The Asian EFL Journal Quarterly
December 2011
Volume 13, Issue 4

Senior Editors:
Paul Robertson and Roger Nunn
# Table of Contents:

**Foreword by Robert Baird** ......................................................... 6-9

1. **Megumi Hamada and Chaehee Park** ......................................... 10-32  
   - *Word-Meaning Inference: A Longitudinal Investigation of Inference Accuracy and Strategy Use*

2. **Xiaoli Jiang and Gerard Sharpling** .......................................... 33-68  
   - *The Impact of Assessment Change on Language Learning Strategies: The Views of a Small Group of Chinese Graduate Students Studying in the UK*

3. **Neil H. Johnson and Jonathan deHaan** .................................... 69-101  
   - *Second Language Development through Technology Mediated Strategic Interaction*

4. **Minh Hue Nguyen** ........................................................................ 102-134  
   - *Learner Self-management Procedures Reported by Advanced and Intermediate ESL Students*

5. **Peter Crosthwaite** ....................................................................... 135-166  
   - *The Effect of Collaboration on the Cohesion and Coherence of L2 Narrative Discourse between English NS and Korean L2 English Users*

6. **Rais Ahmed Attamimi and Hajar Abdul Rahim** ............................ 167-197  
   - *Socio-Economic Orientations in Foreign Language Learning Motivation: The Case of Yemen*

7. **John F. Haggerty** ....................................................................... 198-227  
   - *An Analysis of L2 Motivation, Test Validity and Language Proficiency Identity (LPID): A Vygotskian Approach*

8. **Katerina Petchko** ................................................................. 228-255  
   - *Input Enhancement, Noticing, and Incidental Vocabulary Acquisition*

9. **Ya-huei Wang** ........................................................................ 256-287  
   - *Shattering the hierarchical education system: The creation of a poststructural feminist English classroom*
10. Breda O’Hara-Davies................................................................. 288-299
   - English only? inda kali eh! (not likely!) – Changing the paradigm

11. Gurnam Kaur Sidhu, Chan Yuen Fook and Sarjit Kaur Sidhu....... 300-327
   - Students’ Reactions to School Based Oral Assessment:
     Bridging the Gap in Malaysia
Book Reviews

1. *A Glossary of Corpus Linguistics*
   Paul Baker, Andrew Hardie, & Tony McEnery
   Reviewed by Vander Viana ......................................................... 328-330

2. *Adult Language Learners*
   Ann F.V. Smith and Gregory Strong (eds.)
   Reviewed by Stephanie A. Wilton .................................................. 331-333

3. *Language and Education*
   M. A. K. Halliday. Edited by Jonathan J. Webster
   Reviewed by Servet Celik .............................................................. 334-335

4. *Reading in a Second Language: Moving from Theory to Practice (Cambridge Applied Linguistics Series)*
   William Grabe
   Reviewed by Jim Bame ............................................................... 336-337

*Asian EFL Journal* editorial information and guidelines ......................... 338-346
Foreword

In this, the final quarterly edition of 2011, we have the opportunity to engage with a wealth of research, perspectives, approaches and theories which all have clear implications for the field of English language teaching across Asia.

In *Word-Meaning Inference: A Longitudinal Investigation of Inference Accuracy and Strategy Use*, Hamada and Park, recognising the importance of strategies employed to infer meaning when reading in a second language, investigate the word-meaning inference behaviours of three college-level ESL students. The findings of this qualitative study are discussed in relation to previous literature, and implications for teachers and future research are highlighted.

The second article, *The Impact of Assessment Change on Language Learning Strategies: The Views of a Small Group of Chinese Graduate Students Studying in the UK* by Jiang and Sharpling, qualitatively investigates the experiences and perceptions of Chinese students studying at a UK University vis-à-vis the shifting assessment orientations that they inevitably encounter, both before arriving in the UK and during their time there. The authors offer substantial insights into the influence of assessment on the choice and development of language learning strategies, as well as the diverse factors that influence the experiences, learning behaviours, perceptions and values of the students in their study.

In *Second Language Development through Technology Mediated Strategic Interaction*, Johnson and deHaan give a detailed account of the design and implementation of technology-mediated strategic interaction tools for a group of Japanese university students. They give an overview of problems that Japanese university students in their context typically face, problems that, according to their theoretical discussions and initial results, can be more effectively met with approaches that are driven by a deep understanding of the fundamental nature of language learning alongside contextual awareness.
The next article is Nguyen’s *Learner Self-management Procedures Reported by Advanced and Intermediate ESL Students*, in which the author reports the results of a New Zealand based study of intermediate and advanced ESL students’ self-management procedures, adding to a growing amount of literature in the field. Various research tools were used to gain insights into students’ practices. Similarities and differences between the two groups of learners are identified and critically discussed in relation to current understandings, theories and debates.

In *The Effect of Collaboration on the Cohesion and Coherence of L2 Narrative Discourse between English NS and Korean L2 English Users*, Crosthwaite investigates differences between native English speakers and Korean English learners in terms of cohesive reference maintenance and the role of scaffolding on the accuracy of the latter’s discursive performances. To achieve this, spoken data are analysed and compared in to identify specific grammatical forms that have an important role in maintaining coherent reference to discourse referents. From this, the role of scaffolding is analysed to ascertain its role in maintaining coherence and accuracy among the Korean speakers. Findings from this study suggest an important role for scaffolding in enhancing coherence and easing difficulties managing accurate reference maintenance.

The sixth article, *Socio-Economic Orientations in Foreign Language Learning Motivation: The Case of Yemen* by Attamimi and Rahim, draws attention to the notion of cultural capital, parental economic status and motivation to learn English in a dual-survey study which is supported by interviews. The authors discuss a number of areas of motivation that incorporate cultural capital, student orientations, students’ feelings of or against ‘integrativeness’ with the target language community (or out-groups). Results are presented in the form of quantitative summaries and discussion.

In *An Analysis of L2 Motivation, Test Validity and Language Proficiency Identity (LPID): A Vygotskian Approach*, Haggerty investigates learners’ attitudes, beliefs and motivations in relation to their experiences with high-stakes language assessments. The author’s account advocates the incorporation of Vygotsky’s notion of ‘language proficiency identity’ into research into the impact of assessment and learning experiences on various aspects of identity, values and motivation. The resulting
discussion raises questions about test validity and current assessment practices while at the same time reporting the impact that language tests appear to have on Korean learners in different stages of education.

In the next article, *Input Enhancement, Noticing, and Incidental Vocabulary*, Petchko investigates input enhancement in reading classes by investigating the role of textually enhancing target words (non-words in this study) in noticing, meaning recognition and meaning recall. The author gives a considerable overview of the importance of this area of research for the language teaching community before discussing the implications of this study’s findings and recommending future research directions.

Wang’s *Shattering the hierarchical education system: The creation of a poststructural feminist English classroom* draws our attention to a pedagogical approach designed to improve Taiwanese students’ English proficiency, critical thinking faculties and satisfaction through their learning experiences in English. The author’s poststructural feminist model is discussed and justified drawing on cultural aspects of classroom learning in the Taiwanese context and the results, gained through a variety of methods, appear to show that this approach has been implemented successfully to enhance students’ learning, thinking and experiences.

The next article, *English only? inda kali eh! (not likely!) – Changing the paradigm* by O’Hara-Davies, is both a personal account of realisations that led to shifts in the author’s orientations to language teaching and a call for the field to recognise the importance of heightened awareness of the issues raised by voices in our field and beyond. The author takes an autoethnographic approach to reporting her evolution from a native speaker teacher with a limited awareness of her own limitations and shortcomings to a more critically aware teacher whose approach embodies linguistic and cultural differences. A key result of this account is to help reposition the native speaker expert as often lacking contextual cultural and linguistic awareness, and having to adapt to overcome these shortcomings, rather than the learners.

In the final article, *Students’ Reactions to School Based Oral Assessment: Bridging the Gap in Malaysia*, Sidhu, Fook and Sidhu use quantitative and qualitative research tools to investigate Malaysian students’ perceptions, opinions and understandings of
the Malaysian School Based Oral English Test (SBOET), which represented a shift towards formative testing in 2002. The study is one of the first to research the effects of the introduction of this assessment policy shift, and its findings are reported on a number of levels, informing the perceived effectiveness of the tests, their implementation (including dissemination of information about them) and their perceived value. An engaging discussion of classroom oral assessment is offered, as are suggestions for improving the implementation of the SBOET, and handling educational policy shifts in general.

Perhaps the unifying thread that connects the research presented in this edition, apart from relating to English teaching and Asia, is that each article incorporates our growing understanding of language and language learning into contextualised approaches to English teaching, learning, assessment, use and/or policy. In turn, these authors contribute to our growing understanding of the knowledge that they exploit.

Robert Baird
Production Editor
Word-Meaning Inference: A Longitudinal Investigation of Inference Accuracy and Strategy Use

Megumi Hamada
Ball State University, Muncie, Indiana, USA
mhamada@bsu.edu

Chaehee Park
Kangwon National University, South Korea
cpark@kangwon.ac.kr

Bio Data
Megumi Hamada is an associate professor in the English Department at Ball State University, Muncie, Indiana. Her specialty is second language reading and vocabulary.

Chaehee Park is an English instructor at Kangwon National University.

Abstract
Ability to infer the meaning of unknown words encountered while reading plays an important role in learners’ L2 word knowledge development. In order to provide a longitudinal inquiry into this topic, this study conducted a qualitative analysis of three Korean college-level ESL learners’ meaning-inference behaviors over a 4 week period, focusing on inference accuracy and strategy use. The learners were engaged in weekly reading and meaning-inference training, in which they read academic texts, identified unknown words in the texts, and inferred the meanings of the unknown words. The analysis of the think-aloud protocol indicated that (a) learners with higher inference accuracy used the same types of strategies consistently; (b) learners with lower inference accuracy used a wider variety of strategies more frequently; and (c) learners with higher inference accuracy preferred global strategies over local strategies. Implications for meaning-inference instruction are discussed.

Keywords: L2 word learning, word-meaning inference, reading, incidental word learning, word-knowledge development

Introduction
Word knowledge is crucial in all aspects of second-language (L2) learning. In the past decade, there has been increasing interest in the nature of this knowledge and its acquisition. One growing area of research focuses on word-meaning inference while reading. Word-meaning inference involves “making informed guesses as to the meaning of a word in light of all available linguistic cues in combination with the learner’s general knowledge of the world and awareness of context” (Haastrup, 1991, p. 40). Word-meaning inference is also known as “incidental” word learning due to its
indirect nature. In contrast to intentional word learning, word-meaning inference can occur as a by-product of another activity, such as reading, which can provide learners with more contextualized, efficient, and individualized learning opportunities if they attempt to figure out the meaning of unknown words they encounter during the activity (Huckin & Coady, 1999). Although the contribution of meaning-inference to word-knowledge gain, particularly to word retention, seems to be questionable (e.g., Mondria, 2003), meaning-inference has the potential benefit of vocabulary knowledge gain, and this method is suggested to be “the most important of all sources of vocabulary learning” (Nation, 2001, p. 232). Although meaning-inference has been widely incorporated into teaching, a typical difficulty that teachers encounter is “wild guesses” – the fact that not every student is able to infer the correct meaning of unknown words, although wild guesses could be the beginning stage of incidental word learning (e.g., Kaivanpanah & Alavi, 2008; Kelly, 1990). Although a number of studies have reported on this topic (e.g. Frantzen, 2003; Nassaji, 2003; Paribakht & Wesche, 1999; Pulido, 2003), more findings are needed to clarify factors that relate to the improvement of inference accuracy. Consequently, the primary objective of this study was to provide a qualitative inquiry into the development of L2 word-meaning inference, an area of research that is extremely scarce but vital for instruction. The study analyzed three college-level ESL learners’ inference behaviors over a 4 week period, using the think-aloud technique, focusing on how inference accuracy and strategy use change. The following section presents a review of relevant research.

**Literature Review**

*Meaning-Inference Strategies*

One of the issues often discussed regarding strategy use and inference accuracy is the effectiveness of strategies that utilize local cues vs. global cues. The strategies based on local cues, hereafter called local strategies, include morphological analysis, word-analogy, and grammatical (syntactic) analysis, wherein learners do not necessarily have to apply their understanding of the text in meaning-inference. For example, Huckin and Bloch (1993) investigated college-level ESL learners’ meaning-inference behaviors and reported that the strategy that was most important for accurate meaning-inference was the analytic word-clue strategy.

In contrast, the strategies based on global cues, hereafter called global strategies, involve more contextually-based analysis that requires a deeper level of understanding of the text as well as world knowledge related to the text. It is important to note that
learners may use multiple strategies, including both local and global strategies, in inferring the meaning of a single word. Due to their deeper involvement of context, global strategies are also related to successful reading comprehension (Read, 2000). A number of studies suggest the effectiveness of global strategies over local strategies in meaning-inference, indicating that local strategies tend to yield inaccurate inference (e.g., Chern, 1993; Frantz, 2003; Haynes, 1993; Laufer & Sim, 1985; Morrison, 1996; Nassaji, 2003; Parry, 1993). For example, Haynes (1993) reported that an adult ESL student inferred “the end of spring” for offspring based solely on word-level analysis despite the fact the inferred meaning does not match the context in the passage. Similarly, Nassaji (2003) found that adult ESL learners incorrectly inferred meanings of unknown words based on graphic similarity of words, such as permeated to meat and affluence to influence.

Apart from the effectiveness of local vs. global strategies in inference accuracy, local strategies are more popular than global strategies among learners. Several studies reported that their participants preferred local over global strategies (Bengeleil & Paribakht, 2004; Bensoussan & Laufer, 1984; Haynes, 1993). More specifically, Nassaji (2004) concluded that students with less vocabulary knowledge preferred local strategies more than students with greater vocabulary knowledge. Huckin and Bloch (1993) furthermore reported that the participants first relied on a local strategy, but if they could not infer the meaning of unknown words with that strategy, they tried a global strategy. These findings suggest that local strategies are more easily used by learners, presumably due to the fact that local strategies do not require a global understanding of the text.

The type and frequency of strategy use is another issue investigated in word-meaning inference research. A majority of vocabulary studies that are not in the context of meaning-inference consistently suggest that using a wide variety of strategies more frequently leads to word learning (Fan, 2003; Kojic-Sabo & Lightbown, 1999; Lawson & Hogben, 1996; Sanaoui, 1995). However, meaning-inference research reports that a wider variation of strategy types and higher frequency of strategy use were associated with learners with lower inference accuracy (Bengeleil & Paribakht, 2004; Hamada, 2009). Nassaji (2003) also points out that it is the “quality” rather than “quantity” of strategies used that leads to inference accuracy. The effectiveness of strategy use in inference accuracy seems to be attributed to the selection of a strategy appropriate for each particular unknown word, rather than the number of strategies or strategy types used.
As for which strategies are related to higher inference accuracy, research has shown an agreement on the effectiveness of metacognitive strategies. Metacognitive strategies involve learners’ self-monitoring of their own meaning-inference behaviors. For example, Nassaji (2003) reported that verifying and self-inquiry were related to meaning-inference accuracy more than other strategies. Nassaji (2004), furthermore, concluded that verifying and self-inquiry were used more often by learners with more word knowledge. In addition, Griffiths (2006) reported similar findings, although her study was not specifically in word-meaning inference, but in word-learning in general. She reported that learners with higher proficiency tended to use a management strategy, attributable to their ability to control their own learning. Although metacognitive strategies may be effective, Hamada (2009) also reported that they are the strategies used least frequently.

**Effect of Inference Instruction**

A number of studies have investigated the development of word-knowledge through meaning-inference while reading with native-speaking students (e.g., Buikema & Graves, 1993; Carnine, Kameenui, & Coyle, 1984; Hafner, 1965; Jenkins, Matlock, & Slocum, 1989). The development examined in these studies was the change (an increase or decrease) in inference accuracy after meaning-inference training was given to students for a certain period of time. The training included explicit strategic instruction regarding how to infer the meaning of unknown words in a text by incorporating available textual cues and strategies for meaning-inference. For instance, Carnine, Kameenui, and Coyle (1984) administered a three-session meaning-inference training series to fourth-, fifth-, and six-grade students and found that the group who received the training showed higher inference accuracy than the control group. Likewise, with teen-age native-speaking students, in Buikema and Graves (1993), seventh- and eighth-grade students received a 5-day meaning-inference training series and showed an increase in inference accuracy. The general consensus of these L1 training studies is that the training improved learners’ meaning-inference accuracy, a conclusion consistent with a meta-analysis reported in Fukkink and de Glopper (1998) and Kuhn and Stahl (1998).

As to the effect of training on meaning-inference accuracy, L2 research is extremely scarce. A key study that should be noted is Fraser (1999), in which college-level ESL students engaged in reading and word-learning over 5 months. The study focused on three “metacognitive strategies” for dealing with unknown words the students
encountered while reading: (a) ignore and continue reading, (b) consult a dictionary or another individual, and (c) infer the words’ meanings. In this study, an explicit instruction to attempt meaning-inference was attributed with an increase in inference accuracy and a decrease in ignoring the unknown words. In a recent study, Hamada (2009) investigated whether there would be a change in inference accuracy with college-level ESL students who engaged in inference training over 4 weeks. The results indicated that learners with lower inference accuracy showed a pattern of increase in inference accuracy as their number of strategies and variety of strategy types increased, whereas learners with higher inference accuracy maintained their accuracy at a high level while using the same types of strategies consistently. It should be noted that in both Fraser (1999) and Hamada (2009), the inference training given was to direct learners to use meaning-inference. Unlike L1 studies (e.g., Carnine, Kameenui, & Coyle, 1984) these two L2 studies did not include any specific strategic instruction.

Longitudinal Study on L2 Word-Learning

This section briefly reviews longitudinal studies that addressed the development of L2 word-knowledge in areas other than meaning-inference. Among the few studies available (e.g., Laufer, 1991; Palmberg, 1987; Schmitt, 1998), in short, findings are mixed. For example, Laufer (1991) and Schmitt (1998) reported that their college-level L2 students did not show much development over time, whereas Palmberg (1987) reported that his elementary school children showed a steady word-knowledge increase over time. In order to draw pedagogically useful conclusions, more findings regarding L2 word-knowledge development, in general, and L2 word-meaning inference ability, in specific, are necessary.

Research Questions

As mentioned earlier, the majority of L2 word-meaning inference studies investigated learners’ behaviors at only one point in their incremental learning experiences. In order to provide more findings in a longitudinal inquiry of meaning-inference and to confirm earlier findings (Fraser, 1999; Hamada, 2009), this study attempted a qualitative investigation of how learners’ meaning-inference behaviors might change when given explicit inference instruction. The following specific research questions were examined:

1. Does L2 learners’ meaning-inference accuracy change over time?
2. Does the number of strategies used by L2 learners change over time?
3. Do the types of strategies used by L2 learners change over time?

These questions are concerned with whether intensive meaning-inference experience results in a change in L2 word-meaning inference behavior. In order to simulate the intensive meaning-inference experience, a series of word-meaning inference sessions were employed as tasks. In each session, learners were asked to read a text and infer the meaning of unknown words in the text while reading. Following the previous L2 studies, this study also did not include any strategic instruction during the session. “Time” examined in this study specifically refers to the four weekly meetings in which the meaning-inference sessions were administered to L2 learners. A detailed description of the method follows.

**Method**

*Participants*

Three Korean ESL learners (2 females and 1 male) who were studying in a mid-size university in the United States volunteered as participants. Three points were considered for the participant recruitment criteria. The first was English proficiency level. In order for word-meaning inference to occur, learners should have some basic vocabulary knowledge and reading skill (Coady, 1997). Students who completed basic ESL courses and were enrolled in regular university courses were selected as a target proficiency group, because according to a recommendation from the Director of the Intensive English Institute, this learner population typically has basic skills but struggles in academic reading. The second point was learners’ L1 background. In order to minimize possible L1 influence in the data, students with the same L1, Korean, one of the L1 groups that was well represented in the university, were recruited. Lastly, only those who could complete all of the data collection sessions were selected as participants.

Each participant will be identified by his or her pseudonym henceforth: Adam, Beth, and Cindy. Their age, major, and TOEFL scores are given in Table 1. They had similar backgrounds in English learning. All received education from kindergarten through high school in Korea. They had been in the United States for a short period of time (one semester) at the beginning of data collection and had never lived in an English-speaking environment before beginning their current program. Adam was a 23 year old male undergraduate student in architecture. Beth was a 20 year old female undergraduate student in education. Cindy was a 37 year old female graduate student.
In order to measure their English proficiency level, prior to data collection, the three students were asked to take a grammar and a reading section of a retired version of Test of English as a Foreign Language (TOEFL). Their scores indicated that their proficiency levels differed slightly, Adam being the most advanced, Beth in the middle, and Cindy the least advanced, although all of them had fulfilled the basic ESL requirements at the university.

Materials
Four passages were selected from the Insights series (Brinton et al., 1997a, 1997b) for this study. The series includes readings in various academic subjects targeting college-level advanced ESL students. As in Hamada (2009), the series was chosen for this study by a college-level ESL instructor and the first author because each reading was at an appropriate level for the present participants in terms of vocabulary and syntactic structures, but still contained advanced vocabulary items related to academic concepts that would be unknown to all of the participants and appropriate for meaning-inference. The title of each passage and the number of words in each passage were: (1) “Local Wind Systems” (544 words) from the natural science section; (2) “Genetically Determined Behavior” (451 words) from the biology section; (3) “An Explanation for the Absence of Extraterrestrials on Earth” (485 words) from the astronomy section; and (4) “Temptations of a Superpower” (546 words) from the political science section. The main concerns in selecting these passages were as follows: (a) They were approximately 500 words, a length found to be appropriate for the participants to complete the tasks based on the pilot study; and (b) The subject area in each passage was not any of the participants’ major field. None of the participants had studied the Insights series prior to the present study, nor did they report that they had specialized knowledge of the topics or subjects in the passages.

Tasks and Procedures
A weekly reading and word-meaning inference session was administered for 4 weeks. Each participant met individually with the examiner (the second author) in a quiet room, and each session was carried out in an individual meeting. Each session consisted of the following tasks: (1) passage reading (approximately 30 minutes), (2) comprehension check (2-3 minutes), and (3) word-meaning inference while reading (approximately 30 minutes). For passage reading, in order to have the participants
focus on the activity of reading rather than vocabulary, they were instructed to read the passage at a comfortable speed and to focus on overall comprehension. After reading, they were asked to give a brief summary of the passage in English, in order to check their reading comprehension. All of the participants included key ideas in the summary, exhibiting a basic understanding of the passage.

Following the comprehension check, the participants underlined unknown words in the passage. They then inferred the meanings of the words, using the think-aloud technique. This technique was used because it can reveal what thought process learners are actually using to complete the task of word-meaning inference while reading, as it has been used in many studies that examined reading and/or word learning strategies (e.g., Fraser, 1999; Lawson & Hogben, 1996; Pressley & Afferbach, 1995, Nassaji, 2003, 2004). To ensure that the participants were comfortable using the think-aloud technique, before beginning the first think-aloud session, the examiner modeled and explained the technique until the participants said that they were ready to start the session. The participants were allowed to think-aloud in either English or Korean, freely, throughout meaning-inference. Also, throughout the data collection sessions, whenever the participants stopped thinking-aloud, the examiner reminded them to keep using the technique. Their think-aloud responses were audio-recorded and transcribed for analysis.

**Data Analysis**

The participants’ think-aloud protocols were transcribed in English, and a total of 12 transcripts (3 participants x 4 passages) were compiled. Then, all of the unknown words that the participants underlined and worked on meaning-inference were identified and made into a list for each participant.

In order to measure inference accuracy, two independent raters who hold graduate degrees in linguistics/TESOL evaluated the acceptability of the inferred meanings. The following scale was used to judge accuracy in word-meaning inference: 1 point for a correct meaning (i.e., dictionary-like definition or synonym), .5 points for a meaning that was similar to or exhibited a partial meaning of the word, and 0 points for a meaning that did not exhibit any meaning of the word at all or an “I don’t know” answer. For example, for *glide*, “to slide down or up” was awarded 1 point and “to move” .5 point, and for *rear*, “to foster” was awarded 1 point and “trying” and “to put in” 0 point. Overall interrater reliability was .88. The inferred meanings whose rating was disagreed upon were resolved through discussion by the raters.
The analysis of strategy uses was based on the classification from Nassaji (2004), whose learners were at a similar English proficiency level as in this study. The Nassaji classification included three major types of strategies: *identifying, evaluating, and monitoring*. *Identifying* strategies are those that are used to identify (or infer) the meaning of the new word in the passage. *Evaluating* strategies are those that are used to evaluate and check the accuracy of inferred meanings. *Monitoring* strategies are those that learners use to judge their own inferencing behaviors. After a preliminary analysis, the Nassaji classification was modified by adding three subtypes (1b, 1d, and 1e), in order to best capture the strategies used by the present participants. Table 2 summarizes the classification of strategies used for this study. The subtypes under *identifying* strategies were as follows: (a) *repeating* for repetition of any portion of the passage, (b) *morphological analysis* for making use of morphological structure or knowledge, (c) *word-form analogy* for making use of phonological or visual similarities with other words, (d) *syntactic analysis* for making use of the syntactic function of an unknown word in a sentence, and (e) *contextual analysis* for making use of contextual and semantic understanding of the relevant portion of the passage. The three strategy subtypes added for this study were: (1b) *morphological analysis*, (1d) *syntactic analysis*, and (1e) *contextual analysis*. The subtypes under *evaluating* strategies were as follows: (a) *verifying* for checking the inferred meaning in context and (b) *self-inquiry* for asking questions about a word or the inferred meaning of a word. Finally, *monitoring* strategies used for this study consisted of a single type, *monitoring* for a conscious awareness of inferencing behavior by judging its ease or difficulty.

Following the list of unknown words created for each participant, two independent raters analyzed the transcripts and coded the types of strategies used for each word on the list. Interrater reliability was .83. The words whose strategy coding was disagreed upon were resolved through discussion by the raters.

**Results**

Table 3 summarizes overall meaning-inference accuracy and strategy use. Adam had the highest mean accuracy rate (51.74%), followed by Beth (31.64%) and Cindy (30.93%). Beth underlined and inferred the highest number of words as unknown (57), followed by Adam (40) and Cindy (30), producing a percentage of known vocabulary coverage for each participant of 98.03%, 97.19%, and 98.52%, respectively. These are at the percentage where learners are able to infer the meaning.
of unknown words accurately (approximately 98% is suggested by Hu & Nation, 2000). As for the total number of strategies used, Beth had the highest number (96), followed by Cindy (87) and Adam (69). Cindy had the highest mean number of strategies used per word (2.9), followed by Adam (1.73) and Beth (1.68). The rest of this section presents individual results.

**Adam**

Adam’s meaning-inference accuracy rate and number of strategies used per word are shown in Figure 1. The accuracy rate varied substantially, Passage 2 being the highest (68.75%), followed by Passage 1 (58.33%), Passage 3 (43.75%), and Passage 4 (36.11%). The number of strategies used per word showed a slight change, Passage 4 being the highest (1.94), followed by Passage 2 and Passage 3 (1.75) and Passage 1 (1).

The ratio of the types of strategies used by Adam is shown in Figure 2. *Contextual analysis* was the most frequently used type (50% in Passage 1, 42.86% in Passage 2, and 50% in Passage 3), except that it was the second most frequently used type in Passage 4 (22.86%). A strategy type whose usage increased was *repeating* (0% in Passage 1, 28.57% in Passage 2, 35.71% in Passage 3, and 48.57% in Passage 4). Strategy types whose usage decreased were *morphological analysis* (33.33% in Passage 1, 7.14% in Passage 2, 14.29% in Passage 3, and 8.57% in Passage 4) and *word-form analogy* (16.6% in Passage 1, 7.14% in Passage 2, 0% in Passage 3, and 2.86% in Passage 4). *Syntactic analysis*, *self-inquiry*, and *monitoring* were not used at all, and *verifying* was used less frequently (14.29% in Passage 2 and 17.14% in Passage 4). Overall types of strategies used did not show a noticeable change. Adam used 3 types in Passage 1 and Passage 3 and 5 types in Passage 2 and Passage 4.

**Beth**

Beth’s meaning-inference accuracy rate and number of strategies used per word are shown in Figure 3. The accuracy rate changed, Passage 3 being the highest (45.45%), followed by Passage 1 (44.12%), Passage 2 (25%), and Passage 4 (12%). The number of strategies used per word showed a slight change, Passage 2 being the highest (2), followed by Passage 4 (1.76), Passage 3 (1.64), and Passage 1 (1.53).

