypertext/apa/index.html>
- v) Keywords: All articles must include Keywords at the beginning of the article. List 4-6 keywords to facilitate locating the article through keyword searches in the future.
- vi) Graphs and Charts - either in the body of the document or at the end. In certain cases, a graphic may not appear in the text of the web version of the *Asian EFL Journal* but a link to the graphic will be provided.
- vii) Paragraphs. Double space between paragraphs. Indent the beginning of each paragraph with three strikes of the space bar except those immediately following a heading, quotation, example, figure, chart or table. Do not use the tab key.
- viii) Keep text formatting (e.g., italics, bold, etc.) to the absolute minimum necessary. Use full justification. All lines to be against Left Hand Side Margin (except quotes - to be indented per APA style).
- ix) Abstract  
The abstract should contain an informative summary of the main points of the article, including, where relevant, the article's purpose, theoretical framework, methodology, types of data analysed, subject information, main findings, and conclusions. The abstract should reflect the focus of the article.

x) Graphs – to fit within A4 size margins (not wider)

Thank you for your cooperation.

asian\_efl\_journal@yahoo.com

**Please include the following with your submission:**

Name

School affiliation

Address

E-mail

Phone number

Brief Bio Data noting history of professional expertise

Qualifications

An undertaking the work has not been published elsewhere

Abstract

Any questions regarding submission guidelines, or more detailed inquiries about less common citation styles, may be addressed to the Editorial Board or our Journal Production Editor (Darren Lingley) at: [lingley@cc.kochi-u.ac.jp](mailto:lingley@cc.kochi-u.ac.jp)

**Book Reviews:**

The *Asian EFL Journal* currently encourages two kinds of submissions, unsolicited and solicited. Unsolicited reviewers select their own materials to review. Both teachers and graduate students are encouraged to submit reviews. Solicited reviewers are contacted and asked to review materials from its current list of availability. If you would like to be considered as a solicited reviewer, please forward your CV with a list of publications to the Book Review Editor at: [asianefljournalbookreviews@yahoo.com](mailto:asianefljournalbookreviews@yahoo.com).

All reviewers, unsolicited and solicited, are encouraged to provide submissions about materials that they would like to suggest to colleagues in the field by choosing materials that they feel have more positive features than negative ones.

**Length and Format:**

1. Reviews should be prepared using MS Word and the format should conform to 12 pica New Times Roman font, 1.5 spacing between lines, and 1 inch margins.
2. The reviewer(s)' full names including middle initial(s), title, school affiliation, school address, phone number, and e-mail address should be included at the top of the first page.
3. The complete title of the text, edition number, complete name(s) of author(s), publisher, publisher's address (city & state), and date of publication should be included after the reviewer(s)' identifying information.
4. Reviews should be between 500-700 words.
5. A brief biography of the author(s) should be included after the review.
6. A statement that the submission has not been previously published or is not being considered for publication elsewhere should be included at the bottom of the page.

**Organization:**

Reviewers are encouraged to peruse reviews recently published in the quarterly PDF version of the Journal for content and style before writing their own. While creativity and a variety of writing styles are encouraged, reviews, like other types of articles, should be concisely written and contain certain information that follows a predictable order: a statement about the work's intended audience, a non-evaluative description of the material's contents, an

academically worded evaluative summary which includes a discussion of its positive features and one or two shortcomings if applicable (no materials are perfect), and a comment about the material's significance to the field.

Style:

1. All reviews should conform to the Journal's APA guideline requirements and references should be used sparingly.
2. Authors should use plural nouns rather than gendered pronouns such as he/she, his/her him/her and adhere to the APA's Guidelines for Non-Sexist Use of Language, which can be found at:  
<http://www.apa.udel.edu/apa/publications/texts/nonsexist.html>.