The ratio of the types of strategies used by Beth is shown in Figure 4. The most frequently used type was *repeating* (57.69% in Passage 1, 50% in Passage 2, 55.56% in Passage 3, and 56.81% in Passage 4). Strategy types whose usage increased were
*contextual analysis* (11.53% in Passage 1, 12.5% in Passage 2, 11.11% in Passage 3, and 20.45% in Passage 4) and *verifying* (11.53% in Passage 1, 37.5% in Passage 2, 22.22% in Passage 3, and 18.18% in Passage 4). *Self-inquiry* and *word-form analogy* were not used at all, and *morphological analysis* (3.85% in Passage 1 and 11.11% in Passage 3), *syntactic analysis* (7.69% in Passage 1 and 4.55% in Passage 4), and *monitoring* (7.69% in Passage 1) were used less frequently. Overall, Beth used fewer types of strategies as she experienced more sessions (6 types in Passage 1, 3 types in Passage 2, and 4 types in Passage 3 and Passage 4).

**Cindy**

Cindy’s meaning-inference accuracy rate and number of strategies used per word are shown in Figure 5. The accuracy rate changed, Passage 3 being the highest (56.25%), followed by Passage 2 (28.57%), Passage 1 (22.22%), and Passage 4 (16.67%). The number of strategies used per word showed a considerable change, Passage 2 being the highest (3.71), followed by Passage 1 (3.2), Passage 3 (2.5), and Passage 4 (2).

The ratio of the types of strategies used by Cindy is shown in Figure 6. *Repeating* was the most frequently used type throughout (31.03% in Passage 1, 26.92% in Passage 2, 25% in Passage 3, and 33.33% in Passage 4). *Contextual analysis* was the second most frequently used type, but showed a decrease (27.59% in Passage 1, 19.23% in Passage 2, 25% in Passage 3, and 16.67% in Passage 4). In addition to repeating and contextual analysis, Cindy used morphological analysis, word-form analogy, and syntactic analysis in each session. The usage of syntactic analysis increased (6.9% in Passage 1, 19.23% in Passage 2, 5% in Passage 3, and 25% in Passage 4), and morphological analysis (6.9% in Passage 1, 11.54% in Passage 2, 20% in Passage 3, and 8.33% in Passage 4) and word-form analogy (6.9% in Passage 1, 11.54% in Passage 2, 5% in Passage 3, and 16.67% in Passage 4) were used consistently. *Verifying* was used less frequently (17.24% in Passage 1, 7.69% in Passage 2, 15% in Passage 3, and 0% in Passage 4), and *self-inquiry* (3.85% in Passage 2) and monitoring (3.45% in Passage 1 and 5% in Passage 3) were used even less frequently. Overall, Cindy used fewer types of strategies as she experienced more sessions (7 types in Passage 1 and Passage 2, 6 types in Passage 3, and 5 types in Passage 4).

**Discussion**

Overall, the results indicate that the development of meaning-inference behaviors
varied among the three students. The first research question asked whether L2 learners’ meaning inference accuracy would develop with practice over time. The inference accuracy for each student changed over the sessions, but the change showed neither a constant increase nor decrease in any of the students, inconsistent with Fraser (1999), where increases were observed. Considering in Hamada (2009), only the participants who started the training session with the lowest inference accuracy showed a steady increase in inference accuracy, the lack of increase in this study might be due to the fact that the participants had already had a certain level of inference ability prior to the training session. Another reason for the lack of increase in inference accuracy is attributed to the selection of the texts. This study attempted to analyze learners’ inference behaviors in a more natural environment, and therefore the participants worked on meaning inference on authentic texts, rather than a text that was controlled for linguistic difficulty. Further study involving a more controlled environment will be necessary in order to confirm the current result.

The second research question was whether the number of strategies used by L2 learners would change over time. One of the most important findings regarding this research question comes from when the students used more strategies. Adam and Beth used the highest number of strategies per word (1.94 and 1.76, respectively) when they had the lowest inference accuracy, in Passage 4, and used the least (1 and 1.53, respectively) when they had second highest inference accuracy, in Passage 1. In contrast, Cindy used the highest number of strategies per word (3.71) when she had the second highest inference accuracy, in Passage 2, and used the least (2) when she had the lowest inference accuracy, in Passage 4. These results seem to suggest that learners with higher inference accuracy use more strategies when meaning inference is difficult, but the increased strategy use does not necessarily lead to higher inference accuracy. This interpretation also implies that learners with higher inference accuracy use a smaller number of strategies when meaning inference is not difficult. Conversely, when learners with lower inference accuracy use more strategies, the increased strategy use seems to lead to higher accuracy in meaning inference.

Another important finding is that learners with lower inference accuracy use strategies more frequently than learners with higher inference accuracy, supporting Bengeleil and Paribakht (2004) and Hamada (2009). The mean strategy use per word was the highest in Cindy (2.9), followed by Adam (1.73) and Beth (1.68). Taken together, these findings add further perspective on the frequency of strategy use and inference accuracy (e.g., Haynes, 1993; Morrison, 1996), implying that for learners
with lower inference accuracy, more uses of strategies makes their meaning-inference more accurate.

The third research question asked whether the type of strategies used by L2 learners would change over time. Results regarding this question were similar to existing research (e.g., Chern, 1993; Nassaji, 2004). As for local vs. global strategies, the results showed that Adam used contextual analysis approximately 50% of the time with very few local strategies, while Cindy used morphological analysis, word-form analysis, and syntactic analysis in every session, in addition to contextual analysis, suggesting that global strategies tend to lead to higher inference accuracy. The fact that Cindy used virtually all of the types also shows that learners with lower inference accuracy use a wider variety of strategy types (e.g., Haynes, 1993, Morrison, 1996). Moreover, that Beth’s contextual analysis use increased over the sessions (from 12% to 20%) seems to be indicative of a shift from local strategies to more globalized strategies, consistent with Huckin and Bloch (1993).

The number of strategy types used over the sessions showed an important change in the three students’ meaning-inference behavior. Whereas Adam used either 3 or 5 types in the sessions, the number of types used by Beth and Cindy clearly decreased (6 to 4 types for Beth, 7 to 5 types for Cindy). These results seem to support the claim that learners with higher inference accuracy have a more consistent approach (e.g., Hamada, 2009). That is, learners with higher accuracy selectively use strategies that are more effective rather than trying all possible strategies. Lastly, consistent with Hamada (2009), metacognitive strategies were used very little by the students in this study.

**Conclusions**

This study provided a qualitative inquiry into the development of L2 word-meaning inference while reading, focusing on inference accuracy and the number and types of strategies used by three college-level ESL students. Overall, none of the students showed a considerable increase in meaning-inference accuracy rate, inconsistent with Fraser (1999) but partially supporting Hamada (2009). The reason for the lack of increase is attributed to a possible “threshold” effect, suggesting that the students might have already possessed a certain level of inference ability that was difficult to be improved upon in a 4 week period. Further research in a more carefully controlled design will be necessary to verify the current finding.

Regarding the type and frequency of strategy use, this study provides valuable
developmental perspectives to the existing research. The findings demonstrated clear differences in strategy use between learners with higher inference accuracy and learners with lower inference accuracy. Learners with higher inference accuracy used fewer strategy types consistently, while learners with lower inference accuracy used a wider range of strategy types more frequently. Also, the findings appear to indicate that learners with higher inference accuracy have a consistent approach but increase the number of strategies used in difficult passages, whereas learners with lower inference accuracy, in general, use a wider variety of strategies more frequently, and that increased strategy use leads to an increase in inference accuracy. Nevertheless, the fact that both Beth’s and Cindy’s types of strategies used decreased over the sessions indicates that strategy use by learners with lower inference accuracy becomes more consistent as they experience more meaning-inference training.

Another important finding regarding the type and frequency of strategy use is that learners with higher inference accuracy used global strategies more frequently, but rarely used local strategies.\(^1\) This finding confirms the claim that global strategies are more effective than local strategies (e.g., Chern, 1993; Haynes, 1993) and suggests that L2 meaning-inference instruction needs to stress the importance of global strategies and encourage learners not to rely solely on local strategies. Lastly, limitations and future research suggestions are addressed. As mentioned earlier, the present study is limited in that it provided a qualitative inquiry into the development of L2 meaning-inference. In order to confirm the effect of meaning-inference instruction in inference accuracy, future study needs to involve a cross-sectional design and control the texts in terms of linguistic difficulty.

Notes
\(^1\) One of the reviewers questioned whether it is possible to differentiate between local and global strategies. Current research uses the analysis of think-aloud protocol to categorize different types of strategies (e.g., Fraser, 1999; Lawson & Hogben, 1996; Pressley & Afferbach, 1995, Nassaji, 2003, 2004). Following this tradition, we also used a think-aloud protocol in order to observe the participants’ strategy use, following the analysis schema (see Table 2). As long as the protocol included any of the features detailed in the schema, the usage of the corresponding strategy was recorded. As noted earlier in this article, multiple strategies could be used to infer one word.
References
Prospect, 21, 58-75.


Table 1

Participants’ Characteristics

<table>
<thead>
<tr>
<th>Name</th>
<th>Age</th>
<th>Major</th>
<th>TOEFL</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(max. 40)</td>
</tr>
<tr>
<td>Adam</td>
<td>23</td>
<td>Architecture</td>
<td>33</td>
<td>42</td>
</tr>
<tr>
<td>Beth</td>
<td>20</td>
<td>Education</td>
<td>27</td>
<td>36</td>
</tr>
<tr>
<td>Cindy</td>
<td>37</td>
<td>Mathematics</td>
<td>25</td>
<td>29</td>
</tr>
</tbody>
</table>

Table 2

Strategy Classification and Examples

1. Identifying

   (a) Repeating: The learner repeats any portion of the passage, including the word, the phrase, or the sentence in which the word has occurred. Example:
   “rapacious. The nations we used to consider Cold War allies are now merely rapacious trading partners…merely rapacious”

   (b) Morphological Analysis: The learner attempts to figure out the meaning of the word by analyzing it into various morphological components, such as roots, affixes and suffixes. Example: “inequities. Equity can be equal but the prefix in- means ‘opposite’ so it must mean ‘unequal.’”

   (c) Word-Form Analogy: The learner attempts to figure out the meaning of the word based on its sound or form similarity with other words. Example:
   “hormones. It may have the same meaning as ‘harmony’ since they look alike.”
(d) Syntactic Analysis: The learner attempts to figure out the meaning of the word by analyzing its syntactic function. Example: “….so *glides* will be a verb in this sentence…”

(e) Contextual Analysis: The learner attempts to figure out the meaning of the word based on the meaning of its phrasal-, clausal-, or discourse-level context. Example: “…because it is wind… because the dense air *glides* downslope into the valley, sounds like it means the wind blows towards the valley…”

2. Evaluating

(a) Verifying: The learner examines the appropriateness of the inferred meaning by checking it against the wider context. Example: “…. we and the Soviets *vied* for king of the mountain. here … we and the Soviets ‘pursued’ the same things.”

(b) Self-Inquiry: The learner asks himself or herself questions about the word or the meaning he or she has already inferred. Example: “*reared*… For example, a tree squirrel was reared in isolation…then does it mean something that a squirrel can do in an isolated place?”

3. Monitoring

(a) Monitoring: The learner shows a conscious awareness of the problem by judging its ease or difficulty. Example: “*laymen*…. I don’t think it is a simple literal meaning.”

Note: The classification is based on Nassaji (2004). Strategy types 1(b), 1(d), and 1(e) were added for this study.
Table 3

Overall Meaning-Inference Accuracy and Strategy Use

<table>
<thead>
<tr>
<th>Words Underlined and Inferred</th>
<th>Mean Accuracy (%)</th>
<th>Total strategies used</th>
<th>Strategy per Word</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adam</td>
<td>40</td>
<td>51.74</td>
<td>69</td>
</tr>
<tr>
<td>Beth</td>
<td>57</td>
<td>31.64</td>
<td>96</td>
</tr>
<tr>
<td>Cindy</td>
<td>30</td>
<td>30.93</td>
<td>87</td>
</tr>
</tbody>
</table>

Figure 1. Adam’s meaning-inference accuracy rate and number of strategies per word
Figure 2. Ratio of the types of strategies used by Adam

![Figure 2. Ratio of the types of strategies used by Adam](image)

Figure 3. Beth’s meaning-inference accuracy rate and number of strategies per word

![Figure 3. Beth’s meaning-inference accuracy rate and number of strategies per word](image)
Figure 4. Ratio of the types of strategies used by Beth.

![Graph showing the ratio of types of strategies used by Beth.]

- Monitoring
- Self-inquiry
- Verifying
- Contextual analysis
- Syntactic analysis
- Word-form analogy
- Morphological analysis
- Repeating

Figure 5. Cindy’s meaning-inference accuracy rate and number of strategies per word.

![Graph showing Cindy's meaning-inference accuracy rate and number of strategies per word.]

- Success rate (%)
- Strategies per word

The graph shows the success rate and strategies per word for Cindy across different passages.
Figure 6. Ratio of the types of strategies used by Cindy
The Impact of Assessment Change on Language Learning Strategies: The Views of a Small Group of Chinese Graduate Students Studying in the UK

Xiaoli Jiang  
*Renmin University, China*

Gerard Sharpling  
*University of Warwick, UK*

**Bio Data**

Dr. Xiaoli Jiang is an Associate Professor in the English Department of the Foreign Language School, Renmin University of China. She obtained her PhD in English Language Teaching and Applied Linguistics in 2008, following an MA in English Language Teaching at the University of Warwick in 2004. Her research interest includes learner autonomy, learning strategies and academic acculturation.

Dr. Gerard Sharpling is a Senior Teaching Fellow in English for Academic Purposes at the University of Warwick (UK), where he also lectures in Language Testing and Assessment. He was the co-ordinator of the Warwick English Language Test from 2002 to 2010, and has previously taught at the Universities of Birmingham (UK) and Nantes (France).

**Abstract**

Chinese students embarking on further studies within an English-speaking higher education environment face significant changes in assessment. This study, undertaken at University of Warwick (UK), reports on Chinese graduate students’ retrospective views of their developing language learning strategies, in the light of changes in assessment during their courses. The study charts the students’ perceptions of their own experiences over one year of study, beginning with their preparatory English course and ending upon completion of their Masters’ degree programme. The findings of the study show that the College English Test (CET) in China remains fixed within the learners’ mind-sets, at least in the early stages of their study, but that greater attention is paid to process-oriented learning strategies as their academic studies progress. The increased use of process-oriented strategies is closely connected with the nature of the changing learning environment available to the students, as well as the increased use of formative assessment. Findings from the interview data suggest that the learners demonstrate a variety of approaches to their studies, and that their strategies are, in all likelihood, shaped as much by individual, contextual and pragmatic factors as cultural ones. In spite of this diversity, participants do not seem to recognize the potential transferability of product-oriented learning strategies to their later academic studies.

**Key words:** assessment change, formative and summative test, strategy use, overseas students, Chinese learners
Introduction
Chinese students embarking on further studies within an English-speaking higher education environment face significant changes in assessment, both within language programs and in their subject-based academic studies. They are required to adapt from a system where summative assessment (typically, end of course exams and tests) plays a leading role, to one in which there is a greater balance between formative assessment (written coursework assignments, and to a lesser extent, oral presentations and discussion-based tasks) and final exams or tests. Such a change may create a sense of learning and culture shock (Q. Gu, 2005). This paper reports on a small-scale study, undertaken at the University of Warwick (UK), between 2006 and 2007, to investigate how the assessment changes encountered by a group of Chinese students shaped, and were shaped by the participants’ language learning strategies (LLS). Language tutors and subject lecturers who are unfamiliar with teaching Chinese learners may anticipate relatively restricted LLS use by their students, but the spectrum of LLS utilised by these students is more diverse than their British tutors at first seem to acknowledge; it thus seems essential for professionals within English-speaking higher education settings to fully appreciate the attributes of their students. Data was collected through interviews with a small sample of Chinese graduate students. The participants had completed their one-year Masters programmes at the university, and were asked to look retrospectively at their developing learning experience over time, which began with a preparatory five-week course in English. We assume that assessment change brought about by environmental change might influence students’ LLSs. Two key questions of concern are: Do learners in this position change their LLS once they are assessed differently, such as in an English-speaking environment? Or do they elect to pursue use of their existing strategies regardless of any change in the assessment process?

This paper contributes to wider knowledge about the relationship between assessment change and LLS use by Chinese graduate students within the UK and other English-speaking countries. The study makes four contributions to the field of studies concerning Chinese students’ LLS.

1) The study is designed to be informative to both language teachers (those who teach and assess students on their English proficiency) and subject lecturers (who assess students on content-based material on their degree programmes). Language teachers and lecturers may have little experience of working with Chinese students, and may thus have stereotypical views about their LLS. Although only Chinese
students are discussed, this study may be repeated with students from other first language backgrounds, thereby allowing language teachers and lecturers to evaluate their own students’ LLS more effectively; to provide more targeted feedback on their students’ work; and to pinpoint more readily those skills which the students need to develop to further their chances of academic success.

2) This research fits into the field of test washback studies. The term ‘washback’ refers to the impact of a particular test on teaching and learning within the classroom, the education system and wider society. Weir (2005) regards washback studies as valuable, in that they show us how learners approach tests and assessments, whilst “comparing approaches to tests with the way the same learners approach language-based tasks within the wider world” (p. 269). In China, all non-English major graduate students in China are required to take the CET-4 by the end of the first two years of their College English language course. As with all high stakes tests, the CET-4 (and the higher level CET-6) play a significant role within the mind-set of Chinese university students, both during and after their studies in China, and are a key external factor that affects their LLS.

3) The current study seeks to shed light on the specific nature of LLS, and how they can be developed beyond the immediate, “new” experience of a primarily language-based preparatory English course, as students progress through their academic programmes. Once they come to an English-speaking university context, students are regarded as being involved more in “assessment” than in “testing”. According to the Standards for Educational and Psychological Testing (1999, p. 3), a test is an “evaluative device or procedure in which a sample of an examinee's behaviour in a specified domain is obtained and subsequently evaluated and scored using a standardized process.” These same standards (1999, p. 172) define assessment as “any systematic method of obtaining information from tests and other sources, used to draw inferences about characteristics of people, objects, or programs.” Changes in assessment at university level occur most prominently where students need to refocus their attention on subject-specific content, as well as on language-based assessment. However, even where this change occurs, language skills continue to be important in assessing students’ work, as may be seen in the Masters’ degree criteria in Appendix D.

4) Fourthly, this study contributes to a growing awareness of the methodology needed to investigate the complex nature of language learning, which involves students’ use of a mixture of cognitive, social and metacognitive strategies. In a recent
article on learner beliefs and language learning in *Asian EFL Journal*, Ellis (2008) argues that the relationship between learners’ beliefs and their language proficiency remains largely underexplored. As Ellis notes, this relationship is best investigated through qualitative means, such as interviews, rather than quantitative ones such as questionnaires. Ellis (2008) recommends a “contextual” approach to investigating students’ learning strategies, along with a broadly qualitative framework of enquiry. As Ellis (2008, p. 2) argues, the contextual approach “views learner beliefs as varying according to context” and “involves collecting a variety of data types and diverse means of data analysis.” Unlike Ellis’s (2008) study, our research does not focus specifically on learner beliefs. However, the learner beliefs referred to by Ellis (2008) may be seen in the sample group’s comments on language learning; indeed, the learning strategies they employ may be seen as the operationalisation (or otherwise) of those beliefs.

**Literature Review**

Before outlining the research procedures, some key themes in the relationship between learner strategies and assessment change will be explored. The following broad areas will be considered: a) the impact of language assessment on learners within China, with particular reference to the washback effect of Band 4 of the College English test (CET-4); b) the interconnectedness of language assessment and strategy use, with reference to the interplay between surface and deep learning strategies; c) Chinese students’ strategy use in tests such as the CET, and the implications of this for their new context; and d) the concept of assessment change, insofar as it relates to the experiences of Chinese learners. These themes, whilst seemingly disparate, allow several variables affecting Chinese students’ learning strategies to be brought together. As will be suggested, the participants’ previous experience of English tests and assessments in China, especially the CET-4, will be seen as underpinning their perceived learning experience within the new context, and will consciously or subconsciously shape their use of language strategies, in both positive and negative ways.

**Washback of CET in China**

In recent years, discussions concerning the washback of the CET on LLS have become conflated with issues around the validity of the test. The CET is thought to cause worry and anxiety among Chinese university students, and to be associated with
test dominance, and negative washback. The test is regarded by some as limiting students’ attention unduly to “book knowledge” (Huang, 2007; Ruan, 2007), while it can also draw teachers’ attention unfavourably to the test’s importance (Gan, Humphrey & Hamp-Lyons, 2004). Moreover, the newly reformed CET oral test, whilst useful, is taken only by students who have gained 85% or above in the written part, whereas many teachers and students feel that the oral component would be beneficial to all test takers. It is also felt by some commentators that there is over-use of multiple-choice testing in the CET (Han, Dai & Yang, 2002), and a predominant focus on accuracy and form in the assessment of writing skills (Cai, 2002).

Set against these criticisms are more positive views of the test. The College English Testing Committee, for instance, contends that the CET “has been a good measure of students’ ability to communicate in English” (Jin and Yang, 2006, p 22). Moreover, the CET is reported to have maintained high reliability and validity (Yang and Weir, 1998) and has a well developed and widely respected set of standardized procedures for administering and interpreting raw scores (Yang and Jin, 2000). X. Gu (2005) shows that the CET is held in high esteem by most stakeholders, who appear to value its rigorous design, administration, grading and the new measures adopted in recent years, specifically the availability of oral assessment. Where negative washback does exist, they argue that this is primarily associated with test misuse, rather than the construction of the test itself.

Although it is useful to bear in mind the strengths and the weaknesses of the CET, washback issues are not strongly correlated with a test’s validity. This means that regardless of the CET’s validity, it remains an important factor that influences students’ LLS in tangible ways, though these may not always be easy to measure. This will be further considered when discussing students’ LLS at an early stage in their academic courses within UK HE institutions.

‘Learning for the exam’: surface or deep strategies?
Language learning strategies may be defined as “specific actions taken by the learner to make learning easier, faster, more enjoyable, more self-directed, more effective, and more transferable to new situations.” (Oxford, 1990, p. 8). According to Cohen and Macaro (2007), LLS help learners to “enhance learning, perform specified tasks, solve specific problems, make learning easier, faster and more enjoyable, and compensate for a deficit in learning” (pp. 38-39). The choice of strategies depends on contextual factors, individual factors, and the learning goal itself (Griffiths, 2008).
High stakes tests, such as the CET, exert a significant influence on LLS. One crucial question that emerges is whether the procedure for preparing for exams such as the CET encourages surface learning strategies, premised mainly on the need to pass the test, or whether deeper language learning is promoted. It may seem unusual to apply this distinction to language learning, given that it is applied most usually to the study of learning in content courses. However, a number of researchers have shown that the surface learning/deep learning dichotomy may be specifically applied when examining language learning contexts (e.g. Jonasson, 2004; Aharony, 2006; Qingquan, Chatupote and Teo, 2008).

Surface learning is most commonly attributed to learners in examination-dominated contexts. Li (2007) defines rote learning, when applied to LLS, as involving “repetition of target language items either silently or aloud” or “writing down the items (more than once).” Aharony (2006, p. 853) sees this type of learning as characterised by “a student’s tendency to choose the quickest way to accomplish the task; to acquire the learning material without asking in-depth questions, to study the material in a linear manner; to relate to minimal aspects of material or to a problem without showing interest; or the need to understand it in its entirety; to learn by rote by relying on memory and not on comprehension; and to be concerned with the time needed to fulfill the learning task.”

Within test-dominated contexts, learning is likely to be characterised by a more overt focus on test success, which leads candidates to engage in activities such as recycling notes and memorising chunks of language. Behaviour is frequently goal oriented. This in itself raises problems because, as Cohen and Macaro (2007) remind us, learners cannot always articulate goals because they are not aware of the specific strategies they use. In addition, strategies commonly thought to be surface learning strategies (rote learning, training in examination skills, memorisation) are viewed, within China and other countries, as a strong means of empowering the learner, and promoting knowledge enhancement (Marton, Dall’Alba and Tse, 1993; Lee, 1996; Au and Entwistle, 2001; Li and Chang, 2001). The challenge for teachers is to develop these strategies so that they can be applied to more critical fields of enquiry. Kember (1996) shows that deep and surface approaches are not so much an either/or binary opposite, but are part of the same phenomenon which is combined in different ways by different learners. As Li and Chang (2001) further remark, memorisation among Chinese students operates at both deep and surface levels. When used appropriately, memorisation is a crucial facet in enabling learners to acquire vocabulary, to help
them to gain greater exposure to materials through a process of repetition, to provide psychological comfort, and to train and develop learners’ thinking skills. Thus, learning for any test has beneficial (as well as harmful) effects.

A strong indication of the ability of the CET candidates to switch back and forth between goal-oriented and process-oriented strategies is given by Tang (2005). While Tang acknowledges that the CET constrains the contents, methods and pace of language learning, given its high-stakes nature, he argues that learners are, contrary to expectations, willing to engage in activities that are not directly associated with the CET (such as watching films, practising speaking and extensive reading). Thus, we may assume that the ability (and desire) to switch strategies is latent in many students, and is not necessarily ‘triggered’ by their presence within an English-speaking environment.

**Chinese students’ strategy use in English language learning in CET 4**

A further debate, when discussing Chinese university students’ LLS, is whether their strengths derive from cultural, contextual, personal or individual factors, or perhaps a combination of all of these. This debate cannot be readily resolved; for example, Bedell and Oxford (1996) indicate that ethnicity, as well as culture, strongly shapes LLS. Some early descriptions of Chinese learners as preferring to use surface strategies such as rote learning (e.g. Biggs, 1996; Cortazzi and Jin, 1996) are rather limited, and other empirical research studies reporting Chinese learners’ seemingly passive learning strategies (Hu, 2002; Rao, 2006) are not altogether satisfactory. Fat (2004, para. 11) argues that Chinese students are “more inductive” than their Western counterparts, and have a “binocular” approach, which tends to see “both sides of the coin”. Many such studies find that Chinese official English tests are important factors in shaping students’ LLS. Cortazzi and Jin (1996) have noted that university students perform rote learning of vocabulary lists required for the CET-4. Likewise, Rao (2006) has reported that Chinese university students’ strategy use, such as vocabulary revision, texts and note reciting, as well as doing exam papers, are related to the specific needs of examinations. This association has also been observed by researchers when indicating how mechanical memorization was used to prepare for the much revered Civil Service Exam, which led to highly coveted governmental positions (Lee, 1996; Hu, 2002; Rao, 2006). None of these studies suggests any possible benefits that such strategies might bring about for future language development.
Set against these studies are those which place greater emphasis on the individual and personal characteristics of Chinese students. Research into LLS most often shows that learners have a wide range of individual differences, which occur in specific, unpredictable ways (Benson and Gao, 2008). Hu and Chen (2006) have indicated that stronger and weaker writers preparing for the CET-4 adopt a range of strategies, and do not have recourse to a typecast “blueprint” in terms of what strategies should be used. Jin and Cortazzi (2001, para. 5) have sought to revise and refine their culture-based consideration of Chinese learners, to include a more positive view of their attributes: more recently, they see the strengths of Chinese learners as residing in their ability to achieve “balances between modelling and memorizing; mimicry and mastery with postponed creativity; student preparation and performance, student independent learning and collaborative learning in certain ways.” Huang and Sisco (1994) observe that Chinese students have a somewhat broader spectrum of learning and thinking styles than is often anticipated, even if Chinese students are sometimes seen as more “pragmatic” in their approach to their studies. Huang (1997) notes that Chinese students have an even better ability to categorise information in broad terms than their American counterparts, and in terms of learner styles and preferences, have an equal preference for group work and collaborative learning.

Assessment change
A fourth area of literature relates to how learners are affected by assessment change when working in a different context. Relatively few specific studies have been conducted to analyse the phenomenon of assessment change. Gu (2005) explores the notion of “learning shock” prevalent in the acculturation process, and argues that “Learning shock refers to some unpleasant feelings and difficult experiences that learners encounter when they are exposed to a new learning environment” (p. 42). Kember (2000) provides a revealing account of Action Learning Projects in Hong Kong, demonstrating the challenges of students moving from tightly structured courses, in which content is defined by the teacher, to new and more innovative classroom methods, with a greater focus on reflective writing. As he notes, students initially found the new course formats taxing, but eventually came to appreciate them. Hitchcock and Cross (2007, p. 9) surveyed students’ views of the “acculturation” process within Portsmouth University, UK, suggesting that one of the most obvious differences between Chinese and British university systems is that of the diversity of assessed tasks encountered in the latter system. As the authors observe: “Many
students noted clear differences in the assessment methods of Chinese and British universities with, in their view, Chinese institutions being more likely to use traditional written exams while British universities may use a range of assessments, some of which require spoken language skills (e.g. presentations), and some of which do not require the reproduction of knowledge (e.g. reflective or review tasks).” Despite the promising nature of all of the above studies, little consideration is generally given in the literature to the empowering nature of Chinese students’ existing study skills and LLS, and the way in which the students’ development through a test-oriented culture actually helps, as well as hinders their progress. Yet the interconnectedness between the two systems, old and new, may indeed lead to the development of more positive LLS.

**Purpose of the study**

Thus far, the literature review has drawn out four key themes that are relevant to Chinese students’ experiences of assessment change. First and foremost, the CET in Chinese universities, with its emphasis on summative assessment, creates a strong washback effect, regardless of how one sees the quality and validity of the test. This is likely to exert a high degree of external influence on students’ learning strategies, for better or worse. Secondly, the question of whether a test-oriented culture is helpful for language learning is made more complex, owing to the fact that strategies such as memorization and rote learning are often considered as positive, empowering devices within China. While it is particularly useful to consider what such procedures may bring positively to the learning experience, these considerations are often left out of account in recent studies. Thirdly, we see that in China, the impact of a test such as the CET will elicit a mixed range of attitudes and responses from students, and will encourage the development of different learning strategies. This seems to diverge from earlier, more deterministic attempts to explain Chinese learning strategies through dimensions such as Confucian thought. Finally, the literature review raises the recurring issue that while high stakes exams strongly influence teaching and learning, more research needs to be conducted to show how students’ learning behaviours can be influenced by changes in assessment, and what their perceptions of assessment change might be. It is hoped that the procedures outlined in this study may be replicated by other EFL teachers within similar contexts, in order to further sharpen the picture of the attributes inherent in Chinese students’ learning practice.

On the basis of the above concerns, the research questions of the present study were
formulated, as follows:

1. What are the main differences between the types of assessment encountered within the UK and China, as perceived by the participants of the present study?
2. How do these differences shape students’ developing LLS use during a year of study?

Research Methodology

Guided by the above research questions, the present study used primarily qualitative enquiry, with semi-structured interviews providing the main data source. The interviews were conducted after the students had completed their Masters programmes. The participants were 8 Chinese students, all of whom had finished their first degree in China and also their Masters’ degree studies in the UK. They had all taken a 5-week preparatory English language course, which required completion of a written project as a primary means of language evaluation, and an oral presentation to assess their speaking skills. Some candidates were also required to take the Warwick English Language Test (WELT), a summative in-house proficiency test of grammar, reading and writing at the end of the course. Following this, they moved to their academic departments, which heralded a further conceptual shift, with “written assignment” being the main (though not exclusive) means of assessment. Whilst in their departments, the students’ assessments focused more around their understanding of the course content, though their English language skills continued to play a role in their ongoing assessment. From our discussions with the students, both before and after their Master programs, we found that specific LLS continued to be uppermost in their minds.

The two researchers in the present study occupied different roles during the preparatory course attended by this sample of students: one was a Residential Tutor, while the other was a Course Tutor. Although the students were interviewed after the completion of their Masters programmes, they were selected for the study at an earlier stage, during their English course, on the basis of the following three criterion: first, they were Chinese students; second, their English level differed, even though they were following the same language course; and third, they were about to join a range of Masters’ programs. One researcher (the Resident Tutor) observed the students’ casual talk during their after-class time, and attempted to befriend this group of students. This friendship was maintained during the students’ Master program study, and informal gatherings were held occasionally. Interview data were collected at the
end of the participants’ program of study. The profiles of the participants are given in
the following table, and the interview questions are provided in Appendix A.
Interviews were conducted in Chinese, and were recorded with the permission of the
participants. Each interview took approximately 25-30 minutes to complete.

Table 1: Profile of participants (M, Male; F, Female; WMG, Warwick
Manufacturing Group; MT, Multimedia and Theatre Studies; FM, Financial
Mathematics; AF, Accounting and Finance)

<table>
<thead>
<tr>
<th>Participant No.</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td>M</td>
<td>M</td>
<td>F</td>
<td>F</td>
<td>M</td>
<td>M</td>
<td>M</td>
<td>F</td>
</tr>
<tr>
<td>Course</td>
<td>WMG</td>
<td>WMG</td>
<td>MT</td>
<td>WMG</td>
<td>FM</td>
<td>Law</td>
<td>FM</td>
<td>AF</td>
</tr>
<tr>
<td>CET4/6</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>IELTS</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>WELT</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
</tbody>
</table>

The learners’ voices contributed significantly to the researchers’ understanding of
language assessment, and how this might influence their LLS use. However, we were
aware of the need to adopt other measures to enhance the credibility of our research.
For instance, we drew upon multiple data resources (Bogdan and Biklen, 2006) by
observing participants’ English language learning during after-class time with
reference to the guide for the preparatory English course (Appendix B). We observed
participants’ after-class casual talk, and also referred to the participants’ course results
in their Masters programs (Appendix C).

16 hours of interviews were carefully transcribed by the researcher who undertook
the interviews. Cameron (2001) and Bird (2005) suggest that transcription should be
consistent with the research conventions that researchers adopt. Since the purpose of
the present research was to examine students’ views of LLS in different assessment
environments, transcription provided a focus on content, rather than conversation
analysis or discourse analysis. Transcriptions were sent back to the interviewees for
cross-checking of accuracy (Sikes, 2000) before they were forwarded to the other
researcher. Following this, each researcher analysed the data independently. Data
analysis consisted of open-coding, theme elicitation and co-judgment. First, the
researchers took the position of “finding codes from data” rather than “bringing codes
to data” (Punch, 2005, p200); then, informed by the techniques proposed by Strauss
and Corbin (2000) in their paper on grounded theory, the researchers repeatedly read
the transcriptions, combined focused interpretation with overall judgment to ensure
theme synthesized systematically; third, themes identified by two researchers were
compared and discussed, with reference to the original research questions. Through
several rounds of discussion and exchanging ideas, final categories of themes were established, as presented below in Table 2.

**Table 2: Themes identified in the course of the research**

<table>
<thead>
<tr>
<th>Theme</th>
<th>Interviewee</th>
<th>Assessment mentioned</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Types of strategy use</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Product-oriented strategies</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Timely vocabulary</td>
<td>1, 2</td>
<td>CET</td>
</tr>
<tr>
<td>Memorization</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Exam practice</td>
<td>4, 5, 7, 8</td>
<td>IELTS, WELT</td>
</tr>
<tr>
<td>Training in exam skills</td>
<td>1, 2, 3, 4, 5, 6</td>
<td>CET, IELTS</td>
</tr>
<tr>
<td>Process-oriented strategies</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Content-based English language learning</td>
<td>3, 6, 8</td>
<td>Assignment</td>
</tr>
<tr>
<td>Accessing English language environment</td>
<td>3, 7</td>
<td>Exams and assignment</td>
</tr>
<tr>
<td>Using English</td>
<td>1, 2, 3, 4, 5, 6, 7, 8</td>
<td>Course project, Assignment</td>
</tr>
<tr>
<td><strong>Role of English</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>as a surviving tool</td>
<td>1, 2</td>
<td>Assignment, MSc. thesis</td>
</tr>
</tbody>
</table>

These themes will be further discussed in the following section.

**Findings**

After the interviews, the data was carefully analysed; one aspect of the data that immediately appeared noteworthy was the clear distinction in participants’ reported strategy use between aspects such as ‘learning for the test or exam’ and those which involve gaining greater exposure to English. These types of strategy are referred to, for convenience, as product-oriented learning strategies and process-oriented learning strategies. It is important to emphasise that there was no direct correlation between process-oriented LLS and the availability of increased formative assessment within the English-speaking environment; nor was there any linear development from product-oriented LLS to process-oriented strategies LLS, since it was possible for any of these strategies to occur at any stage in the learners’ studies. Moreover, the LLS identified are not unique to Chinese students; they may well be adopted by other students who come from different first language groups. Further research in different contexts will help to ascertain how far LLS are shared between different first language groups.
Product-oriented learning strategies

Three main product-oriented language learning strategies were unanimously reported by participants in the present study, regardless of the department they were studying in or the subject area they were reading: these were timely vocabulary memorization, exam practice, and engaging in specific training in exam skills. All of these strategies seemed to be closely related to the participants’ main purpose in learning the English language, namely to pass high-stakes exams such as the CET Band-4 and IELTS (or equivalent). As the participants noted, such strategies were most prevalent during the students’ preparatory English course, where language learning was seen as a priority. The participants’ mentioning of these strategies also indicates the interconnectedness between their previous CET-related studies and their early encounters with a very different kind of system within the UK.

Timely vocabulary memorization

The data revealed that Chinese learners used vocabulary memorization strategies to organise and structure their learning. The rationale for using such a strategy was narrated by interviewee 2:

Vocabulary is everything. It determines all the other parts of your performance in exams. For example, the vocabulary and grammar section is directly linked with it….if you know more vocabulary, then you won’t have problems doing reading comprehension…you can express yourself in writing too.

This participant recalled the importance of vocabulary knowledge in the CET-4. There are lengthy, unofficial lists of vocabulary available to help test takers prepare for the CET, and teachers are also given a defined vocabulary syllabus with headwords, in unclassified alphabetical order. Many learners simply learn such de-contextualised vocabulary lists by heart, in case the words come up in the examination. However, unlike other research reports, which recall the debate as to whether Chinese learners used rote or meaningful memorization (e.g. Marton, Dall’Alba & Tse, 1996), the present study shows that students were less concerned with how to memorize vocabulary than finding the appropriate time to do so. All the participants appeared to agree that memorizing vocabulary shortly before taking an exam was the most effective strategy. As interviewee 1 commented:

I didn’t learn English until there were exams. It’s about vocabulary
Upon further probing as to the reasons for this, he responded:

_You would forget anyway. To memorize before exams you will have a deeper impression._

Despite the participants’ overt focus on the need to learn vocabulary, it was noticeable that the interviews revealed little information concerning the value of learning and integrating subject-specific vocabulary later; nor did the participants indicate many of the advantages of learning vocabulary for overall language acquisition, outside of the immediate test framework. Furthermore, the interviews did not reveal whether these approaches might be different now that vocabulary is no longer set as a discrete part of the CET.

**Examination practice**

Where summative testing is high on the agenda, a further important strategy is that of undertaking examination practice. It is often believed that learning can be enhanced simply by exam practice. In line with this presupposition, all interviewees in the present study considered one of the compulsory prerequisites for good exam results as involving the use of relevant exam papers. According to them, the basic purpose of this was to familiarize themselves with the format of exams. As interviewee 5 noted:

_Like IELTS, you would have no idea as what was to be tested if you did not do some simulated papers. The more you did, the better sense you got...you would know what to prepare for._

This view was supported by interviewee 7, who observed that different types of exams placed emphasis on different aspects of learning. Therefore, it was considered useful to know the most important components of exams by doing exam papers.

_CET-4 emphasises vocabulary and grammar. However, IELTS is more practical, with the oral test and writing parts all stressing your language use. You should do exam papers in order to know different focuses in different exams._

In drawing upon an ancient Chinese saying 知己知彼，百战不殆 (Only by understanding your opponents and yourself can you win the battle), interviewee 8 considered that taking exams was akin to a psychological battle between the exam takers and the test designer, and a critical factor was to ascertain the latter’s intentions by trying out relevant exam papers. Interestingly, few interviewees believed that they could prepare for the Warwick English Language Test (WELT.) A common reason for
this was that they could not readily locate simulated test papers to practice on, given that WELT is a small-scale test.

Training in exam skills
A further important aspect, as noted by the interviewees, was the need for training in exam skills. Among other skills, reading was most frequently mentioned by the majority of the interviewees, since it appeared to them to constitute the largest proportion of the total score in all exams. According to interviewee 6:

\textit{It is about your skills to find out the right answers even without understanding the whole passage.}

The above comment demonstrates the inherent perception that examination skills and learning strategies are different: one can learn how to pass an examination with minimal engagement with the input or stimulus material. The participants were not so much concerned with their ability to understand the reading passages as to develop the skills to answer comprehension questions. This, according to interviewee 1, was because

\textit{...time is so short. How can you finish reading them all? There are some skills to answer those questions, for example, only reading topic sentences and concluding sentences, and taking key words of the questions to the text.}

It might have been possible for the participants to relate the technique of speed reading required in the CET to their future academic studies, where the skills of skimming were paramount, but for the participants, the process of completing the paper was seen as most important.

An additional skill mentioned was to prepare language models prior to the exams, including oral test models and writing models. As stated by interviewee 2:

\textit{You know the format of those oral questions. Just follow the samples, memorize some typical answers, and mix them in your answers to the real questions. The writing is the same.}

Indeed, in the students’ view, since different types of exams have different exam focuses, exam skills appeared to vary accordingly. When taking the CET-4, for example, some interviewees (e.g. 2, 3, and 6) reported that the CET-4 stressed students’ abilities to distinguish synonyms or words that had close meanings; therefore, the focus of their training was to ascertain the common tricks test designers intended to play in terms of vocabulary or phrase usage. Likewise, for IELTS, interviewees (e.g. 2, 3, 4, 5, and 6) considered that speaking and writing were the
most strongly emphasised skills, and that they had to train themselves in these aspects. Regardless of the participants’ awareness of the way in which their strategies changed depending on the test they were preparing for, there was little evidence to suggest the transferability of the examination strategies adapted to other learning contexts.

Whilst all interviewees emphasised the paramount importance of employing a variety of LLS to improve their exam results, they were aware that genuine ‘good’ English came from an accumulation of learning in the long term, rather than simply ‘shining the sword just before the battle’. As concluded by interviewee 4:

Perhaps you can prepare some ready-made answers; you can hardly guess what the examiners’ responses are. So it still depends on your real English proficiency. I heard from the training course that IELTS markers were told to deduct marks for those memorized answers, either in oral exam or writing.

This comment indicates an awareness of the difference between positive and negative memorisation. However, as with the previous product-oriented strategies mentioned, the participants gave little or no indication of how memorisation of models might impact positively on their future academic programmes, and their later focus on more extensive reading.

**Process-oriented learning strategies**

As distinct from the product-oriented learning strategies described above, interviewees also reported a range of process-oriented learning strategies, which tended to develop gradually over time, in response to an increased emphasis on formative, as well as summative assessment during their Masters programs. Further examination revealed that these strategies consisted primarily of content-based English language learning, accessing the English language environment, and using English in a practical context.

**Content-based English language learning**

Several interviewees commented on the time they spent searching for English materials for their Master courses. This was regarded as a natural process, since most of the participants were required to write assignments at various stages in the year such as end-of-term paper and the final Master thesis. With the exception of interviewee 5 and 7, all other interviewees were assessed by written assignments in their Masters’ study in the UK. According to these students, it was difficult to obtain a
good grade for an assignment, because they needed not only to understand and document a number of items of literature written in non-native language, but also to try to adopt critical thinking strategies. For example, as interviewee 6 observed:

*In our major, we have to read lots of things. Even though you have done that, you do not necessarily know how to answer the assignment questions, but in general, the more you read, the better sense you have.*

This comment shows the idiosyncratic nature of writing assignments. Often, the insufficiency of language skills alone leads to an inability to answer the question. Interviewee 8 shared similar feelings in searching for English materials to complete tasks required by the Master courses. In her opinion, improving English was no longer the only prominent goal. She felt that her English would automatically improve when English became a major tool used for studying. As she noted:

*You do not particularly need to learn English for the sake of improving it. Now English is the medium, you naturally have contact with it all the time. There is quite heavy pressure from your major and you do not have particular time to learn English.*

As distinct from interviewee 6 and 8, who held that content-based English language learning in itself contributed to their English language learning, since English was read, understood and used in their assignment writing, interviewee 3 expressed a more circumspect approach to English language learning. For her:

*You read a lot because you have to write assignments. However, you can pay much attention to their [author of article] language use, and how they write. You imitate them a few times, and then it became yours.*

These comments reveal the resourceful way in which many candidates read prescribed texts: not only for the instrumental purposes of passing tests and assignments, but also for the more integrative purpose of developing a broader insight into the language competence.

**Accessing an English language environment**

Although most interviewees expressed the fact that they came to be more and more preoccupied by their main academic subject, especially after arriving in the UK, and did not have specific time for English language learning, some reported attempts to take advantage of the English language environment and learn English in their daily lives. As stated by interviewee 3:

*I had a habit of listening to English tapes even when I was in China. Each*
time before exams I would listen to tapes of English texts here in the UK, it is so convenient to watch English programmes on TV and listen to BBC and I spent some of my spare time on them.

Short, repeated bursts of listening to attune the ear to the sounds of the language seem to be of use to students both within China and the English-speaking environment. Upon further probing as to why she was so concerned about watching English programmes or listening to English radio, this participant responded as follows:

*English is a language and you have to get a sense of it. When we speak Chinese, we seldom think of grammar first then speak it out. English is the same.*

A further means of accessing communicative data was reported by interviewee 7, who stated:

*Wherever I go I will try to listen to people’s talk, on the bus, in the shopping centre. Sometimes it was difficult to understand but the more you hear, the more familiar you are with the tones of spoken English.*

The participants’ responses seemed to reflect the positive attitudes of the CET test takers mentioned in studies such as those of Tang (2005), who had sought contact with English-medium materials such as films and natural speaking situations despite realizing that such approaches had little direct relationships with tests and exams. However, both interviewee 3, who wanted to obtain a wider understanding of the English language, and interviewee 7 who expected to be familiar with spoken English, suggested an awareness of ways to improve English language proficiency. This was an important factor determining their performances in assessment, whether summative or formative.

**Using English**

All interviewees considered that they used more English in their ‘new’ environment than was the case in their home country. This was seen as an inevitable part of their studies, and arose because communication with people from other nationalities was needed. Aside from the fact that English was now the only classroom language, some interviewees (e.g. 1, 2, and 5) mentioned that they were involved in some ‘after class’ group projects involving students from different nationalities. Here, English was used more for discussions. In addition, interviewees recognised the fact that living in a hall of residence encouraged diversity. Naturally, they used English to chat with flatmates.
in their daily lives. As articulated by interviewee 4:

_You will never know how different it was from your textbook language and the language used here if you did not come here. For example, we were told to say ‘How are you?’ for a first time greeting. However, here what we heard was often ‘are you all right?’ At first, I did not understand why people asked me whether I was all right or not because I did not do anything wrong. Later on I realized that it was just a common greeting._

This comment revealed the high value placed on the authenticity of materials by graduate Chinese learners. Moreover, for interviewee 3, using English not only meant finding opportunities to speak the language. She recognised that the authorship of textbooks in the UK differed from that of textbooks found in the Chinese context, and that the written language seemed ‘different’ from the textbooks produced within China. Therefore, in her assignment writing, she intentionally tried to use the language she had learned from the UK textbooks.

Clearly, the way in which preparatory English courses and Masters programs are delivered have an influence on the participants’ LLS. As mentioned by the interviewees in the study, group projects provided opportunities for English use, while assignment writing helped to enhance students’ awareness of appropriate language use from the English used both in daily life and textbooks. As shown in Table 3, students formed their own learning groups for after-class learning, based on the format of the assessment they were about to face. Although interviewee 4 and 8 still attempted to practice exam papers, as they were required to take WELT, they demonstrated process-oriented learning strategies when they joined their respective study groups.

With reference to the assessment criterion for Master programs (Appendix D), English language proficiency remained an important indicator of their final results. Typical examples might be interviewees 6 and 3, who noted that the use of process-oriented strategy use led to positive results in their Master programs (Table 4). Since interviewee 5 and 7 took paper-based exams for their Master program, it is more difficult to judge the connections between their LLS and their academic performances.
Table 3: Observed participants’ performance during the preparatory English course

<table>
<thead>
<tr>
<th>Preparatory English language course</th>
<th>Observation on after-class language learning</th>
<th>Assessment faced</th>
<th>Interviewee</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Searching project-related materials, discussion on projects, group study</td>
<td>Written project and oral presentation</td>
<td>1, 2, 4</td>
</tr>
<tr>
<td></td>
<td>Searching project-related materials, discussion on projects, group study</td>
<td>Written project and oral presentation</td>
<td>5, 6, 7, 8</td>
</tr>
<tr>
<td></td>
<td>Exam practice</td>
<td>Written project and oral presentation</td>
<td>4, 8</td>
</tr>
<tr>
<td></td>
<td>Subject study (such as maths)</td>
<td>Written project and oral presentation</td>
<td>3, 5, 7</td>
</tr>
</tbody>
</table>

Table 4: Participant’s average score in their Masters’ Programs

<table>
<thead>
<tr>
<th>Score range</th>
<th>MA. Language related criteria</th>
<th>MA. Interviewees</th>
<th>MSc. Language related criteria</th>
<th>MSc. Interviewees</th>
</tr>
</thead>
<tbody>
<tr>
<td>70+</td>
<td>Well-structured, very well written</td>
<td>7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>60-69</td>
<td>Generally well-written</td>
<td>3, 6</td>
<td>Fluency</td>
<td>1, 2, 4, 8</td>
</tr>
<tr>
<td>50-59</td>
<td>Fairly well structured, conclusions are reasonable</td>
<td>5</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The extent to which decisive and principled approaches to developing LLS positively affected the participants’ test results and learning requires more detailed
investigation; however, according to preliminary evidence, as revealed from observations (Table 3), the assessment of students’ performances during the preparatory course and their final results for Master programs (Table 4) show that this was the case.

**English for surviving rather than for learning**

In addition to the distinctive categories of reported strategy use, interviewee 1 and 2, in particular, expressed one shared similar idea, namely that they had learned English merely in order to survive the system, rather than to gain language competence. As noted by interviewee 1:

*I don’t think I will have many opportunities to use English in the future. I can read, and can communicate with people. That’s all. I don’t deliberately make foreign friends. In fact, my friends are all Chinese. It is unnecessary for me to learn English for the sake of another language.*

Likewise, interviewee 2 observed:

*I admit that if your English is good you have one more skill. However, I will definitely go back home after finishing this degree (MSc.), so what is the use of spending time learning English? If I can pass (the assignment), then it is enough.*

These two remarks indicate that LLS were in some cases developed primarily for instrumental and pragmatic, rather than social concerns. The participants did not elaborate much on the possible beneficial link between stronger social ability in English and their ability to participate in group work of various kinds within their academic programmes.

**Discussion**

From the interviews, a clearer view of the students’ perceptions of the differences in assessment between China and the UK was obtained. At an earlier stage in their academic year of study, the students’ LLS focused mainly on goal-oriented approaches such as matriculating and passing further language tests in their preparatory English course. Later, they had the time and space to consider more process-oriented strategies. Whilst this study has not sought to provide an in-depth study of all assessment types in both countries, the findings reveal that for these students, assessment appears to play a different role in China and the UK. In discussing these issues with the participants, it seems clear that their own experiences
trace a path from the more summative form of assessment practice in China, which focuses more overtly on language knowledge than language use, to a more formative one on their academic programmes in Britain, where assignments (as well as exams) become an important focus. Such differences seem to have a noteworthy, if immeasurable impact on students’ English LLS. At the same time, it is worth noting that the overall language environment change, from EFL (English as a Foreign Language) to English as a native language environment, also has a key role to play in influencing students’ English LLS (Gao, 2003). The following sections summarise the assessment changes, the participants’ LLS use and their opinions of the English language environment.

**Perceived differences between assessment in China and the UK**

According to the students in the present study, assessment practices in China differ greatly from those used in the UK. For these students, language tests in China are highly focused on linguistic skills such as vocabulary, grammar, and reading comprehension. Testing focuses predominantly on what can be tested reliably and accurately, and with maximum objectivity. This seems entirely understandable, given large numbers of test takers and the need for quick and reliable grading. On the other hand, for the participants, a proficiency test such as IELTS appears to be more inclined to test abilities to use the language, by emphasizing the importance of oral exams and creative writing.

In the course of the research, participants were seen to be more aware of the differences between the types of assessment than they were of the possible beneficial links between the assessment types. The students were, on the whole, inclined to view their product-oriented strategies as somewhat negative, if inevitable, whereas in fact, such strategies may well have positive spin-offs in terms of the students’ academic studies, when such skills could be recalled and used to beneficial effect. Students preparing for the CET may, for instance, choose to use a keyword mnemonic approach to learn vocabulary items: this involves the learner devising a key word for the vocabulary item, a word from the first language which sounds like the word being remembered, and a visual image which links first and second language (P. Gu, 2003, para. 56). Combined with other techniques, this approach might be successful at a later stage, when students are studying in their academic departments. A student following a postgraduate Applied Engineering course, for example, might devise key words when attempting to learn and produce vocabulary items linked to the concept
of Total Quality Management (e.g. “six sigma”, “scatter diagram”, “stratification”, “iterative process”). This level of automatic recall may well be necessary for presentations and discussions, as well as examination answers. Despite its potential usefulness, the way in which vocabulary learning techniques encouraged by CET vocabulary preparation books may be applied to future study situations was not explored by the participants.

**Change of assessment and change of strategy use**

As is often noted (e.g. Shohamy, 1993; Wall, 2000), tests have a discernible impact on teaching and learning, as well as on a variety of stakeholders. The data suggests that students’ reported strategy use was strongly associated with the actual means of assessment. For example, when students were assessed by various high stakes exams such as term exams and the CET-4/6 in the university, their reported strategy use was closely connected to the ‘format effect’ of the exams, which focused primarily on vocabulary, grammar, and reading comprehension (Benson and Gao, 2008). On the other hand, where students had to take IELTS so as to be able to study abroad, their strategy use slightly shifted, and became more pragmatic and skills-based, since IELTS measures spoken English, listening and free writing. However, when students found that they were assessed more by formative measures on their preparatory English course, such as through classroom participation, group work and assignments, as opposed to summative exams, their strategy use took quite a different turn. As participants noted, English language learning on the preparatory course did not involve high stakes exams, so their learning became increasingly shaped by the structure of learning in the UK higher education itself, which involves assignment writing and other pieces of continuous assessment. Since a typical means of assessment is to write assignments, students’ attention increasingly became focused on how to write an effective piece of academic writing, rather than simply learning specific segments of language such as vocabulary and grammar. This development may be evidenced by students’ reported LLS use both in the preparatory English course and Masters programs, as well as the Residential Tutor’s after-class observation (Table 3), and the final results of the students’ Masters courses (Table 4).

To take interviewee 6, for example, this student succeeded in obtaining an upper-second level in all his assignment writing, as required by the Master program. Such a level of achievement would not have been possible simply by adapting previously learned examination skills for the CET-4 or IELTS, though naturally, such
examination skills may equip the student with strategies that could be helpful in dealing with the required language work. Indeed, as articulated by most interviewees, assignment writing in the UK appeared to be the major driving force behind their English language learning, not explicitly focusing on learning language knowledge, but more on meaningful writing.

Nevertheless, when students had to take summative English exams such as WELT (e.g. interviewee 4 and 8) by the end of their English course, they did not feel that their English LLS underwent any significant changes. For one thing, they could not find enough simulated papers to practise on; moreover, they did not seem to experience as much pressure when preparing for WELT as was the case for other types of high stakes exams. As the interviews reveal, the participants talked much about how to survive the latter, whereas little information on WELT was given. This suggests that the participants’ strategies were not affected as negatively by the requirement to take summative proficiency tests, as one might have believed. At the same time, there was no immediate indication that they were shaped positively.

**English environment: a critical but not decisive factor**

Although the interviewees pointed out that changes in assessment led to different ways of learning the English language, along the lines of Gao’s (2003) findings, the availability of an English language environment inevitably had an impact on their English language learning, and this cannot be underestimated. Students soon found themselves in a naturally language resource rich environment, which provided practical support to them in helping them to learn and use the language. For example, the textbooks students used were generally written by native speakers; classroom language was in English; and English was used in the students’ daily life. All these factors contributed to the improvement of students’ English language proficiency. They may even have had a more decisive impact on LLS than assessment change itself. We may see that even though English proficiency tests in China such as the CET-4/6 have undergone certain reforms, for example by incorporating an oral test and enhancing the subjective parts in the exams (Jin and Yang, 2006), students in the present study indicated that their English learning strategies were more likely to change through being in an English-speaking environment than purely because of assessment change.

However, as interviewees 1 and 2 noted, students did not automatically take action to improve their English, even when in an English-speaking environment. This view
reflects that of Ellis (2008), whose study shows that living in a native-speaker environment does not lead to automatic proficiency. Neither does the students’ belief that immersion in English is positive and useful guarantee that they will act on those beliefs. Thus, caution should be taken in making any cause and effect assumptions between the availability of an English language environment and automatic changes in students’ English LLS.

**Conclusion**

The current study has focused on Chinese graduate university students’ conceptions of assessment change and its impact on their learning strategies, from a retrospective point of view. The research suggests the following: 1) participants’ LLS were pragmatic in nature, and shaped to a large extent by the dictates of the assessment system in which they were operating; 2) once based in the UK, the participants were more likely to adopt LLS with which they felt comfortable, but their development of process-oriented strategies became more noticeable over time; 3) participants demonstrated a wide variety of LLS, and these cannot all be directly associated with assessment format, though some strategies were found to be prominent when students were required to undertake formative assessments or project based work; and 4) participants’ adaptation of LLS to the “new” assessment environment was longitudinal, and extended in time; it was strongly, though by no means exclusively influenced by the availability of a language-rich environment. However, for some interviewees, surviving the system rather than deriving full benefit from the English-speaking environment was their main goal. In common with the findings of Ellis’s (2008) study, we find that the experience of studying in the UK gave the students experiences that enabled them to provide a more effective evaluation of their language learning, and where necessary, make relevant changes to their approach; nonetheless, wide individual differences remained.

**Implications**

The above findings have implications for both Chinese students who are about to study abroad, where assessment differs from their home country, and for their tutors, who may be unfamiliar with the way in which Chinese students learn. First, it is essential for Chinese students to be aware of the different assessment system in both language course and subject studies and consciously adapt to the “new” system, which can result in satisfactory results in not only one subject area but also in their
general English level. Second, English language tutors and subject lecturers should fully recognise the attributes of their Chinese students’ LLS. Further learner training in LLS needs to be conducted, in order to help Chinese students to make full use of the English-speaking environment, while tutors and lecturers need to become better acquainted with the diversity of LLS that Chinese students bring to the learning process, and develop whilst they are studying. As Ellis (2008, p. 10) argues, “little learning is likely if there is a mismatch between the teacher’s and the students’ belief systems.” The role of the teacher is thus to identify any differences in beliefs. It may also be useful to see some of the LLS that students develop when taking CET-4 as being directly transferable to the students’ future academic circumstances. Skimming, scanning, careful reading and inferring, as well as listening for gist and specific information, are all encouraged by the current test format of CET-4.

Limitations and further research
Although the present study adopts the use of multiple data sources, for example, combing students’ retrospective interview with one researcher’s (Resident Tutor) observations in their after class time, it might have been better if the other researcher (Course Tutor) had kept a record of the talks with participants during their preparatory English course. Moreover, while this small-scale study revealed students’ perceptions of assessment change, LLS use and the role of English-speaking environment, further work needs to be undertaken, both at a research level and within the classroom, to ascertain the transferability of CET-type examination skills and strategies to other types of learning, and to seek to promote such transferability. In further work, it would be particularly useful to use reflective accounts and metaphor analysis, as Ellis (2008) has done, in order to explore the complex, often subconscious link between learners’ beliefs and LLS. Finally, any analysis of Chinese students’ LLS is merely a starting point in this process of discovery; the study may be extended to other first language groups, and such a comparative analysis may reveal further, noteworthy findings.

References


Hu, G. (2002). Potential cultural resistance to pedagogical imports: The case of


and ACER, Hong Kong.


**Appendix A**

**Interview questions**

来英前 (Before you came to the UK)

1. 在大学课堂，你是如何学习英语的？
   In college English language classroom, how did you learn English?
2. 在大学阶段，课后你是如何学习英语的？
   During your tertiary education, how did you learn English after class?
3. 在大学阶段，你的英语通过什么方式进行评估？
   During your tertiary education, how were you assessed in English course?
4. 你需要为这种评估做什么样的准备？
   What did you need to do in order to prepare for such assessment?
5. 你报名参加过补习班来提高英语吗？如果有，请说明。
   Have you ever attended any English language course to improve your English? If yes, please explain
6. 为了申请英国的大学，你参加了何种考试？
   In order to study in the UK, what type of exams have you taken?
   你是如何为这种考试作准备的？
   How did you prepare for such exams?

来英后 (after arrival in the UK)

• 课前语言班 (preparatory course)

1. 来英后，在语言班你是如何学习英语的？
   After you arrived in the UK, how did you learn English in preparatory English course?
   你在上完语言班后需要考 WELT 吗？如果是，你为此作了哪些准备？如果不是,
对你的英语学习有何影响？
Did you need to take WELT after your preparatory course? If yes, what kind of preparation did you need to make? If not, did it have any influences on your English language learning?

• 学期中 (during Master course)

3. 你在专业课开始后是如何学习英语的？
How did you learn English since you started your Master course?

4. 你上过学期中的语言班吗？如果是，对你的英语学习有哪些影响？
Have you ever attended in-sessional English language support classes in term time? If yes, what kind of influences did it have on your English language learning?

上完学期中的语言班，你需要考 WELT 吗？如果是，你为此作了哪些准备？
Did you need to take WELT after in-sessional English language course? If yes, what kind of preparation did you need to make?

结束 (On completion of the Master course)

对于像你这样背景，想来英求学的中国学生，你对他们的英语学习有何建议？
For people like you who wanted to study in the UK, do you have any advice to give for their English language learning?

Appendix B
Teaching guide for preparatory English language course

1. Text based studies
If you are a TBS tutor, this course will be the mainstay of your teaching. The Text-based Studies course is designed to improve students’ skills in reading and writing in English. Text-based Studies also includes language study (e.g. grammar and academic vocabulary).

2. Writing
The exact content of each lesson is for you to decide with the students, but you can probably expect the classes to contain input in the following areas: Note-taking and summarising skills Academic style Using references and quotations Presenting the work of others Writing introductions and conclusions Effective paragraphing Patterns of essay organisation Coherence and cohesion Presenting and commenting on visual material Patterns of vocabulary/grammar structures specific to certain academic fields.
3. Reading
Focus on tackling longer reading texts – academic articles and research reports, chapters from books, for example. The skills involved will be similar to those used in Phase 1 – skimming and scanning, identifying structure, and taking notes – but in dealing with longer texts students will need to improve their ability to identify what is important and necessary to read, and what can probably be skipped or skimmed.

4. Listening and Speaking
Listening and speaking is academic in content and demanding for the students. Because your group will be more subject-specific, you will be able to make use of clips of videos of university lectures for listening and note-taking activities, and to set up seminar-style discussions around topics of academic interest to your students. Students are expected to understand and take notes on longer stretches of monologue in class, and at least one lesson per week should be used to anticipate and/or review the weekly lecture by a visiting speaker.

Appendix C
Sample degree transcript
Name: Interviewee 6
Major: International Economic Law
International Intellectual Property & Policy: Essay 63
Legal Aspect of International Investment & Transnational Corporations. Exam 70
International Economic Law: 1st Assessment. Exam 66
International Economic Law: 2nd Assessment. Exam 62
Legal Research & Writing Skills. P
Dissertation: 68
International Business Transactions & Law: Exam. 61
Legal Aspects of International Trade & The World Trade Organization: Exam 68

Appendix D
Assessment criteria
Preparatory English language course
Assessment
A written project of 1500-2000 words based on research (preferably) around the student’s introductory reading list. An oral presentation of between 7 and 10 minutes (including time for questions), based on research into a subject of special academic
interest: this may cover the same area as the written project

**Master programs.**

**MA**

80+(Distinction):
Work which, over and above possessing all the qualities of the 70-79 mark range, indicates a fruitful new approach to the material studied, represents an advance in scholarship or is judged by the examiners to be of a standard publishable in a peer-reviewed publication.

70-79 (Distinction):
Methodologically sophisticated, intelligently argued, with some evidence of genuine originality in analysis or approach. Impressive command of the critical/historiographical/theoretical field, and an ability to situate the topic within it, and to modify or challenge received interpretations where appropriate. Excellent deployment of a substantial body of primary material/texts to advance the argument. Well structured, very well written, with proper referencing and extensive bibliography.

60-69:
Well organised and effectively argued, analytical in approach, showing a sound grasp of the critical/historiographical/theoretical field. Demonstrates an ability to draw upon a fairly substantial body of primary material, and to relate this in an illuminating way to the issues under discussion. Generally well written, with a clear sequence of arguments, and satisfactory referencing and bibliography.

50-59:
A lower level of attainment than work marked in the range 60-69, but demonstrating some awareness of the general critical/historiographical/ theoretical field. Mainly analytical, rather than descriptive or narrative in approach. An overall grasp of the subject matter, with, perhaps, a few areas of confusion or gaps in factual or conceptual understanding of the material. Demonstrates an ability to draw upon a reasonable range of primary material, and relate it accurately to the issues under discussion. Clearly written, with adequate referencing and bibliography.
40-49 (Fail/Diploma):
This work is inadequate for an MA award, but may be acceptable for a Postgraduate Diploma [although some departments may wish to set the pass mark for a diploma at a level higher than this]. Significant elements of confusion in the framing and execution of the response to the question. Simple, coherent and solid answers, but mainly descriptive or narrative in approach. Relevant, but not extensive deployment of primary material in relation to the issues under discussion. Occasional tendency to derivativeness either by paraphrase or direct quotation of secondary sources. Some attempt to meet requirements for referencing and bibliography.

39- (Fail):
Work inadequate for an MA or Diploma award. Poorly argued, written and presented. Conceptual confusion throughout, and demonstrates no knowledge of the critical/historiographical/theoretical field. Failure to address the issues raised by the question, derivative, very insubstantial or very poor or limited deployment of primary material.

MSc
Mark Range
80% and over (High Distinction) Work which, over and above possessing the qualities of the 70-79% descriptor, demonstrates excellence – the nature of which will vary according to the assignment but may include: comprehensive answers, complete and correct proofs or calculations, project work that extends the original brief, deep and critical analysis, originality, and advance in scholarship, a highly professional approach.

70%-79% (Distinction) The work demonstrates mastery of the subject matter, methodologies, and, where appropriate, laboratory techniques. It also provides evidence of near complete conceptual understanding, high level technical competence, and depth of analysis or mathematical understanding. Where applicable, the statement and proof of theorems is handled with confidence, and their application to unseen material is sound. Accuracy and precision will be strong throughout and, if applicable, presentation will be excellent. Minor mistakes may nevertheless appear occasionally. Where appropriate, the work shows evidence of originality.
60%-69%  (MSc Pass) The work demonstrates a sound and thorough grasp of subject matter and methodologies. Conceptual or mathematical understanding and technical competence are solid, but applications, arguments, or data analysis may contain minor flaws. Examined work will be well organised and structured, while good presentation and a logical approach to the material will be evident in projects or dissertations. Overall, the work reveals a high level of effort and commitment, but lacks breadth, depth, and fluency in parts.

50%-59%  (MSc Pass) The work reveals an underlying grasp of the subject matter, but with areas of confusion or some gaps in conceptual/mathematical understanding or methodology. Answers are fairly well structured but may tend towards the factual or derivative. In project or dissertation work, general conclusions or outcomes are reasonable, but there is room for substantial improvement in the individual’s ability to apply theorems, analyse problems or execute technical skills.

40-49%  (Inadequate for an MSc, but may be acceptable for a Postgraduate Certificate.) Though it reveals some familiarity with the subject matter, and a basic grasp of factual and conceptual material, there are frequent and important gaps and/or misconceptions. Some effort has been made to reflect on and analyse questions or problems, or to apply theorems, but with little evidence of organisation or insight. Technical competence is poorly developed and general conclusions are unreliable or unsubstantiated.

20%-39%  (Fail) The work is insufficient to demonstrate a basic grasp either of factual or conceptual subject matter. Technical competence is at a very low level and, if appropriate, laboratory work has required constant supervision. Data used in project work may be both inaccurate and irrelevant. Overall, answers and arguments reveal little effort towards analysis or conceptualisation. Important issues may have been ignored or seriously misconstrued. There is little evidence of an individual contribution to the material.

Less than 20%  (Fail) Inadequate work: poorly argued, written and presented; conceptual confusion throughout; demonstrates little or no knowledge of the field. Failure to address the issues raised by the question. Project work contains little or no data. Sparse or no evidence for technical competence or individual contributions.
Second Language Development through Technology Mediated Strategic Interaction

Neil H. Johnson  
*University of Aizu, Japan*

Jonathan deHaan  
*University of Shizuoka, Japan*

**Bio Data**
Neil H. Johnson (Ph. D. University of Arizona) is an Assistant Professor in the Center for Language Research at the University of Aizu, Japan. He teaches a variety of English for Specific Purposes courses to undergraduate students who are specializing in computer science. His main research interests are in sociocultural theory and discourse analysis.

Jonathan deHaan (Ph. D. New York University) is an Associate Professor in the Faculty of International Relations at the University of Shizuoka, Japan. His main teaching and research interests are in the areas of educational games, game literacy, and English for Specific Purposes simulations. More information is available at: http://langcom.u-shizuoka-ken.ac.jp/dehaan

**Abstract**
Teaching language proficiency can be particularly problematic in a Japanese university context because of issues with low motivation (Yashima, 2002; Oda, 1993), anxiety and shyness (Kitano, 2001), and practical difficulties associated with monitoring performance and providing effective feedback to large numbers of students. Strategic interaction (SI), as proposed by Di Pietro (1987), uses the scenario as an organizing principle for classroom practice. This involves learners being given different parts or roles in a situation to be resolved through language in unfolding interaction. In this paper, we explore and detail the design of an approach to SI that is mediated by use of an online wiki space and digital video technologies. Participants at a Japanese university engaged in an SI routine within the context of learning politeness strategies for a Business English course. Analysis of performance transcripts using a functional language framework, data from a post-performance discourse completion task, and learner reflections, confirm the potential that technology mediated SI holds for increasing language proficiency in this context. We argue that the data shows evidence of a shift from object-regulation towards increased self-regulation, in the genesis of language development.

**Key words:** Mediation; Strategic interaction; Technology; Sociocultural; Wiki
Introduction
The constraints to teaching English language proficiency effectively in a Japanese university context have been well documented (Oda, 1993; LoCastro 1996), and cultural issues have been used to explain low learner motivation (Dwyer and Heller, 1996; Yashima, 2002), and issues with anxiety and lack of confidence (Kitano, 2001). There may be more procedural difficulties associated with monitoring performance and providing effective feedback to typically large class sizes. Engaging learners in meaningful communication to push development may also be difficult in an EFL monolingual setting, where English is generally not used outside of the classroom. In this preliminary study, we revisit strategic interaction (SI) as an organizing principle for classroom activity to implement effective communication instruction in this context. Interest in SI for language development, based on Di Pietro’s (1987) seminal work, was well documented in the 1990’s (Alatis, 1993), but since then, has fallen from the research and teaching consciousness of the field, witnessed by the lack of recent research studies into application of Di Pietro’s (1987) ideas. Developments in web and video technologies, and the relatively accessible nature of these and related tools, mean that SI, as a dynamic and effective organizing classroom principle (Brown, 1993), can be adapted for use with these technologies, and potentially provide an effectual pedagogical sequence for developing spoken language proficiency.

Strategic interaction
The basic premise of the SI framework is built upon the dramatic tension of completing a scenario in which each participant plays a different and often conflicting role. At the outset, each participant has only a part of the information necessary to complete the given interaction. Dramatic tension arises through the uncertainty of the role that each participant is playing, as well as the exact nature of the situation. A key part to SI in language development is the creation of this tension that captures the real life uncertainty and drama of communicating in a target language. The urgency of having to communicate something in real time, even beyond present levels of development, to solve an emergent situation pushes learner production to its limit and therefore creates optimal conditions for memorizing language, and development of language and communicative proficiency. The suggested routine of activity to enact a scenario, as detailed in the original text (p. 2), is as follows:
Pre-class Preparation:
In this initial stage, a scenario is selected and role cards created for each of the participants. These cards outline the basic premise of the interaction and suggest a goal and desired outcome for that learner.

Phase 1: (Rehearsal)
In rehearsal, students form groups and prepare “agendas” to fulfil the requirements that have been suggested for their roles in the scenario. The instructor guides the learners wherever necessary.

Phase 2: (Performance)
Students perform their roles while the remainder of the class observe.

Phase 3: (Debriefing)
Teacher leads the entire class in a discussion of the student performance.

To fully appreciate the proposal made by Di Pietro (1987), it is necessary to reflect upon his understanding of language. He proposed (p. 6) three important dimensions to language to consider when implementing a pedagogic intervention in a second language:

1. Information exchange (focus on grammatical orientation)
2. Transaction (negotiation and expression of speaker intentions)
3. Interaction (how language works to portray roles and speaker identities)

Di Pietro (1987) points out that in many communicatively oriented language classrooms, the focus is exclusively on information exchange with learners exchanging linguistic tokens according to given criteria. Such exchanges may illicit target forms, but because they are removed from the reality of life outside of the classroom, the tokens are not animated by genuine speaker intention. There is no real communicative act at stake. Building a notion of transaction into classroom interaction, gives notice of the fact that language is always used for some communicative purpose to achieve something, as well recognizing the inherently ambiguous nature of language use in social interaction. Understanding and negotiating an interlocutor’s intentions and different levels of ambiguity, across all three dimensions of language, are important to successful communication. Awareness of the social roles and identities that speakers are always negotiating, similarly shapes the linguistic choices made by speakers in any given context. Adding these dimensions to the linguistic or grammatical aspects of language use shape the design of the SI proposal and places Di Pietro’s (1987) work firmly within a sociocultural
and Hallidayan understanding of language use and language development. For example, Byrnes (2006) writes in summary of the Hallidayan approach: “The aim is a competent level of literacy on the part of the learners that crucially involves awareness of the meaning making consequences of different linguistic resources, at all levels of language” (p.5). This is precisely the goal of the approach outlined by Di Pietro (1987), where learners develop and exercise selective control over the meaning making resources of the target language to achieve the inherent aims and goals of their assignment.

**Sociocultural theory and language development**

The focus on *interaction* in this model, and exploration of “the interrelatedness of the social and cognitive aspects of interactive discourse” (de Guerrero and Villamil, 1994. p 484) is based upon the developmental principles of sociocultural theory (SCT). Two related principles are key to understanding this approach: mediation and internalization (Lantolf and Thorne, 2005). The unique point about the human mind is that it is essentially mediated by semiotic, culturally constructed artifacts. As Wertsch (1994) explains: “[Mediation] is the key…to understanding how human mental functioning is tied to cultural, institutional, and historical settings since these settings shape and provide the cultural tools that are mastered by individuals to form this functioning” (p. 204). Semiotic means of development, i.e. *mediational tools* such as concepts, symbols, schemata, etc. are created and exist firstly on the social plane and become appropriated and integrated into the individual’s own cognitive activity through participation in the culturally organized activities of a society.

In language education, dynamic assessment procedures (Poehner, 2007) and related work on teaching in the zone of proximal development (Aljaafreh and Lantolf, 1994) have made use of implementing changes in mediation, in what is known as the graduated prompt approach (Poehner, 2007). This approach demonstrates the effectiveness of providing different levels of mediatory feedback, dependent on performance level, for learners in instructional and assessment processes. In second language learning research, peer collaboration and collaborative dialogue have been shown to be an effective means of mediating language development (Donato, 1994; Swain and Lapkin, 1998). Here, learners co-create interpsychological activity through dialogue on problem solving tasks that allows for development to take place. In Swain and Lapkin (2008), the language of the learners themselves becomes an artifact for talk about their own language. This allowed learners to notice their own
developmental level and thereby the artifacts served as a mediating tool for further learning.

The convergence of thought with mediational artifacts is known as internalization. This is a dynamic process whereby the learner develops increasing degrees of control, or regulation, over the semiotic psychological tool, and thereby transforms internal cognitive processes (Frawley and Lantolf, 1985; McCaffertey, 1994). In educational contexts, development can be seen in terms of changes in regulation as the learner engages in problem solving activity. Initially, the learner will be dependent on external mediatory means, such as a peer or instructor, and at such a time, can be said to be other regulated (Luria, 1982 p. 91). Reliance on an artifact, such as a diagram or written procedure, for example, is said to be object regulation. In De Guerrero and Villamil’s (1994) study, learners were seen functioning on a task through other regulation. As they engaged in peer revision of writing, the participants clearly did not have control over the linguistic resources necessary to complete the task, and were controlled or regulated by the working draft that they had in front of them. Through engagement in a combination of expert intervention, verbal interaction, and scaffolded participation in appropriately organized learning activity, an individual may effectively gain increasing control over the mediatory means, and will eventually become self-regulating. Frawley (1987) explains further: “language serves to regulate the self…self-regulation is in fact the highest and most critical function of speech” (p. 159). That is, through speech people gain voluntary control over their own mental activity (Vygotsky, 1987).

Data analysis

Self, other, object regulation

The regulatory function of language then means that the goal of research within a Vygotskian framework becomes functional analysis of learner language use in different kinds of speech activity. As Ahmed (1994) explains:

“The task of the researcher is to discover if the speech of the interlocutors shows evidence of object regulation, other regulation or self regulation. This is achieved through careful analysis of the relationship between task factors and specific linguistic forms manifested in the speech of the interlocuters” (p. 160)
This functional perspective of language use provides the theoretical frame for data collection and analysis in this study. To analyze the learner language use, we used two complimentary frameworks to assess the regulatory function of the learner as they use language in the context of the SI framework. We are concerned with the ability of the learner to gain control the target features of the language and discourse, leading to independent and creative use. Tracking longitudinal change, potentially from object and other regulation, towards more self-regulation, we compared language use in a repeated SI as learners progressed through the semester and interacted with the teaching-learning process engendered by the classroom design. The second feature of learner discourse we attended to is the response to direct mediation from the instructor in the form of feedback. The response to feedback, in terms of use in the second performance and subsequent discourse completion task, allows insight into the ongoing developmental level of the learners. Shifts in the requirement for mediation are taken as evidence of language development. Aljaafreh and Lantolf (1994; p. 470) provide five general levels of development through which the novice learner will typically move as changes in regulation take place, from object regulated (e.g. using a given dialogue) intermental (i.e. collaborative task completion) to intramental (i.e. self-regulatory or independent) functioning. The five levels are as follows:

1. The learner cannot notice or correct an error even with assistance.
2. The learner can notice the error, but cannot correct it even with intervention
3. The learner is able to notice and correct the error, but only under other regulation.
4. The learner notices and corrects an error with minimal, or no obvious feedback
5. The learner becomes more consistent in using the target structures correctly in all contexts. Noticing and correcting errors means the individual is fully integrated.

At Level 5, the learner is able to use the target structure consistently and correctly in all contexts and can be said to be self-regulating. From a pedagogical perspective, the key to implementing optimally effective instruction and learning is organizing appropriate classroom sequences that provide effective learning affordances (van Lier, 2000) or instances for different kinds of mediation to take place so that assistance can be targeted appropriately to lead development. Here, an affordance is understood as a “learning opportunity that can be used by an active and engaged learner to take action over his/her language” (p.135).
Microgenetic analysis

Microgenetic analysis focuses on the moment-to-moment uses of language in social interaction and reveals moments where interpsychological assistance can lead to awareness and shifts in regulation (de Guerrero and Villamil, 1994; Belz and Vyatkina, 2005; Guiterrez, 2008). This method allows the tracing of the history (i.e., genesis) of developmental processes in and through social interaction, in order to understand "how the human mind functions as a consequence of its formation in cultural activity" (Lantolf & Thorne, 2006, p. 57). As Vygotsky states: “we need to concentrate not on the product of development but on the very process by which higher forms are established...to study something historically means to study it in the process of change” (1978, p. 64-65). In order to understand the developmental effect of participation in this classroom activity design, we focus on a dyad of learners through microgenetic analysis of initial and repeat performances in a language scenario, and on a subsequent discourse completion task (Cohen, 1996).

The goal of the functional microgenetic analysis is insight into how the affordances provided by the pedagogical procedure are taken up by the learners. “It becomes important to determine if the locus of control resides in one interlocuter, is distributed between the two or lies in the external context of the task itself” (Ahmed; 1994 p. 160). The procedure for analyzing the data involved taking the transcripts of each performance on the online wiki and firstly checking for accuracy against the actual recording. The language data for each participant was then organized into units of meaning, using a conversational analysis framework (Heritage, 1989) in which conversation is analyzed as discreet categories of meaning, such as a greeting, request-response, or closing of conversation. These units of meaning were then traced longitudinally to assess changes over time in the language used to express them. Each unit of meaning was analyzed firstly in terms of the given context of the task itself, including the task prompt and the model dialogue the learners had studied. Secondly, the correction from each unit of meaning was analyzed in terms of the scale produced by Aljifraah and Lantolf (1994). Finally, the second transcript was analyzed in terms of changes from the first performance, and the wiki transcription data allows for direct comparison. Changes in performance for each unit were noted and related to changes in regulation from either the participants. Finally, the units of text were related to the final discourse completion task, which was similarly broken up into discreet conversational units of meaning, using a conversational analysis approach (Heritage, 1989). The analysis protocol is shown in Table 1.
Table 1: Analysis protocol for conversation units

<table>
<thead>
<tr>
<th>Category</th>
<th>Unit of Talk</th>
<th>Transcript 1</th>
<th>Student Corrections</th>
<th>Transcript 2</th>
<th>Discourse Completion Task</th>
</tr>
</thead>
<tbody>
<tr>
<td>Language Use</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Analysis of Learner</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Regulation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Analysis of Developmental Level</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The protocol for analysis allowed the researchers to trace development of target forms, such as modal verb use for politeness strategies, across the semester.

**Designing a web-based strategic interaction**

We now describe the ways in which this framework was implemented in a Japanese university setting. Our goal has been to maintain the integrity of the instructional proposal made by Di Pietro (1987) yet also adapt the procedures to the learning affordances offered by web 2.0 and digital technology. Warschauer and Grimes (2007) define the key innovation of web 2.0 technologies as “interactive participation” (p. 2). Relative ease of access and use means that networking sites, such as a wiki, can promote rapid and collaborative production and publication of digital material that can become central to learning activity. In the cycle outlined below, participants transferred their videos directly into Apple iMovie software, for brief digital editing and then uploading into the wiki space. This takes from between one and three minutes and then students have access to a recording of their own performance. In class the learners were instructed, firstly, to transcribe their own conversation exactly as it had taken place and this text then appears below the video screen as shown in figure 1.

---

1 This can be done in QuickTime for PC users
The instructor created a free and private (the students shared one login and password) www.pbworks.com wiki site for the project. A PBworks wiki was selected because its pages can contain text and embedded video, and there is no storage limit for video. This functionality means that students and the instructor can view and edit performances, transcriptions, and error corrections. The “2.0 version” of the SI procedure developed in this research, can be therefore be summarized as follows:

Pre-class Preparation:

Before the scenario, classroom work details target structures and outlines grammatical and pragmatic issues of use. Learners become familiar with the technology. A scenario is selected and role cards created by the instructor for each of the participants. These cards outline the basic premise of the interaction and suggest a goal and desired outcome for that learner. Basic vocabulary and suitable phrases for the task are provided. A model dialogue is prepared by the instructor and studied in class and for homework.

Phase 1: (Rehearsal)

Students form groups and prepare “agendas” to fulfil the requirements that have been suggested for their roles in the scenario. The instructor guides the learners wherever necessary.

Phase 2: (Performance)
Students perform their roles. The interactions are recorded on digital video and the performance is then uploaded into the class wiki. Learners transcribe their performance and copy and paste their scripts into a wiki page directly below their videos. Other students can review online.

Phase 3: (Debriefing)
Teacher works in discussion with each of the groups, asking questions and guiding the learners in reflection on both positive and weak aspects of the interaction. Learners then work together to re-work their script into a possible ideal performance. Errors in grammar or pragmatics should be noticed and corrected.

Phase 4 (Second Performance)
The learners either act out a second performance with an extension made to the scenario to maintain dramatic tension or, if the first performance is not deemed successful, they may re-enact their first performance in light of feedback. The second performance is recorded and transcribed for comparison with the first. Learners reflect on the two performances.

There are a number of changes here from the original suggestion in Di Pietro’s (1987) work. Firstly, in phase one, we provide a model dialogue to help learners with lower proficiency gain some concrete understanding of the kinds of interaction that might be expected. This should not be too close to the actual scenario, otherwise the dramatic tension will be lost. In phase two, the learners perform out of the view of the rest of the class to allow for better recording. The performance is later available for all students to review and comment on. The debriefing is more personal in this design, with the instructor giving feedback to each pair or group in turn. Important points can be shared with the whole class by playing the recording back to the whole class, in keeping with the original proposal. The use of a second performance was seen as an important way to push or extend learner development, in keeping with Di Pietro’s (1987) ideas. However, the option of a repeat performance, without extending the scenario, was built into the design to help students with lower proficiency, who may initially struggle to complete the task. If the first performance was not successful i.e. the participants simply could not solve the problem built into the scenario, it was decided that the learners should have the opportunity to redo this performance in light of instructor feedback.

In this pedagogical design, the second performance, strictly speaking, will not be a strategic interaction, since the roles and outcomes of each party will now be known.
Nevertheless, the learners are still required to perform the scenario in real time with a camera running, and without external aids, other than the scenario calendar for their respective role, reproduced from the role cards (see appendix 2) and visible in both performances. This artifact lessens the burden on the memory of the participants and facilitates more concentration on language production. It was important to allow these generally low speaking proficiency learners the option to repeat their performance, so that if they had struggled with the scenario, they could process and act upon feedback from the debriefing stage. Another teaching possibility, not followed in this case, would be to allow students to switch partners for the second performance. This would maintain a degree of the dramatic tension. It should be noted that numerous attempts to engage the learners in discussion activities and tasks in the target language had failed to produce satisfactory turns at target language talk in the previous weeks of class. This fact impacted on the classroom design and allows us to suggest that these learners, in common with previous cohorts for the class, had a generally low speaking proficiency. The wiki-space allows for direct comparison of performances, which may prove developmentally important, as well as motivational to learners in this context. The use of learner transcripts as a mediational tool to foster learner awareness and development is suggested by Swain and Lapkin (2008). The research question guiding this study then is as follows: Do the affordances provided to the learners in the strategic interaction sequence of study lead to development of target language knowledge?

Study design and implementation

The setting for this study is a computer science university in Japan where the learners typically develop upper-intermediate levels of literacy in specialized aspects of technical writing, for example, but often struggle with basic levels of spoken communication. The participants were 18 students taking an elective course in the third year of their program. The title of the course was Speaking and Writing for Business English, and learners explored the discourse of various business related texts while assuming the roles of Japanese staff at a major international company. A feature of this course in previous semesters has been a difficulty in engaging the learners in communicative tasks, and their lack of willingness and/or ability to take part in classroom discussions, or work that required oral participation. The spoken element of the course was designed to provide some degree of training for these learners so that they may begin to develop the requisite language skills necessary to function...
effectively in a workplace environment where English is used. The implementation of SI is therefore held as a possible pedagogical framework for improving and developing speaking and interactional skills. The goal of this study is to analyze the outcome of the implementation of a pedagogical design that was based upon Di Pietro’s (1987) ideas, with a view to future implementation and development of the pedagogical sequence. The participants met for one 90-minute session per week over 14 weeks. The speaking module, designed around the features of SI, took place over 3 weeks with an additional discourse completion procedure in week 14. Within the broad goal of developing speaking and interactional competence, we were also interested in development of one of the main target structures for the course: the use of modal verbs for politeness strategy in a business or professional context. Hinkel (1995) has suggested that modal use is particularly important in English, as their deployment is a marker of discourse community participation and values. The following features were implemented as part of the research design.

**Reflection**

At the outset, participants completed a questionnaire and were asked to reflect upon their English language learning strengths/weaknesses. At the completion of the SI module they reflected back on their experiences with the strategic interaction and other elements of the course.

**Discourse completion activity**

In week 14 (four weeks after completion of the SI), the learners were given a follow up writing task (see Figure 2), designed to elicit and thereby assess control over target structures (politeness requests and refusals using modal verbs) and vocabulary related to scheduling and appointments. This task was familiar to the learners and was given under timed, examination conditions. Though this was not a spoken task, the time pressure (20 minutes) of the exercise in examination conditions, meant that learners were producing a near spontaneous linguistic performance. Cohen (1996) has also used written discourse completion tasks as an effective means of evaluating pragmatic development. Negueruela (2003) argues that spontaneous production, either spoken or written, can provide insight into learner development. The criteria for assessment of this particular task were accuracy and appropriateness of language use, and achievement of a solution to the problem written into the task.
Figure 2: Discourse completion activity

Business Meeting Scenario (20 minutes)

No dictionaries or other materials may be used

Two business people, Mr. Tanaka and Mr. Jones, are having a short meeting to discuss a problem with a project. They are meeting in Tanaka’s office. Tanaka needs Smith to rewrite the introduction of the report he has submitted. The report has some mistakes in grammar and content. Smith is leaving the country for a business meeting tomorrow so will need more time. The deadline is tomorrow but Tanaka can give him one more week.

The conversation is started for you – try to continue it. Try to use politeness strategies for requests and refusals where appropriate:

Tanaka: Thanks for coming to my office, Mr. Jones.

Jones: You are welcome. You wanted to talk about the report?

Focus subjects

For the purposes of this paper, we focus on two participants. Shin and Jun (names have been changed), both male undergraduate students in their third year of the English program, initially described their English level as “low” and characterized by “lacking in confidence to speak English.” Their performance is selected for analysis because out of the nine participating pairings, they were the only pair that did not initially rely upon hand written notes to complete the task. While the use of notes had been forbidden in the class prior to the performances, all but one pairing felt that this was necessary for them to take part in the recording session. This is disappointing from a teaching perspective and confirms that the level of the task was perhaps too high for this group of learners, especially as they were unfamiliar with the SI procedure. Interestingly, all of the pairs managed to complete the task without notes in the case of the second performance, suggesting that even using handwritten notes may have had some mediational benefit. Using the functional perspective of language use, consistent with a sociocultural theoretical framework, we trace microgenetic and visual analysis of learner performance throughout the SI cycle, and on to a discourse completion task, to better understand the developmental impact of the instructional sequence, through analysis of changes in regulation. The recordings took place in an office away from the main body of students and typically took between 2 and 3 minutes to complete. The pairs were announced immediately before the recording, so
that learners from the two different groups could not rehearse their different parts. The two transcriptions of the SI performances were checked against the digital recordings, to ensure accuracy.

**Results**
The focus participants produced the following transcription after their first performance. The text in bold is the correction work done by the learners themselves, as it appeared on their wiki page, after consultation with the instructor and discussion between themselves. The interaction between instructor and learners involved two kinds of feedback. The first was an overall commentary on the scenario and how the learners had solved the situation. In this case, it was pointed out that essentially the pair had not solved the problem and through deferring to the supervisor, they had only really postponed the solution to another time. The pair was asked to consider other ways that the scenario could unfold and the scheduling issue be resolved. The learners agreed at this point that it had been difficult to solve the question and they requested the opportunity to try the same scenario again. The second type of feedback involved going through the transcript and pointing out grammatical and pragmatic issues. Correct or alternative versions were not given to the learners, however. So in line one for example, it was pointed out that in “I Jun Hideki” that there was a helping verb missing from this introduction. The participants each took notes as the instructor gave feedback and the pair was then given time to find alternative forms.

*Transcription 1: Week 9 of Semester*


*I'm Jun Hideki, I'm head of marketing in Tokyo.*

2. Shin: Nice to meet you too. I'm Michael Smith. I'm working Yahoo from…from Yahoo at Osaka.

*I'm from Yahoo in Osaka.*

3. Jun: Welcome to our office Mr. Smith. Ah…er, Let’s start meeting.

*Let's start the meeting*

4. Shin: There... There is a problem... There is a problem are you sending ah... a brochure brochure.

*There is a problem with the brochure that you sent.*

5. Jun: Really?
6. Shun: Yes. There is mistake spelling and grammar, grammar mistake ah...
also there is there is mistake a page 3.

There are some mistakes with spelling and grammar, also page 3 is missing.

7. Jun: I'm sorry, ah…I'm sorry.

I'm sorry.

8. Shun: Erm…I wondered if you if you repair it... for you to ah... to repair this week?

if you could correct it this week?

9. Jun: Oh I'm sorry, I have to…go to…to America to…tomorrow…tomorrow morning.

I'm sorry, I have to attend a conference in San Francisco, so I’ll leave tomorrow.

10. Jun: So, It's going to be difficult, get going to be difficult to get that ah…finish…finish it. When is the deadline?

So it is going to be difficult to get that finished this week.

11. Shin: Our deadline is Friday.

12. Jun:…Could I…could I ask you for more time?

Could I ask you for more time on this?

13. Shin: No. Ah... it is dif… difficult. But I...I talk to ah...my supervisor.

It is difficult. But I will talk to my supervisor.

14. Jun: Thank you ah…ah I'm sorry I would appreciat…I appre…I would appreciate more time.

OK thanks, I'm really sorry for this problem and I would really appreciate more time.

15. Shun: I understand such for ah... such kind attention to meet meet me.

Thank you.

I understand. Thanks for coming in today, thank you.


Performance 1 – Discussion

Transcription 1 suggests that even though the performance was rehearsed, the learners still struggled with the task. A marked feature of the participants’ turns at talk (e.g. lines 2, 4, 8, 9, 10) is the lack of fluency, evidenced by pauses, false starts, and repetitions as they search for the appropriate lexical and grammatical items to achieve
the joint communicative goal. This is suggestive of learners operating at an interactive and proficiency level just beyond their developmental level. This was perhaps a result of the task design being set at a level beyond present capability, as evidenced by the failure of other groups to complete the task without notes, though in the rehearsal stage it was clear that the scenario and required language were generally not beyond comprehension. There were, however, numerous errors in word choice (e.g. repair instead of correct in line 8), grammar (e.g. missing auxiliary in line 1, problem are you sending in line 3) and there is also the struggle to maintain an appropriate rhetorical tone for this task in the assumed social context. There is an appropriate use of a modal verb on one occasion (line 12), the target of the class instruction, to create a polite question form. Then again, we also see politeness as problematic in both line 10 where the use of no is a little abrupt and also the use, in line 14, of such kind attention to meet me, which is rhetorically inappropriate, showing too much deference, for a business-meeting context. The transcriptions of the interactions accurately reflect the stuttering and unsure nature of much of the performance.

An unexpected outcome of this routine, and clearly evidenced in the other dialogues in the class also, is the use by the learners of the structure and composition of the model dialogue that was studied in the preparation stage (see appendix 1). It is clear from the organization, turn taking, as well as vocabulary and grammatical structures employed, that the model has been studied, memorized and is being utilized by the learners. For example, in the model, the following turn at talk is given as a way to negotiate the deadline on a project:

**Jones:** When is the deadline? Is this urgent?

**Yamamoto:** Actually, our deadline for the printing of the new brochure is next Wednesday.

**Jones:** Could I ask you for more time on this? I might need two more weeks to get the furniture information to you.

**Yamamoto:** Well, it’s going to be difficult, but I will talk to my supervisor and I’ll let you know.

This is clearly the strategy used in the strategic interaction by the two participants, who yet, in preparation, did not know the precise nature of the scenario that they were going to take part in, since they each only had access to half of the situation. This point was confirmed in the review stage after the second performance. The utilization of the supervisor was not a part of the scenario that had been given to the learners as a possible outcome, yet having studied this model; they were able to use this aspect of
the given model as a resource to resolve the communicative problem that they faced. In the debriefing stage, the learners acknowledged the fact that without this assistance, the task would have been too difficult for them.

This use of the model was not foreseen at the outset of the teaching cycle, yet in retrospect, it might have been expected, since the response by the learner to assistance has been shown in dynamic assessment to reveal developmental level. This is the case in this instance, reflecting a task beyond present development. It is, however, an interesting example of learners exercising their agency (van Lier, 2008) in adopting and using this artifact of instruction for the communicative task, indeed, the learners actually change the scenario to suit their own capabilities, and this object mediation allows these learners to operate, albeit with difficulty, at levels beyond their current developmental level. Recalling the Aljaafreh and Lantolf (1994) framework, there is evidence in the data of examples where the learners have been able to correct errors with feedback from their instructor, such as the grammatical and pragmatic points that the learners offer as appropriate corrections. This suggests operation at developmental Level 3 for grammar and pragmatic elements, where feedback can be taken and used to improve performance. This is similar in nature to the object regulation found in De Guerrero and Villain’s (1994) study, mentioned above.

The corrections in bold suggest that both the discussion with the instructor and time spent jointly focussed on the performance data, in transcribing and self-analysis, has created sufficient awareness to allow the learners to focus more effectively on the linguistic errors, rather than on the more pragmatic or interactional issues that we see (e.g. the rather abrupt statement in line 11). Indeed, the learners were unable to take up the feedback they received about this aspect of their performance and yet were able to work on the more grammatical features of the interaction. This was either because they did not fully understand the feedback that they were given by the instructor, or it was simply too difficult to deal with all of the suggestions given. Recall that at levels 1 and 2, learners are unable to correct a performance even with appropriate mediation. The re-writes that do occur imply that they were operating at between levels 2 and 3 on Aljaafreh and Lantolf’s (1994) scale, i.e. able to notice and correct some of the errors with assistance. These corrections include the mistakes in word choice, grammar and pragmatics. One week later, the same learners produced and transcribed the following performance:
Transcription 2: Week 10 of semester

15. Jun: Nice to meet you. I am Jun Hideki. I'm head of marketing in Tokyo.
17. Shin: Welcome to our office Mr. Smith. Let's start meeting.
19. Shin: There is a problem with brochure that you sent.
20. Jun: Yes, I sent you the brochures last week.
21. Shin: There is...ah there are problem something...ah...there are some problem...with spelling and grammar...also page 3 is missing.
22. Jun: Oh really?
23. Shin: I wondered if you could correct in this week?
24. Jun: Oh, I'm sorry, I have to attend a conference in San Francisco…I leave tomorrow, so it's going to be difficult to get that finished this week. When is the deadline?
25. Shin: Actually, our deadline is Friday.
26. Jun: Could I ask you for more time?
27. Shin: It is difficult, but I will talk to my supervisor.
28. Jun: OK thanks, I'm really sorry for this problem and I would appreciate more time.
29. Shin: Thanks for coming in today, thank you.

Performance 2 – Discussion

As mentioned above, the second performance is no longer a strategic interaction in the sense that the participants now know the outcome and understand the dual roles. Here, the learners are able to draw upon the corrections that they have worked upon collaboratively and access them in their performance. Performance 2 is markedly more fluent with fewer false starts and repetitions. This might be expected since it is essentially the same dialogue that they are enacting; however, the point is that the learners themselves have chosen to ignore feedback related to trying to solve the scenario in a different way and have chosen to do a simple repeat performance. The inability to make corrections, even with corrective feedback, places the discourse task at level 1 or 2, and therefore beyond their present level. Direct evidence for improvement is in line 24, where the turn at talk is taken from the correction that appears in transcription 1. There are also fewer errors in terms of accuracy, fluency, and
pragmatic appropriatness, especially for Shin, who produces less language than his partner, but what he does produce is basically accurate. On the three occasions in the first performance where Shin repeated or gave a false start he was able to produce fluent and accurate lines in performance 2 (lines 24, 26, 28). Evidence from the initial scenario suggests that he almost had control of the target forms, and the feedback and repetition of the performance allow him to further his control over those forms. In other words, there is evidence of learner operation functioning at developmental Level 3. There are still grammatical errors from Jun, in lines 17 (missing article) and 21 (subject-verb agreement), and still an issue with fluency in line 21, which features three false starts. Pragmatically, there is also still awkwardness within the interaction, which is clearly evidenced in line 19, with a very abrupt introduction of the main topic of the discussion, which is responded to in line 20 with a slightly awkward conversational response that might be said to be inappropriate. This awkwardness suggests that the learners still do not have full control over the pragmatic or linguistic properties of language use in this context, and are in a sense, still trying to exchange correct linguistic tokens to complete the task, rather than engage in a meaningful interaction. This might be expected to some degree, given the fact that it is a second performance and the dramatic tension has been lost. However, their performance here is still other and object-regulated; in so far as the language that is being used can be traced back to the interaction that took place following the first performance, and the learners do not appear to be exercising independent control, which would allow for a more creative and autonomous performance.

Interestingly, despite clear instructor feedback given to the whole class on different ways that they might solve the task, more within the confines set out in the task rubric, the learners chose not to try and work towards a different solution, and instead worked at performing the second performance as a more accurate and fluent version of performance 1. When asked to comment about this, these participants acknowledged that they “haven’t done as it said” but expressed a satisfaction with the improvement between performance 1 and 2 in the second debriefing session, and noted that many of the basic grammatical errors (missing auxiliary verbs for example) had been correct in the second performance. Inability to correct performance even with other regulation and feedback, is suggestive of a developmental level of 1 or 2 on the scale used in this paper. In other words, the solution to the problem was beyond the capabilities of the learners at their given level. Again, the learners exercise agency in electing not to take
up mediational means deemed by the learners to be beyond their own developmental level.

**Discourse Completion Task**

Though there was evidence in improvement in accuracy and fluency in the second performance, it was still important, since this was essentially a different kind of task, to have the learners undertake a final assessment activity to better gauge if actual development had taken place. The discourse completion activity was given in the final class of the semester as part of a timed review procedure. Learners had been asked to review the work of the semester, prior to the final class.

The focus participants in the study, within the allotted twenty-minute time frame, wrote the following dialogues:

**Shin – Assessment Scenario 1: Week 14 of Semester**

31. Tanaka: Thanks for coming to my office, Mr. Jones.
32. Jones: You are welcome. You wanted to talk about the report?
33. Tanaka: Yes, I want to ask if you could rewrite the introduction?
34. Jones: Oh, really?
35. Tanaka: Yes, there are some problems with the grammar and also the content.
36. Jones: I am sorry about this problem. When is the deadline?
37. Tanaka: I would like this to finished on tomorrow. Our deadline is tomorrow.
38. Jones: I’m sorry, it’s going to be difficult. I am going to leave for America tomorrow. Could I ask you for more time on this?
39. Tanaka: It’s difficult, I have to talk to my supervisor.
40. Jones: I would really appreciate more time.
41. Tanaka: OK. Well, maybe I can give you one more week.
42. Jun: Thanks. I’m sorry for the problem. I really appreciate.
43. Tanaka: Thanks. Goodbye.

**Jun – Assessment Scenario 2: Week 14 of Semester**

44. Tanaka: Thanks for coming to my office, Mr. Jones.
45. Jones: You are welcome. You wanted to talk about the report?
46. Tanaka: Yes, there is a problem in the introduction.
47. Jones: Oh, really? I’m sorry for your trouble.
48. Tanaka: Yes, there are problems with grammar and content.
49. Jones: I am sorry about it. When is the deadline?
50. Tanaka: Our deadline is tomorrow.
51. Jones: I am going to leave for America tomorrow. Can I ask you for some more time?
52. Tanaka: It’s difficult. I will have to ask to my supervisor.
53. Jones: I would really appreciate more time.
54. Tanaka: I understand.
56. Tanaka: You are welcome.
57: Jones: Thanks, I will be in contact next week.

Discourse completion task – Discussion

It is interesting to note in both of the discourse completion dialogues, the way in which the language forms used are appropriated from the text that came from the earlier SI work. This is clearly evident by accurate use of appropriate expression and modal verbs in making polite requests:

I want to ask if you could rewrite the introduction
…there are some problems with the grammar and also the content
Can I ask you for some more time?
I will have to ask to my supervisor.
…maybe I can give you one more week
…it’s going to be difficult. I am going to leave for America tomorrow

This suggests that the process of preparing, performing, reflecting, re-writing and re-performing the initial SI, has had some impact on learning, in the formation of memorizable chunks of language related to these target forms, that has lead to accurate use on this later task. Memorized language has been described as developmentally important, and according to Ortega (2009), the way in which a “multi-layered repertoire of creative plus memorized language” (p. 116) may be of use to the learner is of great interest in understanding the development of L2 fluency. From an SCT perspective, imitation of language use can serve as a transitional point between intermental and intramental functioning. This is not imitation as mindless copying, but as Lantolf and Thorne (2007), suggest, “it involves goal directed cognitive activity that can result in transformations of the original model” (p. 203).
The phases of collaborative activity built into the design of the SI procedure can be seen as mutually dependent in creating learning affordances. Firstly, the provision of a suitable model and scenario provided opportunity for learners to work beyond present levels of development and to collaboratively prepare for and execute their performance. The creation of an analyzable artifact in the form of recordings and transcripts provides further affordance in the debriefing stage, where instructor feedback is important as a pedagogical intervention. As we have seen, not all of the suggestions were taken up by the learners, who elected to work on their own transformed version of the scenario. It is these artifacts, however, that allow the learners to notice and focus upon the differences between their L2 production and the target language (Swain, 2000). Further reflection and comparison between performances is then provided by the repeat performance and second debriefing stage. Interestingly, in the completion task in this study, both participants are able to reproduce both sides of the task dialogue, having previously worked with their partner on each others’ role, as well as their own. Of particular note is the correct form used, by both learners, when expressing the idea about the problems with the documents, as required by both the scenario and the discourse completion exercise:

In Performance 1, Jun says:

Yes. There is mistake spelling and grammar, grammar mistake ah... also there is there is mistake a page 3.

This is subsequently corrected and posted online:

There are some mistakes with spelling and grammar, also page 3 is missing.

In Performance 2, Jun says:

There is...ah there are problem something...ah There are some problem...with spelling and grammar...also page 3 is missing.

In the discourse completion task he writes:

Yes, there are some problems in the introduction...there are problems with grammar and content.

In this sequence it possible to trace the development of the target structure, ‘there are some problems with’, until, finally in week 14, the plural form of ‘problems’ is used correctly and appropriately. The continued reliance on forms found in the earlier classroom material suggests that the learners still do not have full control over the target language, again, since creative and independent use is still not evident. We suggest that they are operating between Level 4 and Level 5 on the regulation scale, producing correct target forms, but without full self-regulation. This may be because the task itself...
asked for a similar kind of conversation, and assessment in further research might usefully offer different kinds of discourse tasks that elicit more creative and independent uses of language.

Another feature of this discourse task performance is Jun’s correct use of the definite article (Lines 45, 46 and 49), which had been problematic in both of his SI performances. This highlights the uneven nature of development from a sociocultural point of view. The source of this correct form could have been the correct usage that was used in the text to set the writing scenario (talk about the report). Dynamic assessment procedures, as described above, demonstrate that when a feature is almost under the control of a learner, an implicit suggestion can mediate an accurate performance. It is also possible that further awareness may have developed during the reflection and feedback session that followed the second performance. The article system in English is, however, notoriously difficult to use (Master, 1990) and so the performance here can be seen in terms of a gradual increase in control over that linguistic feature.

The discourse data in this study shows evidence of a developmental shift towards increasing self-regulation. Successful re-contextualizing of the language forms, seen here in the discourse completion task, represents a developmental step towards internalization, control and development of the target linguistic forms and their appropriate use. As Lantolf (2000) explains it “The convergence of thinking with culturally created mediational artifacts…occurs in the process of internalization” (p.13). In the initial interaction, the learners are struggling and relying heavily on the model dialogue that has been introduced in class (appendix 1) to try to achieve their communicative goals. Their performance here is therefore object mediated, and while the learners are interacting to some degree successfully, they do not appear to have control of the target forms or interactional resources of the language. As they moved towards the second performance, it can be said that they have shifted towards other regulation. Through a collaborative pooling of resources, they manage, through noticing weaknesses in their own performance and through selecting from the given feedback, to regulate each other’s second performance. Through the activity required to achieve this performance, the participants have gained further control over the target language suitable for the given social situation. By the time they produce the final discourse completion language, there is some evidence, through the completion of the task and increased accuracy and appropriateness of language use, that target structures
have been memorized and internalized and are now being used in a more independent way to mediate and control linguistic performance.

**Reflection on the learning activity**

The 18 participants in the study were asked to reflect on the use of technology and SI in their language learning, at the completion of the course. These were administered anonymously and two main themes were found from the written feedback. Representative comments are annotated below:

**Use of Technology in SI:**

- I enjoyed recording the role play and the technology was easy
- The technology helped me learn English
- I am good at using computers so this was good for me
- Wiki was easy to use
- This class was fun and I enjoyed using the wiki

**Improvement in spoken language:**

1. My speaking improved. I found it possible to speak by using a model.
2. I learned how to speak polite language
3. I think I come to be able to speak more smoothly than before
4. This class increased speaking time, because of the roleplay…I liked the recording
5. I feel my speaking skills have improved a little. I enjoyed the roleplay
6. I felt through the roleplay I could speak clearly and understand what to say next. I felt my skills improved
7. My speaking skills improved a little because I was able to think about using better words in English
8. Recording a performance was a good chance to improve speaking. I will never forget!
9. When we did performance a second time, I could speak more fluently than the first time.
10. My speaking did not improve, I need to practice more.
11. We need another chance to do the role play
12. It became possible to speak more naturally
13. It was unique and interesting that I had recorded a conversation
14. I got to be able to speak according to a model, so I improved
15. I want to record more conversations with different students
16. I learnt to write from a model and I learned how to choose words
17. I still can’t write sentences by myself, I write using example sentences
18. When I could understand how to use words, I felt skills improved

The suggestion that the three different performances (SI 1 and 2 and discourse task) represent different stages of regulation, and therefore of development, is to some degree supported by the comments made in reflection at the end of the task cycle. Several different students describe the effect of using a model as helpful in their learning processes and the idea that they are starting to make conscious choice (how to choose words) in language use was also mentioned by four of the participants in the study. Several students also noticed and suggested that they needed more time to develop the necessary further self-regulation. This awareness and control is the key towards advanced literacy development from a sociocultural perspective.

**Limitations of the study**

There are important aspects to this study that limit the extent to which the findings can be generalized. Firstly, we have followed only two students through a semester of work and therefore the analysis of longitudinal development within their data sets is obviously limited. In future work, we will seek to measure developmental changes in performance over much larger classroom samples. One feature of microgenetic analysis is the large volume of data that is produced, as learner speech and written interactions are noted and analyzed. This means that large scale longitudinal studies become very difficult to manage in certain contexts.

Further, there is no pre-test measurement in this study, which would have been useful to gauge the level of awareness of target forms before the classroom intervention began. This weakens the findings, though the instructor’s own understanding, based on experience with this student cohort, was that the learners did not have control over these forms at the outset. The initial performance, even after instruction, confirms this notion, and it is still possible to trace the development of acquisition of certain forms, through analysis of the production of those forms from first performance through to the discourse completion task.

The use of discourse completion tasks as a measure of language development is also potentially problematic. Their use has been widely employed in the teaching and
assessment of pragmatics in particular and critiqued for the differences between the
data elicited by this method and oral production (Rose, 1994). However, research on
enhanced completion tasks, where context and details are provided for the learner,
suggests that there is some validity in this kind of data elicitation as a means of
assessing proficiency in target forms (Billmer and Varghese, 2000). In future research
and teaching, we will address this issue with different types of summative assessment,
and one possibility will allow time for a final SI routine as a way of assessing learner
development.

**Future practice and research**

This was an exploratory study conducted in an atypically small class to enable control
of the teaching/learning process, and exploration in some detail of the developmental
processes taking place. In future teaching cycles, it seems prudent to prepare the class
for an extended SI such as the one reported here, by enacting a much shorter and less
demanding SI sequence earlier in the semester. This will familiarize learners with the
procedure and help develop some basic interactional strategies that should facilitate a
stronger performance in the later SI. Further research into the design of the SI
scenarios is also required for this population of students, so that they receive effective
training and are able to complete the given task with the aid of appropriate feedback.
Further investigation, with class sizes that can extend to 35 or 40 learners, is clearly
also required to explore the dynamics and logistics of the teaching design for the
Japanese context, which is one of the stated goals of this research. Having participants
record, upload and transcribe their own scenarios may be necessary with a larger class
and could provide beneficial motivation for learners who otherwise may not be so
active in their own learning processes. Recording and analysis of the debriefing
sessions may also provide further insight into learner appropriation of target forms.

**Conclusion**

The aim of this study was to investigate the efficacy of technology enhanced strategic
interaction as a principled way of developing language proficiency in a particular
context, characterized by particular barriers to successful learning outcomes.
Specifically, we engaged the participants in a meaningful cycle of activity that
provided learning affordance through engagement in a type of communicatively
oriented meaning-making. Participation in this design sequence therefore, allowed
the learners to co-create conditions for their own language development, even though
the task itself can be seen as having been too difficult for the learners, since none of the 9 groups were able to successfully solve the problem in their first performances. The technology provided an important element to the sequence and allowed the learners to use their own performance as an artifact to mediate their own development. This self-analysis of performance and the subsequent repeat of the scenario provided the basis for further control and awareness of language use in the given context.

This study highlights the understanding of learner development provided by microgenetic analysis of learner talk, by revealing shifts towards greater self-regulation in the process of language development. These findings reflect previous work in the SCT framework (Frawley, 1985; Ahmed, 1994; Gutierrez, 2008) and the study provides a basis for understanding the developmental impact that an SI design can have on relatively low proficiency learners. In particular, our study highlights the potential importance of mediational affordances that can be provided to low proficiency learners, such as the use of a model conversation that the learners were able to appropriate into their own interaction and ultimately their own language repertoire, as suggested by the successful use of forms from the model in the final discourse completion task. The transcripts of their own performances, available through the affordances of available technology, also became an important artefact for reflection and further language use, in common with earlier research on the value of learner language as a part of the development cycle (Lapkin and Swain, 1998).

The participants generally reported a positive experience with the SI procedure and this sequence provided an intensive learning experience in this context for learners for whom spoken communication is often not a primary concern. Evidence of learner agency was seen in appropriation of language from the instructional model, selective uptake of the feedback used to scaffold subsequent performance, and in changes made to the scenario itself to better suit learner capabilities and goals. Through artifact-mediated development, higher levels of language performance and self-regulation were seen. This learner agency serves as a reminder that in the original spirit of Di Pietro’s (1987) work there was a deep concern and respect for the humanist elements of learning a second language: “To speak is to be human, and to speak another language is to find new ways to express that same humanity” (p.12).

Acknowledgments
The authors would like to thank reviewers for comments and instructive feedback on
an earlier version of this paper.

**Further Information**

Materials and further information regarding implementation of Strategic Interaction in language proficiency classes are available to download at the following:

http://langcom.u-shizuoka-ken.ac.jp/si

**References**


Negueruela, E. (2003). *A sociocultural approach to the teaching and learning of*
second languages: Systemic-theoretical instruction and L2 development.

Unpublished Ph.D. dissertation. The Pennsylvania State University, University Park, PA.


Appendix 1 – Model dialogue

In preparation for your dialogues next week, study the following conversation between two business colleagues. They are discussing a problem in a project. Notice how they use modal verbs and politeness strategies. Underline examples that you see in the dialogue.

Today there is a meeting at the Tokyo head office of an import company. Mr. Yamamoto is meeting Mr. Jones who is coming to the office from the United States to discuss a project they have been working on:

Yamamoto: Nice to meet you. I’m Yamamoto and I am head of marketing here in Tokyo.
Jones: Nice meeting you too. I am Tom Jones, from the office in San Francisco.
Yamamoto: Welcome to our office, Mr Jones. Please take a seat.
Jones: Thanks!
Yamamoto: So, about the online marketing project. I have a question about the online information.
Jones: Yes, I sent you the files last week. I hope you received them OK?
Yamamoto: That’s right, no problem. But there seems to be some information about the new product missing
Jones: Oh really? What’s missing?
Yamamoto. Yes, the information about the new furniture line was not finished. I wondered if it would be possible for you to finish that this week?
Jones: Oh, I’m sorry, I have to return to America tomorrow, so it is going to be difficult to get that finished this week. When is the deadline? Is this urgent?
Yamamoto: Actually, our deadline for the printing of the new brochure is next Wednesday.
Jones: Could I ask you for more time on this? I might need two more weeks to get the furniture information to you.
Yamamoto: Well, it’s going to be difficult, but I will talk to my supervisor and I’ll let you know.
Jones: OK thanks, I’m really sorry for this delay and I would really appreciate more time.
Yamamoto: I understand. Well, thanks for coming in today, it was nice talking with you.
Jones: Thanks, I will be in contact through email next week.
Appendix 2 – Scenario cards

Role A: Marketing Assistant Manager – Roppongi Branch

Identity:
- Name: Your own name: ______________________
- You are a native speaker of Japanese
- You are working for Yahoo Japan in Roppongi. Your job is to advertise and market the Yahoo group.
- You are working to produce new advertising brochure for your company

Situation:
- It is Monday.
- You have a meeting with the Osaka branch manager to discuss the new brochure (Tuesday).
- The deadline for the brochure is Friday – it MUST be finished by then or your manager will be disappointed and the work delayed.
- The manager from Osaka has made some of the brochure but there are some mistakes in his/her work. You are not satisfied with the quality – you need to ask him/her to check for spelling and grammar mistakes. Also, there is a page missing. Page 3.
- You know the brochure will be good when it is finished but you need to ask your colleague to finish their work and get it to you by Thursday at the latest.
- The Osaka manager might be very busy, but they need to finish their work.

Your Calendar:

<table>
<thead>
<tr>
<th>Sunday</th>
<th>Monday</th>
<th>Tuesday</th>
<th>Wednesday</th>
<th>Thursday</th>
<th>Friday</th>
<th>Saturday</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(today)</td>
<td>1:00-2:30: Meeting with Osaka branch manager</td>
<td></td>
<td></td>
<td>Deadline for the printer: 5pm</td>
<td>10:00 – 6:00 Printing</td>
</tr>
</tbody>
</table>

Location:
- Roppongi office (it has a closed door, several desks and chairs, and a computer)

Relevant Information:
- You have made an appointment to speak with the Osaka manager during office hours.
- It is the middle of November. You must finish the brochure project and submit it to the printer this week. This project is very important for your work and reputation in the company.

Goals:
- To finish the project.
- To get an agreement with the Osaka manager about changes that need to be made.
- To be respectful of the branch manager of Osaka.
**Role B: Yahoo! Japan! Osaka Staff Working in Advertising**

**Identity:**
- Name: Michael or Michelle Smith, from the US.
- You are a work for Yahoo! Japan in Osaka
- You have been working in Marketing for 5 years and have some experience
- You are helping the new staff in the Tokyo office produce their new advertising brochure
- You submitted the files that they asked from you last week and now you will have a short meeting in Tokyo (Tuesday) to discuss finishing the project
- If the Tokyo staff want you to do more work on the project – it might be difficult. You have to go to San Francisco for a conference. You need some time to prepare for this.

**Situation:**

**Meeting:**
- You will meet in Tokyo, before you go to Narita. You are trying to finish the brochure. You have submitted all the work that was asked of you and now you want to help the Tokyo office put the product together. However – you do not have much time to help them.

**Research:**
- You are writing a paper (and making Power Point slides) for a conference in San Francisco on Thursday. You need more time to finish that work. You may have more time next week.
- You will take the train to Narita Airport Tuesday night, spend the night at a hotel, and fly to San Francisco on Wednesday morning for the Thursday conference.

**Your Calendar:**

<table>
<thead>
<tr>
<th>Sunday</th>
<th>Monday</th>
<th>Tuesday</th>
<th>Wednesday</th>
<th>Thursday</th>
<th>Friday</th>
<th>Saturday</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Today)</td>
<td>Meeting in Tokyo to discuss the brochure. 1:00 – 2:30 Afternoon train to Narita</td>
<td></td>
<td>10:00 Flight departs (Narita Airport)</td>
<td>10:30 SF conference begins</td>
<td>- Conference</td>
<td>5:00 SF conference ends</td>
</tr>
</tbody>
</table>

**Location:**
- The Roppongi Head office (it has a closed door, several desks and chairs, and a computer)

**Relevant Information:**
- You have submitted your work for the brochure
- You have to leave for the US tomorrow.

**Goals:**
- To help and advise the staff in Tokyo to finish the brochure.
- To have enough time to write an excellent paper and Power Point slides for the conference.
Learner Self-management Procedures Reported by Advanced and Intermediate ESL Students

Minh Hue Nguyen

Monash University, Australia
Vietnam National University, Vietnam

Bio Data
Minh Hue Nguyen (M.A. in Applied Linguistics at Victoria University of Wellington) is a PhD candidate in the Faculty of Education, Monash University, Australia. She has taught in undergraduate and postgraduate programs in TESOL at Vietnam National University. Her main research interests are in the areas of TESOL teacher education, a sociocultural perspective on language teacher education, and language curriculum development. She works as a reviewer for Asian EFL Journal.

Abstract
This study aims to investigate the learner self-management procedures that advanced and intermediate ESL students used in their three-week preparation for a five minute seminar as part of their English Proficiency Program at a New Zealand university. The study used learner diaries, follow-up interviews, and classroom observations to collect data from 4 advanced and 6 intermediate ESL students. All of the students from both proficiency groups reported going through a range of self-management procedures including planning, self-monitoring, and problem solving. Both groups reported a limited amount of planning, i.e. setting goals, setting criteria, analyzing the task, and setting a timeline. However, within the planning procedure itself, the groups revealed different focuses. Moreover, the advanced students monitored their preparation more frequently and were better at problem solving than the others. These findings add new insights into the self-management procedures that students of low and high English proficiency followed in three weeks of preparing for their presentation.

Key words: Learner self-management, learner self-management procedures, pre-task planning

Introduction
Since much of learning occurs outside the classroom (Wilson, 1997), the study of conditions that contribute to student learning is essential to enhance students’ learning. Studies showed that self-management in language learning is a useful condition that helps students learn more effectively. Research into learner self-management and its relevant literature mostly focused on either expert – novice learners or successful - unsuccessful learners (Rubin, 2005) and contributed greatly to
our understanding of what strategies, processes or procedures can help students learn a second/foreign language effectively. However, past research did not adequately document the self-management procedures of low and high proficiency students. In addition, even the growing research of pre-task planning mainly focuses on the effects of planning on task performance (Foster, 1996; Skehan, 1998; Mehnert, 1998; Ortega, 1999; Foster & Skehan, 1999; Bygate, 2001; Yuan & Ellis, 2003; Kawauchi, 2005). Among a limited number of studies on the planning process (Ortega, 1999; 2005; Sangarun, 2005; Kawauchi, 2005), only Ortega’s (2005) study looked into the planning process of learners of different proficiency. More research is needed to shed light on self-management procedures used by learners of high and low proficiency in task planning.

In this study I examined effective self-management process and procedures in language learning. More specifically, the study aimed to identify what self-management procedures learners of high and low proficiency went through in preparing for their learning tasks and investigate the similarities and differences between the two proficiency groups in their self-management procedures. Two general questions were addressed in this study. The first question is: How do high and low proficiency students manage their independent preparation for the learning tasks? The second question is: Is that process different for low and high proficiency students?

This study has its contributions in terms of theory and practice. First, the paper introduces a research-based model of effective learning that can be applied in different language learning contexts. Second, the research helps teachers to understand the assistance that students need for their self-management in language learning. Since most of the participants came from an Asian background, there is the possibility of generalizing the findings to Asian EFL contexts. Moreover, being the first study to investigate low and high proficiency students’ self-management procedures in an extended period of learning tasks, the research adds new insights to the existing literature and stimulates more studies in this field.

**Theoretical framework**

Serving as the theoretical framework for the study, this part reviews theory and research on learner self-management and the procedures involved in successful self-management. It also reviews relevant research into pre-task planning with a particular focus on research findings on how learners of different proficiency levels
perform their pre-task planning. The section ends with a distinction of the two key terms *planning* and *preparation* and identifies the research questions based on the gap in the previous research.

**Learner self-management**

In the L2 literature, the terms *self-regulation* and *self-direction* have been used commonly with a close meaning to *self-management*. Zimmerman (2002) describes self-regulation as the selective use of specific processes, including task analysis, self-motivation, self-control, self-observation, self-judgment, and self-reaction. According to Wenden (2001), in the cognitive literature, self-regulation is described as the processes of learners planning for the task, analyzing the task, and monitoring task implementation, and self-direction refers to the same processes in the learner autonomy literature.

Rubin (2005) developed a model of learner self-management based on the theories built on over thirty years of research. This model (Figure 1) represents the interactive relationships among the procedures, among the knowledge and beliefs held by the learners, and between the procedures, knowledge, and beliefs (Rubin, 2005). More cognitive procedures are included in the new model, altogether including *planning*, *monitoring*, *evaluation*, *problem-identification* and *problem-solution*, and *implementation of problem-solution*. The knowledge and beliefs learners use in self-management consist of task knowledge, self-knowledge, beliefs about learning, background knowledge, and strategy knowledge. In Rubin’s model, every single procedure, kind of knowledge or beliefs has an interactive relationship with another procedure, kind of knowledge, and belief in shaping the way the learner self-manages.
Rubin (2005) also provides clear evidence that knowledge about tasks, strategy knowledge, background knowledge, and self-knowledge, together with beliefs, play an important role in successful self-management procedures. Therefore, in an educational setting, teachers’ instructions are important in promoting these kinds of knowledge, especially task knowledge and strategy knowledge, and thus fostering the success of learner self-management process. Adopting Rubin’s model, the present study focuses particularly on the self-management procedures that its subjects report using in their seminar preparation.

The learner self-management procedures

Planning: Rubin (2005) specifies a comprehensive and systematic set of steps in self-managed planning. The steps consist of setting goals, setting criteria to measure goal achievement, task analysis, and setting a timeline. Firstly, in goal setting, learners determine specifically what they wish to achieve within a predefined period of time (e.g. learn 30 new words each week). Secondly, in setting criteria, self-managed learners establish measures to assess their goal achievement. Thirdly, task analysis, which includes three components: task purpose, task classification, and task demands, is how the learners plan to approach the task. Task purpose is the pedagogical or real life objectives that learners want to achieve by doing the task (e.g. to pass an exam, to survive in a native country). Task classification is the identification of the characteristics of the task which helps learners decide task demands, i.e. knowledge, skills, and strategies they need in order to complete the task. Finally, self-managed learners set a realistic time line for completing the task.

Monitoring: According to Rubin, in monitoring, self-managed learners notice any difficulties they may have in their learning. For example, they may find it difficult to concentrate on the task, or to think of appropriate language items to use. They continually monitor their understanding and outcome and note the causes of their difficulties.

Evaluation: The self-managed learners, following Rubin’s model, evaluate their goal achievement based on the criteria they set while planning. They then decide whether they have performed appropriately and whether they need to solve any problem to meet their goals.

Problem-identification and problem-solution: Based on the problems noticed during their monitoring and evaluation, learners begin to consider the causes of their
difficulties and lack of success. After that they consider what they can do to solve the problems in order to attain their goals. The problem solutions may entail returning to earlier stages of their learning process such as setting goals, setting criteria and adjusting some of them.

**Implementation of problem-solution:** The self-managed learner implements the possible solutions to see whether they work for them. They may need to make changes to other procedures if necessary in order to fulfill the task.

The model of learner self-management was built on decades of research into metacognition (e.g. Rubin, 1975; Naiman, Fröhlich, Stern, & Todesco, 1978; Abraham & Vann, 1987, Wenden, 1987, 2001). Its procedures have been found to be used by successful/expert learners (Rubin, 1975, 2005; Goh, 2005; Bygate, 2005; Weigle, 2005). However, metacognition research has largely focused on good and poor language learners. In the L2 literature, while good/poor language learners are defined based on the learning strategies they use (Rubin, 2005), high/low proficiency learners are largely defined according to the levels of their language performance (Ellis, 2005; Kawauchi, 2005). Although good and poor language learners are found to differ in their ability to self-manage learning, whether or not high and low proficiency learners can be distinguished by their self-management ability is yet to be confirmed. One of the aims of this paper is to answer this query and the answer would have useful implications for teachers of different levels in improving their learners’ self-management.

**Pre-task planning and oral performance**

Ellis (2005) distinguishes two major types of task-based planning: pre-task planning (planning that takes place before the task performance) and within-task planning (planning that occurs during task performance (Bygate & Samuda, 2005). **Pre-task planning** can be further divided into rehearsal and strategic planning. Rehearsal refers to the opportunity for task repetition, of which the earlier performances are preparatory. Strategic planning concerns the opportunity to prepare for the task content and how to convey it to an audience. Studies on pre-task planning have mostly been following a product approach and/or a process approach.

**A product approach**

The last decade has seen a growing body of research into the effects of pre-task planning on oral performance as a product approach. Increased quality of oral
performance as a result of pre-task planning has been found in all three aspects of language production: fluency (Foster, 1996; Mehnert, 1998; Ortega, 1999; Foster and Skehan, 1999; and Kawauchi, 2005), accuracy (Mehnert, 1998; Foster and Skehan, 1999; and Kawauchi, 2005), and complexity (Foster, 1996; Skehan, 1998; Mehnert, 1998; Yuan and Ellis, 2003; Bygate, 2001; and Kawauchi, 2005).

The aforementioned studies asserted the important roles of pre-task planning in improving learners' speech production and contributed significantly to the literature for the reason that they addressed the conditions in which pre-task planning can be most beneficial to learners. However, the type of planning investigated in all of these studies is short-term planning, which lasts for a maximum of 10 minutes in the classroom. It is different from the long-term 3-week planning studied in this project. Therefore, the positive effects of short-term planning as discussed above might not be generalizable to the long-term planning concerned. In addition, most of these studies did not involve what processes learners went through in their planning. It is very important to understand the planning process since they may enable teachers to better facilitate learners in their task planning.

A process approach
To date, only a few researchers have investigated the process of pre-task planning (Ortega, 1999; Ortega, 2005; Sangarun, 2005; and Kawauchi, 2005). Ortega (1999) set out to establish what learners actually did when given 10 minutes to plan for their story-telling task with an aim to find out whether learners took that opportunity to focus on form. The planning opportunity was found to facilitate learner-initiated and learner-driven focus on form. Besides, learners were found to use a number of strategies such as focal attention to problem-solving, rehearsal, memory-related strategies, etc.

While Ortega (1999) investigated learner behaviors in unguided planning, Sangarun (2005) used guidance on three foci (form, meaning, and both form and meaning) to investigate its effects on the students’ cognitive planning processes for an instruction/argumentative task. The three types of instruction on form, meaning, and both form and meaning were reported to guide learners’ attention to form, meaning and both form and meaning respectively. Sangarun’s findings highlight the influence of teacher instructions on planning.

Kawauchi (2005) also took the process of planning into account. However, rather than trying to discover what planning activities learners engaged in, the researcher
looked at how each of the three pre-specified planning activities (writing, rehearsal, and reading) influenced task performance and reported that the different learners doing three different planning activities gained similar improvements regardless of which planning activity they engaged in.

In her following study, Ortega (2005) extended the scope of her previous research on planning and investigated strategic processes that low-intermediate and advanced learners engaged in while planning. The participants were discriminated based on their course levels (semester 4 vs. semester 6 and beyond) and by means of complexity, accuracy and length of utterances observed in their narratives. Although no significant differences were found in the number of strategies used by the two proficiency groups, the results did show that given the planning opportunity, advanced learners put more balanced effort on retrieval and rehearsal, and stronger engagement in self-monitoring strategies. On the other hand, low-intermediate speakers paid more attention to retrieval strategies to solve lexical and verbal morphology difficulties.

These studies have touched upon a domain which has been much focused on lately: the learning process. However, the first three studies have not focused specifically on what actions, strategies, or behaviors learners took in focusing on form/meaning/form and meaning, in writing, rehearsing, and reading. Only in Ortega (2005) were the students’ planning processes described in detail. More research into this area is necessary in order to find out a better description of pre-task planning processes in different task conditions, and this study was conducted to fill in this research gap.

**High and low proficiency students in previous pre-task planning research**

Although many studies have been done on the relationship between L2 proficiency and strategy use (Griffiths, 2003, 2004; Green & Oxford, 1995), only a few studies have taken strategies used by high and low proficiency students in pre-task planning into consideration. Significant results have been found about the differences between these two groups of students in planning. Wigglesworth (1997) found that only high proficiency students benefited from the one-minute unguided planning and increased fluency, accuracy and complexity of their language use in an oral test. Learners of low proficiency did not benefit from the planning opportunity.

Following Wigglesworth (1997), Kawauchi (2005) investigated the effects that Low, High, and Advanced proficiency had on learners’ pre-planned oral performance. She reported that High proficiency learners gained significant fluency and complexity whereas Low proficiency students gained most in accuracy; moreover, the Advanced
students benefited the least from the opportunity to plan. The researcher considered a ‘ceiling effect’ as a possible reason why the Advanced students gained the least from the planning opportunity because their high proficiency might enable them to handle the task well either with or without planning. Ortega (2005) investigated planning processes for learners of different levels (as determined by fourth and sixth semester students) and found that advanced and low-intermediate students employed different strategies in pre-task planning which in turn contributed to the difference in quality of performance. This finding is very significant to teachers in planning instructions.

The literature review suggests that pre-task planning plays important roles in improving L2 oral performance. Meanwhile, some limitations of the studies should be addressed. Firstly, all studies used short-term tasks such as narrative (Skehan & Foster, 1997) and decision making (Foster & Skehan, 1999) for which the students had a few minutes in the classroom to plan. Task planning which takes place within an extended period of time and occurs mostly outside of class has not sufficiently drawn researchers’ attention. Research into this kind of long-term planning will help the study of procedures that are not available to observe in short-term planning research such as how learners self-manage the extended planning time and the way through which they make use of the resources available to them from both inside and outside of class. Secondly, the current body of research mainly focuses on the product of pre-task planning. The few studies that looked into the processes of planning were either limited to short-time planning and/or general teacher-guided processes rather than learner self-managed planning. Therefore, studies on the self-managed planning process of learners of different proficiency are needed.

**Planning vs. preparation**

As far as terminology is concerned, the literature review suggests that the term *planning* has been used in both the literature on learner self-management and that on pre-task planning. However, it is perceived that these uses refer to different processes in the two cases. In the learner self-management literature, it indicates a metacognitive procedure which self-managed learners go through mostly to make an action plan for doing the task. In the pre-task planning literature, it refers to a broader process of preparing for the performance of the task which may include the learner self-management planning and other procedures such as writing, rehearsing, and reading (Kawauchi, 2005). This study investigated the process of pre-task planning that involves several self-management procedures including planning. In order to
avoid terminological confusions, planning will be used only as one of the learner self-management procedures and preparation will be used to refer to pre-task planning.

**Research questions**

In order to identify self-management procedures for learners of high and low proficiency in preparing for their seminars and find out the similarities and differences between the two proficiency groups in their self-management procedures, the following questions are examined:

1. What self-management procedures do advanced ESL students report using in preparing for their seminars?
2. What self-management procedures do intermediate ESL students report using in preparing for their seminars?
3. What are the similarities and differences between the two proficiency groups in their use of self-management procedures in their seminar preparation?

**Research design**

**Context and participants**

This study was conducted in the English Proficiency Program (EPP) at a university in New Zealand in 2008. This twelve-week program prepared ESL students for academic studies at the university. As an alternative to IELTS and TOEFL, the EPP scores were used as language proficiency proof to seek admission to academic studies at the university. After successful enrolment for the program, students took placement tests on dictation, reading comprehension, and writing. Based on the test scores, the students were classified into proficiency levels and placed into classes 1 to 4 from highest to lowest proficiency. At the end of the program, the students took four tests on listening, speaking, reading, and writing which were similar to the IELTS in terms of scope and testing rigor. In addition, they conducted a 10-minute seminar presentation, the score of which made up a part of the speaking score.

Formal permission for this research was obtained from the Head of School, the Program Director, the teachers, and the Human Ethics Committee through the human ethics procedure. Students from an advanced and an intermediate class (as assessed by the EPP placement tests) were purposefully selected for this study because they consisted of two groups of English proficiency levels taught by one instructor with the same curriculum. The participants completed the consent form before the study. By
the end of the data collection period, 4 sets of diaries from the advanced students and 6 others from the intermediate group were handed in (3 intermediate students withdrew from the study). Nine participants came from Asian countries including Korea, China, Japan, Myanmar, and Thailand, and one student is from a European background (Germany). The participants are described in Table 1.

Table 1: Participant characteristics

<table>
<thead>
<tr>
<th>Factor</th>
<th>Learner characteristics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Proficiency level</td>
<td>Advanced (n=4)</td>
</tr>
<tr>
<td>Age</td>
<td>18-30</td>
</tr>
<tr>
<td>Language learning experience</td>
<td>7-15 years, EFL</td>
</tr>
<tr>
<td>Gender</td>
<td>3 male, 1 female</td>
</tr>
<tr>
<td>L1 background</td>
<td>1 European, 3 Asians</td>
</tr>
</tbody>
</table>

**Data sources**

In her reviews of the larger literature on learning processes/strategies, Chamot (2001, 2005) found that self-report is a meaningful data source to investigate learners’ strategic processes for a language task since it helps researchers get an insight into learners’ unobservable thinking processes. According to Chamot, self-reports in learner strategy studies have been carried out through stimulated recall interviews, questionnaires, think-aloud protocols, written diaries/journals, and retrospective interviews associated with a particular learning task.

In order to gain insights into learning processes, this study used multiple data sources as suggested by Chamot: interview, journals, and classroom observations. The triangulated data sources allowed me to gain an insight into the subjects’ preparation processes from different angles. I was able to collect as much data as possible from a small number of participants within their learning setting in a rather extended period of time.

Diaries were chosen as the primary source of data for this study for three main reasons. First, this was considered an appropriate method for investigating the independent preparation process in an extended period of time (Adams, 2007). Second, diaries enabled the participants to report their internal processes and thoughts about the preparation experience which might not be accessible from the researcher’s perspective (Mackey & Gass, 2005). In addition, written diaries allowed learners to freely reflect on their learning processes rather than being constrained to a pre-defined
set of strategies as in the cases of questionnaires and structured interviews, which is important for this exploratory study.

Follow-up interviews were chosen as a secondary data source. The interviews were used to compensate for the weaknesses of diaries such as students’ unwillingness to write and limited ability to self-reflect on their preparation, etc. The interviews enabled me to elaborate on the self-managed preparation procedures reported in the students’ diaries and gather further information relevant to their self-management.

Classroom observation was conducted with an aim to gain more information about the participants’ self-management in the seminar preparation during class time. The observation also aimed to see if the teacher provided any instructions and feedback on self-management procedures which might influence the students' self-management. Through observation, I planned to ensure that similar seminar instructions were given to the classes.

Materials and piloting
The study used a five-minute seminar presentation task that all students of the program required as part of their learning program. After a session introducing the seminar program given by the teacher, the students chose their own academic topics and had three weeks to prepare for their presentations. Each class met once a week for a one-hour seminar session. The meetings were opportunities for them to get instructions and feedback on giving a seminar. After the three preparation weeks, each student gave a presentation individually. The audience included their classmates and another teacher other than his/her own teacher.

A handout of guidelines for writing a diary about preparing for the seminar was used in this study. The diary guidelines (Appendix A) followed the structure used by Krishnan and Lee (2002). They were given and explained to 5 low proficiency students other than the subjects to test comprehensibility. The students were asked to write one-day diaries following the guidelines. The results of the materials testing showed that the students understood the guidelines well and were able to follow the guidelines.

Data collection
In order to address the three research questions, learner diaries, follow-up interviews, and classroom observation were used to collect data. The research procedure was summed up in Table 2.
Diaries

One week before the subjects started preparing for the seminar, I conducted a training session with them and the teacher for 45 minutes to train the subjects in keeping a diary. A sheet of the diary guidelines (Appendix A) was given to each student. They had a chance to look at the guidelines carefully and were welcome to ask questions for clarification. To avoid undesirable training effects on the research results (e.g. the students might use the same preparing strategies as modeled), I modeled writing an example diary about independent vocabulary learning instead of preparing for the presentations and strategies associated with learner self-management were avoided. More clarification was given two days later based on the subjects’ questions.

Table 2: Data collection schedule

<table>
<thead>
<tr>
<th>Week</th>
<th>1</th>
<th>2-5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Activities</strong></td>
<td>Materials testing; Training in diary writing</td>
<td>Task was introduced; Diaries were written almost every day and checked weekly; Classroom observation and weekly interviews were conducted</td>
<td>Diaries were examined to prepare for follow-up interviews; Follow-up interviews were conducted</td>
</tr>
</tbody>
</table>

Diary keeping requires a high level of commitment from the diarists (Mackey & Gass, 2005; Adams, 2007). To increase the participants’ commitment, prior to the study, I discussed with the teacher and participants about the teaching/learning benefits of diaries and the value of the present research in promoting self-management as well as presentation skills. Each participant received a book voucher, chips, and chocolates for taking part in the study.

The diary writing took place within three weeks, from the second week of the program. As soon as the presentation task was introduced and assigned by the teacher, each participant was given a diary notebook with guidelines and asked to keep diaries about his/her preparation process. They were encouraged to write every day during the preparation period. I collected the students’ diaries once a week at the end of each seminar session to make sure that the students were on the right track and returned to them later that day or the next day. The subjects finished writing diaries and returned
them to me on the days when they gave their seminars.

*Follow-up interviews*

Together with diary studies, the follow-up interviews helped to collect information that would answer the three research questions. During the preparation period, I met each participant at least once a week and asked them questions about their preparation process during the week. As soon as the diaries had all been collected, I examined them and prepared questions for final follow-up interviews based on the diaries. Then I made appointments with individual participants. The initial examination of diaries and the interviews took place within 9 days. Each participant was interviewed for around 10 minutes. To create a friendly, comfortable and relaxing atmosphere, the interviews were conducted at places familiar to the participants (Mackey & Gass, 2005) such as the classroom and the library. All interviews were audio-recorded.

*Classroom observation*

I attempted to observe every seminar instruction session with the two groups to gather more information about the students’ preparation processes and gain an insight into the teacher’s instructions and feedback given to each class. The observation was unstructured with notes taken and the lessons audio recorded.

*Data coding and analysis*

Before data were coded and analyzed, each participant’s real name was replaced with a code. Each code comprises of either Ad. (for Advanced) or Int. (for Intermediate) and a number to distinguish the student from the others in the same group. All advanced students were named from Ad.1 to Ad.4 and intermediate students from Int.1 to Int.6.

The data for this study consisted of students’ written diaries, transcripts of follow-up interviews, classroom observation recordings, and observation notes. Data was initially coded as directed to learner self-management procedures in Rubin’s (2005) model namely planning, monitoring, evaluation, problem identification, and implementation of problem solutions. Planning was further coded as setting goals, setting criteria, task analysis, and setting a timeline. Task analysis was coded as task purpose, task classification, and task demands. The coding system was illustrated as follows with “to make it in another language” and “to make it good” coded as goals:
Interview excerpt 1:

Setting goals: The only goal was…to make it in another language and try to make it good

Although the study followed Rubin’s procedures in data coding, these procedures were not imposed on the data. Rather, occurring themes were allowed to reveal naturally (Mackey & Gass, 2005, Gan, Humphreys & Hamp-Lyons, 2004; Adams, Fujii & Mackey, 2005) to exploit the richness of qualitative data. As a result, some emerging patterns were developed from data, and the identified learner self-management procedures were categorized into three broad groups (planning, self-monitoring, and problem solving) instead of five procedures as in Rubin’s model. Detailed explanation of this will be offered in the discussion of findings.

The coded data was then quantified to identify the percentage of participants performing self-management procedures and the mean frequency (M) for each procedure was then calculated for each group by taking the average of the group members’ frequency of using each self-managed procedure.

Results

Adopting Rubin’s (2005) model of learner self-management, this study focused on discovering the self-management procedures reported by ESL students of advanced and intermediate level. The findings from diaries, interviews, and observation reveal the learning processes that the two groups of students reported going through in preparing for their seminar task. In this section, the results are presented and discussed in order of the research questions stated at the end of the literature review.

Self-management procedures reported by advanced students

The first question was as follows: What self-management procedures do advanced ESL students report using in preparing for their seminars? Data from the advanced level students’ diaries and interviews were coded and results were displayed in Table 3:
Table 3: Advanced students’ reported LSM procedures (n=4)

<table>
<thead>
<tr>
<th>LSM procedures</th>
<th>%*</th>
<th>Mean frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>1. Planning</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Setting goals</td>
<td>50</td>
<td>1.5</td>
</tr>
<tr>
<td>Setting criteria</td>
<td>75</td>
<td>2.3</td>
</tr>
<tr>
<td>Task purpose</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Task analysis</td>
<td>100</td>
<td>3.0</td>
</tr>
<tr>
<td>Task classification</td>
<td>100</td>
<td>10.5</td>
</tr>
<tr>
<td>Task demands</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Setting a time line</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td><strong>2. Self-monitoring</strong></td>
<td>100</td>
<td>19.0</td>
</tr>
<tr>
<td><strong>3. Problem solving</strong></td>
<td>100</td>
<td>9.3</td>
</tr>
</tbody>
</table>

* % = percentage of advanced students performing the procedures

**Planning**

Overall, although every advanced student reported going through a planning procedure, the reported amount of planning was moderate. The steps that the subjects took in the planning procedure consisted of setting goals, establishing criteria, and analyzing the task.

*Setting goals:* All of the students were asked in follow-up interviews whether they set any goals for their seminar task. The data from interviews together with diaries reveal that 50% of the advanced ESL students did not report setting any goals. The remaining two students, Ad.2 and Ad.4, reported setting a common goal of making a good or excellent seminar:

*The only goal was...to make it in another language and try to make it good.*

(Interview excerpt 1 - Ad.2)

I want to have an excellent 5 minute seminar. (Diary excerpt 1 - Ad.4)

Both Ad.2 and Ad.4 broadly wanted to do the seminar well in the target language, which was challenging enough for them and this was a worthwhile goal to aim at. However, this was too general a goal and an examination of the two students’ diaries and interviews shows that Ad.2 did not break down his goal into more specific short-term goals to achieve at different stages of the task. Only Ad.4 reported setting specific sub-goals such as “I wanted to find something about my seminar” and “I want
to understand those data” to work toward the final goal of making an excellent seminar.

**Setting criteria** An analysis of the data from the advanced students’ diaries and interviews shows that 75% of the students reported setting criteria to measure their success. A participant reported in her diary:

"...a seminar that has a good structure. It should be clear. It should be understandable to the other person. Probably there should not be so many mistakes, especially grammar and pronunciation. (Diary excerpt 2 - Ad.2)"

**Task classification** In general, among the planning steps, task classification was one that received significant attention from the advanced students. All of them were found to classify the seminar with regard to its characteristics and genre. While the four advanced students classified a seminar as generally having an introduction, a body, and a conclusion, Ad.3 and Ad.4 also realized that finishing the seminar within 5 minutes was a challenging job. Particularly, Ad.4 also perceived it as an academic seminar in which she should use some academic vocabulary.

**Identifying task demands** Careful examination of the subjects’ diaries and interviews indicates that 100% of the advanced students reported considering the task demands. However, they did not globally consider the strategies and knowledge necessary to do the task at the planning stage. Rather, the students followed the seminar program schedule, which divided the seminar task into smaller weekly tasks such as forming a topic, gathering sources of information, making a structure, and so on. The extract from Ad.2’s first week diary entry below is a typical example of the task demands the students identified during the first week of the preparation period in response to the subtask of finding a topic. She made no reference to the demands of other parts of the big seminar task:

"The most important thing is to find the right topic that you are interested in, you are able to speak about.... The next thing to do is to get English sources.... I went to public city library just to get an overview of a possible topic. I started to read two nights...about medication to decide on the topic. (Diary excerpt 3 – Ad.2)"

All the advanced students reported identifying task demands at different stages of their preparation, which helped them self-manage different parts of the task. However, none of them globally considered the task demands of the whole seminar task at the early stage of the preparation period so that they could make an appropriate and flexible global action plan for themselves in preparing for their seminars. This
indicates that the participants lacked global strategic planning for the task.

**Self-monitoring**

The data analysis shows that all students from the advanced groups reported self-monitoring in preparing for their seminars at a very high rate, with an average frequency of 19 times. In their diaries and interviews, the advanced students reported that they frequently *self-evaluated* their learning activities and the outcome of those activities. Sometimes they evaluated their performance against their established goals. Ad.4, for example, against her goal to “make it in another language and try to make it good”, evaluated her rehearsal as “the first try was terrible” (Ad.4, diary). Even though when goal setting was not reported, the subjects still self-evaluated their learning process and outcome:

*When I practiced the seminar, I used the portable recorder. It’s very helpful because you can hear repeatedly what you say and you can do it over and over again.* (Diary excerpt 4 - Ad.1)

*So upset was I today. In the class all students could speak fluently but I could speak for only a few minutes.* (Diary excerpt 5 - Ad.3)

They also identified problems and tried to locate the sources of those problems:

*Before I wrote this diary I surf the net. I wanted to find something about my seminar.... Although it shows lots of detail about my topic, all of it is academic words and special words.* (Diary excerpt 6 - Ad.4)

*My topic is not very specific. So it’s difficult to solve the problem. When I chose the topic I thought it is easy to suggest some solutions but actually I found out it’s not so easy to suggest. The more information I get the more abstract the presentation will be.* (Diary excerpt 7 - Ad.3)

**Problem solving**

The data analysis indicates a common problem-solving pattern among the advanced students. The pattern was based directly on the result of problem identification, which is part of self-monitoring. The results typically involved three main self-management procedures: *problem identification/problem anticipation, problem solution, and implementation of problem solution*. Once the students noticed their own difficulties in the preparing process and the possible causes of them, they selected or adapted strategies to solve the problems and implemented the solutions. Following are some extracts from the subjects’ diaries and interviews to illustrate this finding:

*The amount of information is so much so I was overwhelmed (monitoring).... I*
found that it’s really important to decide what I want to talk about in the lecture (problem solution). So I started to plan the lecture: the causes of decreasing number of international students, the second is the effects of it and the third is what the government expectation or something like that (implementation of problem solution). (Interview excerpt 2 - Ad.1)

Every minute when I thought about it, I felt kind of unsure, unconfident, not really frightened but kind of negative feeling (monitoring)...I changed the topic (problem solution). When I really get nervous or thinking about the seminar...I had to do the shopping, to prepare a cake or something else just to avoid that I am thinking about the negative issues (implementation of problem solution). (Interview excerpt 3 - Ad.2)

However, in several occasions the problem solution and implementation of problem solution procedures were not distinctly reported but the subjects actually carried out both procedures like in the following extract:

I wanted to find some information about...New Zealand but that was very difficult...and all the information is about UK (monitoring). So I changed the topic to UK (problem solution and implementation of problem solution). (Diary excerpt 8 - Ad.4)

It is found that the advanced students took active control of their learning. They reported noticing their own difficulties, thinking of ways to overcome the difficulties on their own, and trying out the solutions to see if they worked for them. This ability is central to learner self-management.

**Self-management procedures reported by intermediate students**

The range of self-management procedures found among the advanced students was also reported by this group and they were also categorized into planning, self-monitoring, and problem solving. The findings were presented in Table 4.
Table 4. Intermediate students’ reported LSM procedures (n=6)

<table>
<thead>
<tr>
<th>LSM procedures</th>
<th>%*</th>
<th>Mean frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>1. Planning</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Setting goals</td>
<td>83</td>
<td>2.0</td>
</tr>
<tr>
<td>Setting criteria</td>
<td>17</td>
<td>0.5</td>
</tr>
<tr>
<td>Task purpose</td>
<td>67</td>
<td>1.0</td>
</tr>
<tr>
<td>Task analysis</td>
<td>83</td>
<td>1.3</td>
</tr>
<tr>
<td>Task demands</td>
<td>100</td>
<td>12.5</td>
</tr>
<tr>
<td>Setting a time line</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td><strong>2. Self-monitoring</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>11.0</td>
</tr>
<tr>
<td><strong>3. Problem solving</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>6.2</td>
</tr>
</tbody>
</table>

* % = percentage of intermediate students performing the procedures

**Planning**

Setting goals: 83% of the intermediate students reported setting goals for their seminars. The goal of making a good five minute seminar was commonly set by three of them, including Int.2, Int.4, and Int.5. Noticeably, Int.1 just wanted to “make” or to pass it and Int.6 only reported setting a specific goal of speaking more fluently. Int.4 was the only student who reported setting both a general goal and specific goals towards achieving it. He reported in an interview:

_I want to do a better seminar. I want to attract the audience to listen to my seminar. And I want to do my body language better because I think my pronunciation is not very good so I want to do some body language to attract the audience…. I hope I can speak fluently._ (Interview excerpt 4 – Int.4)

Setting criteria: Int.1 was the only student of the intermediate group who set criteria to evaluate his task performance. According to him, making a good seminar meant:

_Speak clearly that everybody understands you and to make up the best structure you can have…. You have to fit the information in 5 minutes and you are not supposed to hurry in terms of speed._ (Diary excerpt 9 - Int.1)

Identifying task purposes: It is noticed that 67% of the intermediate students considered the task purpose as to practice giving a seminar in preparing for a ten minute one which they would give at the end of the course. Another purpose commonly identified by them was to improve their speaking ability. One of the intermediate students wrote in his diary:

_I think it’s a good chance for me to practice myself and improve my English_
speaking because in the future I will use it.... I hope...I get the good mark in the ten minute seminar. (Diary excerpt 10 - Int.5)

Task classification: 83% of the intermediates reported classifying the task. The students identified a common characteristic of a seminar presentation, i.e. its introduction-body-conclusion format. Int.5 also classified the seminar as an academic presentation. Some other intermediate students further classified their own particular seminar as having a cause-effect-solution structure (Int.2, Int.5) or involving the use of visual aids (Int.3, Int.6).

Identifying task demands: Similar to the advanced students, 100% of the intermediate students identified task demands and adapted or invented strategies to meet the demands but none of them globally considered the requirements of the whole seminar task at an early stage of their preparation. In fact, their self-reports show that they relied on the class schedule and their use of task specific strategies followed the weekly homework requirements. For example, during the first preparation week, they mainly focused on finding a topic for their seminars and during the second week they focused on collating information for their topics:

Today I printed information from the internet about my seminar but I thought it's old one. Maybe tomorrow I'll search again. (Diary excerpt 11 - Int.5)

Meanwhile, during the third week, typically the students wrote about the task demands of making a structure for the seminar:

I wrote the structure about the causes and effects and solutions. In my mind I think causes are linked to the effects. I was thinking about how many effects.... I wrote my key points down in the sheet. (Diary excerpt 12 - Int.2)

Self-monitoring
All the intermediate students reported self-monitoring. This group of students often evaluated their strategy use and the outcomes of their learning activities and noticed their problems in preparing for the task. Following are some examples to illustrate how intermediate students self-monitored their seminar preparation.

Today I just thought about my seminar structure. There is not enough information. I downloaded a lot and chose. Oh no too much now. (Diary excerpt 13 - Int.2)

I was worried about the time when I stand in front of the audience. There are difficulties about my pronunciation. When I stand there, I forget something. I worry about that... (Interview excerpt 5 - Int.3)
We just had 2 minutes to finish it. But I have not talked all the parts in the practice. At night I practiced two times for my seminar. I finished it in about 5 minutes. I think that it is not a fluent seminar when I practiced it. (Diary excerpt 14 - Int.4)

I practice in front of other class students. It was so helpful for my seminar. Before I start practice, I thought 5 minutes is too long for me. But I noticed it is a short time. I worry about answering the questions. (Diary excerpt 15 - Int.6)

From the data that show the students’ self-monitoring, it is found that by noticing difficulties in the process of preparing for their seminars, the students actually started the first step of a problem solving procedure the findings about which are presented in the following part. Therefore, some evidence suggested a strong connection between self-monitoring and problem solving in the subjects’ self-management.

**Problem solving**

All the intermediate students also reported problem solving procedures. A number of problems encountered in the course of their seminar preparation were identified and solved. Their problem solving normally involved identifying the problems, finding solutions and implementing the solutions. A typical characteristic of the intermediate students’ problem solving is that they relied on assistance from other people such as the teacher and classmates. Let us look at the following illustrations:

*The teacher said my topic is not so specific (problem identification) so I’d like to choose a suitable topic (problem solution). (Diary excerpt 16 - Int.2)*

*I show the topic to the teacher but I couldn’t collect any information for the main points (problem identification). So I talked to the teacher about the topic and I searched for information on the internet (problem solution and implementation of problem solution). (Diary excerpt 17 - Int.3)*

Asking for help is considered a good strategy, especially for low level students who are not yet able to solve some problems by themselves. For the intermediate participants, they might lack certain knowledge to identify and solve their problems and therefore seeking help from other people seemed to work well for them. However, it is generally observed that most of them did not have a holistic look in their problem solving process. Sometimes they identified their problems but did not go further to find solutions to the problems and implement them. Similarly, sometimes they failed to define the specific nature of problems before seeking appropriate solutions.
The similarities and differences between the two groups in self-management procedures

The third research question was as follows: What are the similarities and differences between the two proficiency groups in their use of self-management procedures in preparing for their seminars? In order to answer this question, a contrastive analysis of data from both proficiency groups’ diaries and interviews was conducted. The findings from classroom observation were also analyzed to see whether the teacher’s instructions and classroom schedules for the two groups could have caused the noted similarities and differences.

Table 5 shows that there were two common characteristics of the two groups in their self-management procedures. First, both groups reported going through the same range of learner self-management procedures which were grouped into planning, self-monitoring, and problem solving. An analysis of the data from classroom observation showed that there was no explicit instruction on these self-management procedures. There was only one time during the final preparation week when the teacher asked both groups to do an unprepared 4/3/2 activity in which the students told their partners about their difficulties and solutions. However, the activity required each student to speak about their problems and solutions three times to three different partners and there was no interval time between each speech. Therefore, they did not receive any feedback from their teacher and partners during this activity. Moreover, the data from their diaries and interviews show that they had carried out problem solving earlier than when the 4/3/2 activity was conducted. Therefore, the students’ use of those procedures were probably mainly self-initiated, and no direct influence of instructions and feedback on the students’ self-management was observed.
Table 5: Advanced (n=4) and intermediate (n=6) students’ reported LSM procedures

<table>
<thead>
<tr>
<th>LSM procedures</th>
<th>%*</th>
<th>Mean frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Advanced</td>
<td>Intermediate</td>
</tr>
<tr>
<td><strong>1. Planning</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Setting goals</td>
<td>50</td>
<td>83</td>
</tr>
<tr>
<td>Setting criteria</td>
<td>75</td>
<td>17</td>
</tr>
<tr>
<td>Task purpose</td>
<td>0</td>
<td>67</td>
</tr>
<tr>
<td>Task</td>
<td>100</td>
<td>83</td>
</tr>
<tr>
<td>analysis</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Task demands</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>Setting a time line</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td><strong>2. Self-monitoring</strong></td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td><strong>3. Problem solving</strong></td>
<td>100</td>
<td>100</td>
</tr>
</tbody>
</table>

* % = percentage of students performing the procedures

An analysis of the learner background and learning context suggests two possible reasons for this generally high use of self-management procedures. The first possibility is that although the students were from different proficiency levels, they were all experienced adult learners of English. According to Hong-Nam and Leavell (2006), experienced language learners tend to use more strategies including metacognitive ones than less experienced learners. Another explanation could be that the learning environment might have an important influence on the students’ self-management. Because all of the participants were dependent on their EPP scores to get admission to academic studies at the university, they were likely to be much more motivated to prepare efficiently for the seminars to get satisfactory scores than, for example, students who might take a language course simply because it is required. Additionally, the students were preparing for the seminar as a part of the course and they had seminar homework to submit every week. To meet the course requirements, they were encouraged to conduct effective independent learning. Similar circumstances are also reported by Hong-Nam and Leavell (2006).

Second, the two groups both reported a limited amount of planning. Not all students from each group reported setting goals and when a student did, he/she did not report very specific and measureable goals that could guide his/her preparation. The same situation was found in setting criteria. Noticeably, none of the students from both groups reported making a time line for their preparation and all they did was
following the deadlines. The possible explanation for this finding is that most of the planning steps are internal processes. The steps might have been internalized and their deliberate attention was not necessary. As a result, the students would not report using the steps. It could also be possible that the students were not familiar with those planning steps, and thus paid limited attention to them. Another striking finding regarding planning is that none of the students from either group planned global strategies for the task. My observation revealed that the seminar program was scheduled according to weekly sub-topics such as formulating a seminar topic, using overhead transparencies, etc, and there were no explicit instructions and feedback on planning strategies and strategies needed to complete the task. This made the students rely on the schedule and did not go further than preparing for the weekly tasks. This could possibly be the reason why none of the students from either group planned global strategies for the seminar task. Further research to identify the causes of this limited amount of planning would have useful implications for teaching and learning.

Besides the similarities, the two groups also exhibited three striking differences in how they went through these procedures. First, although both groups reported a moderate amount of planning, the intermediate students tended to set goals and define task purposes more frequently than the advanced students (83% vs. 50% and 67% vs. 0% respectively). Meanwhile, setting criteria was found less frequently among the intermediate students (17%) than among the advanced group (75%). The observation reveals that the two groups had exactly the same seminar program and received almost exactly the same instructions. Therefore, it can be said that the role of instructions in causing the differences was strictly limited.

In seeking an explanation to the finding, two factors need to be considered in the present context. Because of the intermediate students’ lack of experience in preparing and giving a seminar (only two of them had previously given a seminar) and their lower proficiency, there might be quite a distance between the task requirements and their current ability. In order to fulfill the task and satisfy the proficiency proof requirement for admission to academic studies, they were likely to set goals and purposes to work towards. On the other hand, the advanced students, with their high proficiency and experience associated with seminars (three advanced students had experience in giving seminars and one had attended some), might find the task less challenging. They were also likely to have a better idea of what made a good seminar than their intermediate counterparts and therefore indicated it in their reports as criteria for assessing their performance.
Second, the advanced students reported that they monitored their progress much more frequently than the intermediate students (M=19 and M=11 respectively). This result supports the findings of Ortega (2005) where monitoring is more characteristic of high than low proficiency students. It is also consistent with the finding of Chamot, O’Malley, Kupper & Impink-Hernandez (1987) where the use of metacognitive strategies increased as the course level increased. It is also found that while the advanced students tended to identify the sources of their difficulties and seemed to be ready to solve the problems, most of the intermediate students were more inclined to notice their problems without reporting their causes.

Since both groups were instructed by the same seminar teacher and the teacher strictly followed her common lesson plans for both groups, these differences were minimally caused by the instruction. The fact that the advanced subjects were more active in setting criteria than the intermediate participants indicated that they were more prepared for monitoring their progress by self-evaluating their performance against the criteria. Another possible explanation is that because the advanced students reported better problem solving skills, they were likely to be more fluent in self-monitoring during their independent learning (Belfiore & Hornyak, 1998).

Third, the advanced level students reported a generally better ability to solve problems than the intermediate learners. A larger number of problems were identified and solved by the advanced students than the other group. The advanced students also took active control of their learning by initiating their own problem solving. In contrast, most of the intermediate students often depended on feedback from their teacher or friends in identifying their problems. In addition, the advanced students had a more systematic approach to problem solving than their intermediate counterparts. Once they noticed their problems, they moved on to seeking solutions and trying out the solutions. The intermediate students, on the other hand, tended to fall short in one or another of the procedures namely problem identification, problem solution or implementation of problem solution.

The observation data show that apart from the 4/3/2 activity during the final preparation week which was administered in the same way in both groups, the teacher gave no explicit instruction on problem solving procedure. Therefore, instruction was not an important factor that influenced the students’ problem solving and the differences between the two groups. As far as feedback is concerned, my observation shows that during the preparation period, the participants received some teacher and peers’ comments which helped them realize some problems regarding language, ideas,
structure, and presentation skills. The intermediate students received more comments because they had more problems than the intermediate students. However, it was found that they identified and solved fewer problems than their advanced counterparts. Thus, it can be said that the feedback was not important in distinguishing the two groups’ problem solving ability. Rather, according to Lan (1998), researchers have found that more self-monitoring can result in increased problem solving ability. If this is true, the advanced students’ reported greater ability to self-monitor their seminar preparation can possibly a reason why they were better at problem solving than the intermediate students. Another possibility is that the high level students with resort to their good repertoire of knowledge and strategies might have been able to locate and access resources to help them solve their problems better.

**Discussion and Implications**

The findings of the present study show that the subjects from both groups reported going through a range of self-management procedures similar to that described by Rubin’s (2005) model of learner self-management. However, the study found some new information about learners’ self-management procedures. First, while Rubin’s model separates monitoring from problem identification, the findings of the present study found an overlap between them. That is, problem identification was an aspect of monitoring. This can also be seen in Rubin’s description of monitoring as the process of learners identifying their problems. Therefore, it is recommended that subsequent research take into account the overlap of the two terms when categorizing self-management procedures. Second, the findings reveal that it is more reasonable to consider evaluation as another aspect of monitoring because in evaluating their performance, the students actually monitor their progress. The inclusion of evaluation and problem identification in monitoring is also supported by Butler (1997), Belfiore and Hornyak (1998) and Wenden (2001).

Another new insight is that while Rubin’s model separates problem identification and problem solution from implementation of problem solution, the present study found that in many cases the two procedures were not distinctly reported but the students actually went through both. Therefore, the study has presented findings about problem solution and implementation of problem solution in a broader category namely problem solving, which was carried out based on problem identification done during monitoring procedure.

The findings about the two proficiency groups’ self-management procedures also
reveal an interactive relationship between monitoring and problem solving procedures, including self-evaluation, problem identification, problem solution, and implementation of problem solution. This finding is congruous with Rubin’s characterization of the model of learner self-management. Also, it was proved that monitoring can trigger problem solving (Lan, 1998) and problem solving skills in turn can enhance monitoring of independent learning (Belfiore & Hornyak, 1998). However, the relationship was not clearly reported between planning and other procedures. This might be caused by the subjects’ lack of deliberate attention to the planning procedure which resulted in the students not returning to their plan while working on other procedures.

Interestingly, the present study shares some findings with previous research on short-term pre-task planning, including problem solving (Ortega, 1999, 2005) and self-monitoring (Ortega, 2005). More importantly, this study found that, with more extended preparation time and space, the students under research took advantages of the resources available outside the class in self-monitoring and problem solving. Additionally, the present study found that the participants also paid attention to setting goals, setting criteria to assess goal achievement, and analyzing the task.

Triangulated data sources generated results that have important implications for TEFL. Because all of the participants have EFL background, the majority of them are from Asian countries, and they had been in New Zealand for a very short time, they are likely to share certain traits, especially learning processes, with Asian EFL students. Following are some important implications put forwards for both the researched context and Asian EFL context.

First, according to Rubin (2005), setting goals, establishing criteria, analyzing the task, and making a time line are important steps in approaching a task. They help direct the students’ attention, understand the task better to make appropriate steps towards accomplishing it, and manage their time on the task. Because the students did not do these steps adequately in their seminar preparation, and they lacked global planning strategies, it is recommended that teachers train them in these steps and strategies. For example, teachers can ask the students guide questions and create a common planning checklist with them. Most importantly, time management is very essential for doing tasks. Hence, teachers should train the students in the habit of making a time frame for doing a task.

Second, because the students exhibited a limited ability to self-monitor and solve problems, the teacher should teach them the skills needed to do so and at the same
time raise their awareness of the importance of these procedures in independent learning. At an early stage the teachers may need to scaffold the students’ monitoring and problem solving and gradually withdraw the assistance when the students are ready to take up the procedures independently. For the advanced students, teachers should encourage them to maintain and further develop their self-monitoring and problem solving skills. This not only helps them better prepare for the upcoming ten minute seminar but also manage other learning tasks.

Third, since knowledge and beliefs are among the variables that may have great influence on the success of learner self-management (Rubin, 2005), class instructions should create opportunities for students to learn the knowledge needed to do tasks and develop facilitative beliefs. Task knowledge and strategy knowledge are of great importance and can be promoted by modeling task analysis and instructing strategies.

With respect to the relationship between L2 proficiency and strategy use, the study reveals that the higher level students qualitatively and quantitatively performed greater monitoring and problem solving strategies than their intermediate counterparts. This is congruous with that of relevant research in different contexts (Chamot, O’Malley, Kupper & Impink-Hernandez, 1987; Green & Oxford, 1995; and Griffiths, 2003) and confirms the positive relationship between proficiency and strategy use. The finding, therefore, can be extended to other contexts and suggests that it is important for EFL teachers, especially those of low proficiency classes, to raise students’ awareness of the availability and effectiveness of those strategies and facilitate the use of them.

Because the study did not investigate the relationship between the students’ use of self-management procedures and their seminar performance, it is recommended that future research goes further to investigate the influence of reported self-management procedures, knowledge and beliefs in task preparation on task performance. In addition, although knowledge and beliefs are an integral part of learner self-management (Rubin, 2005), this study did not elicit the different kinds of knowledge and beliefs used by the subjects in their self-management. Subsequent studies that take account of both the procedures and the kinds of knowledge and beliefs that students use in their self-management may gain even more insights into the learners’ self-management and therefore could yield results that have further implications for EFL learning and teaching. Finally, it is recommended that variables that may possibly influence the students’ self-management such as instructions, context, nationalities and cultures should be studied in future research to ascertain the
In summary, this study investigated the learner self-management procedures that advanced and intermediate ESL students from a New Zealand university used in their preparation for a five minute seminar as part of their English Proficiency Program. Findings from triangulated data sources namely learners’ diaries, interviews, and observation revealed that all of the students from both proficiency groups reported going through a range of self-management procedures including planning, self-monitoring, and problem solving. However, both groups did a limited amount of planning, and revealed different focuses. Moreover, the advanced students monitored their preparation more often and were better at problem solving than their intermediate counterparts. These findings have important implications for the studied classrooms and can be extended to other contexts.

References


Appendix A:

Diary guidelines

What is a language learning diary?
A language learning diary (or journal) is a record of your daily learning experience. Everyday you spend some time thinking about your own learning experience and write your thoughts and feelings about it down in a diary.

Why should I keep a language learning diary?
There are important reasons for keeping a language learning diary:
1. It encourages you to reflect on your learning experience, evaluate your strengths and weaknesses, find out what works for you and what does not, and make improvements based on your day by day evaluation of your own learning.
2. It helps other people such as your friends and teachers understand your strengths and weaknesses in language learning and help you make improvements if it is necessary.
3. It gives you practice in writing English.

What should I write in my diary?
Each day, spend at least 15 minutes writing about what you have done during the day to prepare for the seminar, your thoughts and feelings about your seminar and about what you did for it. Following are some questions you might ask yourselves:
1. What thoughts and feelings do I have about the presentation?
2. What do I prepare in my mind?
3. What action did I take today to prepare for my seminar?
4. What do I need to do next to prepare for my seminar?

Note: These questions are just guidelines. You might want to write beyond answering them. You might write about anything relevant to your preparation for the seminar. Do not worry about making spelling, grammatical, or organizational mistakes in your diary. Your diaries will NOT be assessed in any way. What we need is information about how you prepare for your presentation which can be used to help you and other students learn better. You can write merely in English or in you first language or you can write using both languages.
The Effect of Collaboration on the Cohesion and Coherence of L2 Narrative Discourse between English NS and Korean L2 English Users

Peter Crosthwaite
University of Cambridge
prc34@cam.ac.uk

Bio Data

Peter Crosthwaite is currently a doctoral student at the Department of Theoretical and Applied Linguistics (DTAL), University of Cambridge. He has also taught English for 7 years in South Korea, teaching at Pukyong and Dong-Seo universities. He is interested in cohesion and coherence of spoken and written discourse and is currently investigating Korean and Chinese L2 English discourse for referential typology and L2 acquisition sequences. He is also interested in language assessment, is a qualified IELTS examiner, and has been director of studies for EF language schools in Cambridge, Clare College.

Abstract

This research looks at differences between how native speakers of English and Korean L2 English learners manage cohesive reference maintenance, as well as the effect of scaffolded interlocutor collaboration on the coherence and cohesion of extended L2 narrative discourse. Scaffolded and unscaffolded narratives were elicited from 10 Korean learners of English as an L2 and were compared against the narratives of 5 native speakers of English, to compare the grammatical means used to maintain coherent reference to discourse referents within and across clauses, as well as to see the effect that any scaffolding had on the L2 participant’s ability to maintain coherence during performance. A link was found between the coherence of NS narrative discourse and accurate use of co-referential & distant anaphoric grammatical referential devices, and the presence of scaffolding was found to increase the accuracy of non-native speakers’ use of these devices. The implication of these results is that scaffolding helps L2 learners to create and hold more accurate reference to discourse referents, and instances of unscaffolded narrative discourse present increased difficulty for the L2 speaker. Finally, as L2 learners have more difficulty managing accurate reference maintenance, the overall coherence of their discourse is reduced.

Key words: Coherence, Cohesion, Reference, Narratives, Discourse, Korean.
Introduction
According to Cribb (2003), ‘further research is needed into discourse coherence in non-native speech’ (2003, p. 464). This research focuses on adult L2 English learners’ production of coherent and cohesive oral narratives. In particular, the effect of collaboration or scaffolding by native speaker interlocutors during the L2 learners’ production of the narratives will be a central variable of this research. It is hoped that by focusing on the difficulties faced by learners when producing extended non-scaffolded narrative discourse, a clear picture might be gained of the common L2 language learners’ experience of producing such complex linguistic events and the coherence and cohesion inherent within.

Review: Cohesion and Coherence in English

Cohesive devices and their role in coherence
Halliday & Hasan (1976) gave one of the first comprehensive overviews of cohesion, defining cohesion as something that occurs ‘where the interpretation of some element in the discourse is dependent on that of another’ (1976, p. 4). Cohesion is then split by Halliday & Hasan into two categories, that of grammatical cohesion, and lexical cohesion, with grammatical cohesion including devices for reference (personals, demonstratives and comparatives), and lexical cohesion including devices for reiteration & substitution (after the match / after the game). Accurate use of these cohesive forms is one of the pre-requisites for coherent discourse, along with the need to maintain a clear sequence of temporality, aspect and causation. From this point onwards, this research will be concerned with referential cohesion and coherence – cohesive devices that refer to discourse entities.

Of interest to L2 discourse cohesion, the production of coherent and cohesive discourse is a feat considered difficult for second language learners, as shown in von Stutterheim (2003) who found that even advanced learners still have problems in applying cohesive forms in context. The primary reason for this difficulty is that while discourse pragmatic principles such as the marking of information in discourse are universal (principles such as the ‘given/new’ hypothesis, where discourse referents are introduced/maintained in discourse), ‘the devices available to mark the relevant distinctions differ across languages’ (Hickmann & Hendriks, 1999, pp. 419-420).
Tanskanen (2006) also notes that ‘cohesion may not work in absolutely identical ways in all languages, but the strategies of forming cohesive relations seem to display considerable similarity across languages (2006, p. 38),’ using the examples of Enkvist (1975) in Swedish and Danes (1987) in Czech. Researchers of discourse cohesion and coherence therefore find it likely that typological differences between how different languages handle L1 cohesion may cause difficulty for instances of L2 cohesion.

Therefore, when considering the context of EFL, the coherence of any non-native discourse in English is likely to depend on the L2 users’ accurate management of English referential cohesive devices, the grammatical means of which will be the focus of this study.

**What makes ‘coherent’ referential discourse?**

Givon (1995) defines coherence as ‘the continuity or recurrence of some element(s) across a span (or spans) of text (1995, p. 61).’ Given the possibility of variation in cohesive marking grammatically between different languages, an overall framework for comparing the differences in reference maintenance between any source language and English can be taken from Givon (1995, p. 71) from his comparison of cohesive devices that signal *continuity/discontinuity* (or grounding) of referents in discourse. For cataphoric grounding in English (where new referents are identified as those that will be ‘important, topical and thus persistent in the subsequent discourse’) (Givon, 1995, p. 65), indefinites are used. The indefinite articles (‘a/an’) and determiners such as ‘this’ are used in English to mark indefiniteness. For anaphoric grounding (where the referent is ‘retrieved’ from the mental discourse structure) (Givon, 1995, p. 68), a definite expression would be used, such as a pronominal form, or a full NPs with the definite article ‘the’, as in English.

In English, the common pattern of grounding across co-referential clauses (where the antecedent of the cohesive device is found in the same or previous clause) would generally be of the form indefinite to pronoun for characters that have just been introduced into the discourse. For example:

‘A man entered. He went upstairs.’
Zero anaphora (the element of cohesion with the most activation) is typically only used between clauses with an additive conjunction in English or ‘lists’ of clauses where the referent has not changed and keeps the same semantic role:

‘A man entered and (ø) went upstairs.’

‘He turned, ø looked, ø screamed and ran’

Discontinuous or ‘distant’ reference (where the antecedent of the cohesive device is further ‘back’ than a single clause) might occur when a new referent is introduced into the discourse, and the initial referent must be ‘reactivated’ through a definite expression, marked by the definite article ‘the’ in this example:

‘A man entered. A woman came in. The man said ‘hello’.’

Languages have a sliding scale of cohesive devices for discourse continuity (Gundel et al. 1993, Givon, 1995, Ariel, 2008) with ‘zero’ anaphora being the most ‘continuous’ method to refer to a referent, followed by pronouns (with unstressed pronouns considered more ‘continuous’ than stressed pronouns), then followed by nouns with definite articles, and finishing with full lexical nouns (including modifiers) respectively. Referential access of this kind can also work on a ‘frame based’ approach where our pragmatic world knowledge can come into play when reference is made, as with ‘part-whole’ or ‘possessor-possessed’ relations (ex: the house was a mess, the roof leaked – for this reference to be accessible, we should know that houses have roofs). These are known as ‘bridging descriptions’ (Clark, 1977) but are special cases of reference that generally go against the scales above and will not be discussed further in this paper.

The continuity of reference within and across clauses is highlighted as a way of measuring overall coherence, achieved through the accurate and appropriate management of co-referential and non-co-referential (distant) cohesive devices within a text. When accurately managed, chains of reference within the text will be properly maintained for the listener, who will be able to correctly follow the flow of information through use of the appropriate devices for retrieval of referential information.

Following Hickmann & Hendriks’ (1999) methodology, a suitable method for the observation of cohesive reference maintenance and subsequent coherence is
through the elicitation of narratives, a kind of discourse to which we will now turn our attention.

*The Importance of Collaboration for Coherent and Cohesive Narratives*

Narrative discourse is a complex verbal task that is perfect for the analysis of linguistic reference maintenance. Labov and Waletsky (1967) suggest that narratives contain a *referential* function that needs to be fixed in time according to *the principle of natural order*. Barthes (1977) also suggests that narratives contain a referential function, which is ‘a seed that is sown in the narrative, planting an element that will come to fruition later – either on the same level or elsewhere on another level (1977, p. 89).’ Through observing L2 learners’ narrative production, we can get a clear picture of how an L2 learner maintains this referential function over discourse. This approach is validated by Loschky and Bley-Vroman (1993) when looking at how languages approach cohesive article use, with their recommendation that “in creating tasks for developing knowledge of articles, the task designer … should consider using narrative tasks for the definite/indefinite distinction” (p. 133).

However, of importance to this research, Reismann (1993) claims Labov’s (and others) model ‘leaves out the relationship of teller and listener […] a teller has a fundamental problem: how to convince a listener who was not there that something important happened (1993, p. 20). Solving this problem may require something more than the linguistic skills of the speaker – fundamentally, a second party may well be involved in the production of the narrative. This *collaborative* aspect of narrative-making is of vital importance to the coherence of the finished product, and this aspect is the primary focus of this research.

As mentioned, researchers interested in cohesion and coherence view coherence as more than a linguistic ‘text’, in that a fully coherent text is a *collaborative* negotiation ‘for the common ground of shared topicality, reference and thematic structure – thus toward a similar mental representation (Gernsbacher & Givon, 1995, p.vii).’ This is also touched upon in Clark (1996) who mentions that language use ‘is really a form of joint action (1996, p. 3).’ Tanskanen (2006) notes that: ‘there are still notable gaps in our understanding of the effects on the use of
cohesion of the different contexts in which speaker, writers, listeners and readers operate and communicate (Tanskanen, 2006, p. 2).’

For narratives, Clark & Wilkes Gibbs (1986) show that interactions between speaker and listener normally involve both participants and can take several turns to accomplish. Goodwin (1995), working under the framework of turn-taking proposed by Sacks, Schegloff & Jefferson (1974), shows the common turn-taking pattern taken when a narrative is to be performed in Fig.1. (The blank lines intended to represent the telling of the extended narrative).

\[\begin{array}{|c|c|}
\hline
\text{Teller:} & \text{Narrative Preface} \\
\hline
\text{Recipient:} & \text{Request to hear narrative} \\
\hline
\text{Teller:} & \text{Narrative} \\
\hline
\ldots & \ldots \\
\hline
\text{Recipient:} & \text{Response to Narrative} \\
\hline
\end{array}\]

*Fig.1. Narrative Turn-Taking Structure. From Goodwin (1995, p. 126)*

Taking this further, Goodwin is quick to note that ‘processes of interaction within the turn at talk have strong consequences for the flexible organisation and maintenance of coherence on a number of different levels.’ (1995, p. 122). To Gibbs (1995), this comes back to the concept of grounding mentioned previously in terms of the collaborative process, in that ‘in conversation, the process of grounding a contribution divides conceptually into presentation and acceptance phases [...] in which the participants look for evidence that they have satisfactory mutual interpretation of the action. (1995, p. 244).

Pellegrini and Galda (1990, pp. 118-120), while observing experimenters who were asking children to perform narratives, devised a very extensive list of measures used by interlocutors to scaffold the narrative process. Examples of such measures include asking for extensions, reinforcements, role clarifications, evaluations and
even reprimands, all made by the recipients during the child’s performance of the narratives. Linell (1998) calls this collaborative process ‘dialogism’ or ‘individuals in dialogue with partners and contexts’ (1998, p. 8). Tanskanen (2006) also proposes that ‘collaboration can be realised for example as feedback between participants in the form of completions, clarifying questions, or other types of acknowledging that the participants have understood what their fellow communicators were saying. (2006, p. 24).

Thus, an experiment where the kind of interlocutor interactions suggested by Pellegrini & Galda (1990) are controlled for may shed light on the contribution such interactions make to the maintenance of cohesive devices and the subsequent coherence of a linguistic text as defined above. This leads the researcher to pose the following research questions:

**Research Questions**

From the literature reviewed above, this research will seek to answer the following questions related to the creation of coherence through cohesive devices, as well as the effect of collaboration on cohesion and coherence.

1) What kind of devices will the non-native speakers employ for cohesion and coherence during their performance of the narratives, compared to those of the native speakers?

2) What kind of variation can we find in the use of cohesion between scaffolded and unscaffolded instances of narration within and between users?

3) What is the relationship between the use of co-referential reference maintenance and distant non-co-referential reference maintenance between native and non-native speaker groups? How does this relationship affect the coherence of the text?

**Hypotheses**

In terms of grammatical referential cohesion, the native English speakers are likely to introduce characters into discourse using indefinite articles.
A man went into a house.

For definite co-referential reference, they are likely to use personal pronouns (such as he, she, they etc.) between adjacent clauses to describe the actions of the main protagonists in each narrative (as in the first example below), and for longer distance non-co-referential chains, are likely to re-introduce the referent through full NPs with definite articles (example 2 below, and following the findings of Hickmann & Hendriks, 1999).

Ex: 1) A man went into a house. He went upstairs, and then he ate dinner.
   2) A man went into a house. A butler was there. The man ate dinner (not the butler).

In Korean, which lacks a grammatical article system, new mentions may take the numeral form han (in a use similar to the English numeral ‘one’ e.g. ‘one man’) as with the example below:

(New Mention)

Ex: Han-namja ga chib-uro gatda.
   One man house into went

Korean has personal pronouns marked for gender, but commonly for referents in co-referential contexts (where the topic of the reference has not been replaced by another, additional referent), zero anaphora are normally used. In terms of the preference for zero anaphora in Korean, this is related to the existence of the pro-drop/topic drop parameter for anaphora, and the positive setting of this parameter may have consequences on the coherence of referents in topic/subject/object position. Korean is known as a ‘pro-drop’ language, where reference to entities in certain discourse contexts are omissible when pragmatically inferable, as in Korean ‘it is stylistically more natural not to explicitly mention anaphors in subordinate clauses that are co-referential with nominal expressions in the main clause’ (Mitkov, Kim, Lee & Choi, 1994, p. 23; see also Huang, 1984). Subject relationship within a sentence in Korean is determined by a suffix on the noun (Namja-Ga).

1) (Co-referential zero anaphora)

Ex: Namja-ga chib-uro gatda. Ø oui-chung-uro ola-ga jonyok-ul mokkotda
   Man house into went (zero) upstairs went (zero) dinner ate

142
Where Korean differs from English is that zero anaphora may be used *between* sentences to refer as with the example above, while in English, it may only be used within a sentence after a conjunctive while the same referent is in subject position or as part of a ‘list’ of clauses of the same type (he looked, ø turned, ø screamed and ran). The discourse pragmatic distinction between the use of the subject marker vs. use of the topic marker are complex and lie outside the scope of this paper, yet it is enough to say at this stage that either subjects or topics may be omitted in Korean when pragmatically inferable. A discussion on this phenomenon can be found in Huang (2000), who claims that ‘there are some grounds for believing that in a pragmatic language like Chinese, Japanese or Korean, when syntax and world knowledge clash, world knowledge frequently wins. By way of contrast, in a syntactic language like English, French and German, there is a conflict between syntax and world knowledge (2000, p. 265).’

Distant definite co-referential expressions are sometimes maintained by the use of the demonstratives ‘that’ (*ku*) for the distant mentions given the lack of definite articles. This use can occur if the correct spatial relationship exists between the speaker and the referent, but this marking is optional rather than acting as an obligatory definiteness marker (as occurs in English).

(Distant mention)

Ex: **Han-namja-ga chib-uro gatda Han Yeoja-ga kogi-ae issotda.** **Ku-namja jonyok-ul mokkotda**

*One man house into went. One woman was there. That man ate dinner.*

An additional complication for Korean L2 English learners may come from the complex honorific system for reference to person in Korean. A great deal of variety in the use of referring expressions in Korean is determined by the speaker’s relationship to the discourse subject in question as well as the speaker’s relationship to his audience. This is commonly evidenced in verb suffixes, with as many as 6 methods to express the English sentence ‘John hit Bill’ depending on whether the register is formal, polite, blunt, familiar, intimate or plain (Chang, 1982), but may also take nominal form (as with the pronominal forms for ‘I’, ‘*na*’ & ‘*cho*’, with ‘*na*’ being the common form, and ‘*cho*’ the respectful form when in the company of someone of a higher social status). Such a complex referential system is not found in English, and in discourse with multiple referents of different ages or social positions, there may be confusion for Korean learners about how to label these referents in L2 English.
Given the clear differences between cohesive reference maintenance between English and Korean, one would expect that the Korean learners might struggle in terms of maintaining the main protagonists of each narrative in discourse in the same manner as that of native speakers, even though the discourse pragmatic principles the speaker utilises to do so are universal (Hickmann & Hendriks, 1999, pp. 419-420).

At lower proficiencies, according to Givon (1995), the NNSs at early/beginner stages of English acquisition will take the form of a ‘slow, analytic pre-grammatical mode of discourse processing [that] is heavily vocabulary driven’ (Givon, 1995, p. 78). For Givon, grammatical forms of coherence ‘evolved as a mechanism for speeding up the processing’ (Givon, 1995, p. 78) of coherence, and what this would translate to in this study should be the use of vocabulary-driven reference in place of grammatical-driven reference in lower-level learners, who have not yet fully acquired the grammatical means to accurately maintain reference. This performance is not to be confused with lexical cohesion, as lower level learners would not likely have access to the amount of synonyms needed to perform the kind of substitution necessary for true lexical cohesion. What is more likely to occur is that inaccurate bare forms of referential phrase (where the grammatical marker is missing – ex: A man went into a house, man went upstairs) are likely to be found in the NNS data, and what is more, that the use of these bare forms will only lower the overall coherence of the narrative, as accurate reference maintenance will suffer from the lack of definiteness marking. At intermediate levels, over suppliance of articles and inappropriate pronominal mentions are expected, eventually falling in line with native speaker norms at higher levels following a ‘u’ shaped pattern of development.

In addition, due to the lack of linguistic means available to maintain a narrative in a second language, coupled with the added cognitive strain of doing so, it is expected that while NSs will be able to maintain reference to the main protagonists in each narrative accurately, the NNSs may jump from referent to referent depending on what they are able to comment on as they perform the narratives (a bottom-up approach). L2 coherence is divided by Anderson (1995) into global and local coherence within a text, where ‘there is [generally] local consistency but global inconsistency and where the text is produced in a bottom-up unplanned manner with flexible, shifting and negotiated perspectives’ (1995, p. 2). By ‘bottom-up’, we mean that L2 learners (especially at lower proficiency levels), have difficulty drawing on contextual cues or world knowledge to aid them in the processing of text (a
top-down approach), and instead can only focus on building text one-word-at-a-time, paying close attention to grammatical and phonological accuracy. This pattern was also characterised in Carroll & von Stutterheim (2003), who found that L2 learners of lower proficiencies may ‘run into trouble since they do not construe sets of events as larger units that are linked’ (Carroll & von Stutterheim, 2003, p. 394). As proficiency improves, Karmilloff-Smith (1981) found that English (and French) speakers use a top-down cognitive approach to personal reference for narratives, organised around a central character, after having initially started out with a bottom-up strategy for achieving this kind of reference. This discontinuity is likely to be characterised by the existence of errors, including the incorrect use of the indefinite article to signal existing referents within the narrative (thus creating redundant ‘new’ referents that should not have been introduced into the events of the narrative), or incorrect uses of pronouns/demonstratives/zero forms that will create ambiguity for the listener as to who is being referred to.

Method

Participants

10 NNS participants were selected for this study from a private university in Pusan, South Korea in 2010. The participants were all freshman college students of 21 years of age who volunteered to participate without payment for the purposes of the research. They were all native Korean speakers learning English as a foreign language. The students were either English majors or were majoring in travel & tourism degrees that require a degree of English ability to complete, and had TOEIC scores of just above or below 250, having not taken any other standardized tests where their proficiency could be measured such as TOEFL or IELTS.

An additional 5 NS participants were selected to provide the NS data against which the NNS data would be compared, and 2 of them were from the U.K., 2 from the U.S.A., and one from New Zealand. They are all teachers of English at the university where this research was carried out, and all in their 30’s.
Stimuli

The stimuli for this experiment were two picture sequences taken from a popular comic series. The picture sequences were controlled for the length/topic of the narratives so that the results could be generalized among the spread of participants’ responses. These particular pictures were chosen as they are meant to be read in sequence as a coherent narrative in their natural context. The sequences were edited to remove all character speech (except !, ? symbols) from the speech bubbles present in the pictures so that the narrators would avoid falling into ‘reported speech’ while telling the narratives, as it is possible that the narrator would shift strategies for reference depending on which perspective they took, as the finding of Carroll & von Stutterheim (2003) suggests that even advanced learners ‘face a problem at the level of perspective taking […] where the basis for the inappropriate use of certain linguistic forms lies (2003, p. 393).’ The participants were informed in the instructions for the task that they did not have to provide speech for the characters but were not explicitly told not to do so, allowing the participants the option to do so if they wished.

The narrative sequences were pre-tested for length on two native Korean speakers (who had IELTS 6.5 proficiency – ‘competent’ users). Each speaker took narrative 1 first, with one speaker allowed scaffolding, and the other allowed scaffolding only on the second narrative. Narrative 1 (unscaffolded) took 2:25 to complete on the pre-test, and Narrative 1 (scaffolded) took 2:52 to complete with two instances of scaffolding from a NS. Narrative 2 (unscaffolded) took 2:50 to complete, and narrative 2 (scaffolded) took 2:44 including three instances of scaffolding from a NS.

The picture sequences used were taken from the animated books Tintin in America (Herge, 1932) and Tintin and the Seven Crystal Balls (Herge, 1948). As the order of the vignettes was changed from the originals in the books, as well as due to the modification of the images, the publisher (Moulinsart, France), were not able to give permission to reproduce the images in this publication, and therefore written descriptions of the picture sequences will be provided in appendices 1 & 2.
Procedure

The participants were invited into the experiment room after regular class hours at a time arranged with the participants’ co-operation. The participants were encouraged not to discuss with other classmates any information about the experiment in order to avoid revealing the nature of the picture sequences before elicitation. The picture sequences were stored on Powerpoint slides, and could change the slides to show the next/previous images in the sequence at any time during the experiment so that the students’ ability to recall information would not be a factor in the experimental design. The full instructions to participants are shown in appendix 3. Participants were allowed some time to study the picture sequences before starting the narratives in order to reduce the cognitive load on the participant from retelling previously un-experienced events. They were given around three minutes to do this by the examiner but were not told in advance how much time they were to be given to avoid the pressure of time. They were also not told how long they should be narrating for, again to avoid any pressure from time constraints. The students did not have to make a comment on every picture in the sequence, although they were free to do so if they wished. The instructions to candidates were provided in English with accompanying Korean translation to ensure the participants’ full understanding and co-operation. Identical conditions were imposed on the NSs’ performance of the narratives.

In terms of the scaffolding used, a list of our interlocutor’s permissible interactions is included in appendix 4, and is taken from Wilkes-Gibbs (1995) and Pellegrini and Galda (1990). In total, there were 143 instances of scaffolding for narrative 1 (avg. 28.6 per narrative) and 92 for narrative 2 (avg. 17.6 per narrative), which was not significantly different (F=4.812, P=0.060).

After the data had been collected, five other NSs (not previously used in the study) were selected to analyse the coherence of the NNS speakers’ narratives on a 10-point Likert scale, with a score of 1 being described as ‘totally incoherent’ and a score of 10 described as ‘totally coherent’. These scores were then collected and attempts were made to correlate the overall coherence rating of each narrative given in the Likert scale to the use of cohesive devices found in the narratives to see if any distinctions could be drawn about the use of these devices and the coherence of the narrative to a NS.
Analysis

Transcriptions and word counts were made of each narrative. Initially, there was concern that the NS narratives may have been substantially longer than those of the NNS narratives, but a one-way ANOVA of the word lengths between the groups did not show a significant difference (F= 1.548, P <0.252, mean = 202.6 words). For narrative 1, seven obligatory animate referents were selected (Tintin, Snowy the dog, the butler, the cat, Captain Haddock, a Mexican knife thrower and his Indian participant), and for narrative 2, six were selected (Tintin, Snowy the Dog, the taxi driver, the police, the man with the boomerang, and the doctors in the ambulance). Subsequent mentions of these referents within the narratives were noted and totalled. These mentions were coded as pronominals (zero, personal, relative), or nominals (bare, definite, demonstrative, possessive, indefinite) as with Hickmann & Hendriks (1999). Repetitions within the same clause by the NNS were not included in the total counts (as the NNSs tended to repeat themselves when a gap in fluency occurred) and direct repetitions made by the NNS as a result of the NS interlocutor mentioning that referent directly (as with the examples below) were also not included as these forms could have been considered as putting words directly into the participants’ mouths. (I=Interlocutor, P= Participant):

Ex: I-he opens the door? P-he opens the door and Kevin rides the taxi [pause]

Ex: I-so who is he? What does he look like? An Entertainer? P-Yeah an entertainer, and he can throw?

Relative pronominals and demonstrative nominals were scarcely found in the data (only four instances between NS and NNS data each) and so were not included in the final analysis. This was surprising given the Korean tendency to use demonstratives to refer to given referents across adjoining sentences or longer distant anaphora as suggested in the hypothesis. I account for this trend post-poc by referring to Kang (2009) who claims that even at low levels, Korean learners were ‘aware’ of the typological differences regarding encoding of reference between their L1 and L2 English due to their classroom based learning experiences, and therefore made attempts to follow the referential strategy of the target language. A higher (yet still small) use of possessive nominals were found in the data, but ANOVA between the seven referents in narrative 1 showed their difference in use between NS and NNS data was not significant (F=0.62. P<0.242). A similar pattern of use was
found in narrative 2 for possessive nominals. The use of possessive nominals is quite different between within-same-clause co-reference (syntactic) and between adjacent clause co-reference (discourse-based), therefore these forms will not be discussed further in this paper. The remaining factors to be analysed in this research were those of zero anaphora, personal pronouns and bare, definite, and indefinite nominals.

Each reference was additionally coded for anaphoric ‘distance’ with coding for ‘new’ referents (likely to have been made in error, linked semantically to a previously mentioned referent), as well as coding for co-referential expressions (where a referent in clause X was mentioned in clause X or X-1), and coding for non-co-referential expressions or ‘distant’ reference (where a referent in clause X was mentioned in clause X-2+). Examples of each kind of coding can be seen in table 1 below:

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>P-The man finds a cat (initial mention)</td>
<td>P-One man (initial mention) has a pet and they (co-referential) go to a party, he (co-referential) said ‘where is the master’ to the waiter</td>
<td>P- A man (initial mention) take a tram? With his dog and arrived safe, he take a taxi?</td>
</tr>
<tr>
<td>I-OK, so the man goes to find the cat and [pause]</td>
<td>I-OK</td>
<td>I-Taxi</td>
</tr>
<tr>
<td>P-Go house […] a cat fight (subsequent mention of previously introduced referent with indefinite article)</td>
<td>P-And his (co-referential) dog meet another dog</td>
<td>P-Taxi with his dog. And the driver has blind down</td>
</tr>
<tr>
<td></td>
<td></td>
<td>I-He closes the blinds?</td>
</tr>
<tr>
<td></td>
<td></td>
<td>P-yeah he closes the blinds. The man (distant reference) is very nervous and he arrived strange space.</td>
</tr>
</tbody>
</table>

Table 1: Examples of coding for references

ANOVA were used to find the differences between the use of grammatical cohesive devices between NS and NNS data for narrative 1 and narrative 2, then additional ANOVA were performed to see whether the use of these devices was any different depending on whether scaffolding was provided for the NNSs for the first or second narrative. Repeated measures ANOVA was also used to measure any variation in anaphoric distance between NS and NNS (scaffolded or unscaffolded) data, so that the hypothesis regarding the co-referential/non-co-referential grounding of the main protagonists between the NS and NNS data could be tested.

Finally, the results of a 10-point Likert scale coherency judgement task performed by an additional five NSs (NSs who were not used in the experiments) on the NNS narratives were collated to see whether there was any correlation between
the perceived coherence of the narrative by a NS and the use of cohesive devices for reference within that narrative, using a correlation matrix determined using SPSS.

**Results**

All data was collected in line with the methods of analysis mentioned above. With reference to the 3 research questions listed above:

1) What kind of devices will the non-native speakers employ for cohesion and coherence during their performance of the narratives compared to those of the native speakers?

With the exception of relative pronominals, demonstrative nominals and possessive nominals (due to their lack of use in the data as described above), the following is a table of the grammatical cohesive devices used to maintain reference to the animate referents chosen for each narrative. Table 2 shows the references used in narrative 1 and table 3 shows the references used in narrative 2.

<table>
<thead>
<tr>
<th></th>
<th>NS data (n=5)(total = 182 references)</th>
<th>NNS data (n=10)(total 233 references)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Zero</td>
<td>23 (12.64%)</td>
<td>5 (2.15%)</td>
</tr>
<tr>
<td>Personal</td>
<td>76 (41.76%)</td>
<td>71 (30.48%)</td>
</tr>
<tr>
<td>Bare</td>
<td>0 (0%)</td>
<td>58 (24.89%)</td>
</tr>
<tr>
<td>Definite</td>
<td>82 (45.05%)</td>
<td>79 (33.90%)</td>
</tr>
<tr>
<td>Indefinite</td>
<td>1 (0.55%)</td>
<td>20 (8.58%)</td>
</tr>
</tbody>
</table>

*Table 2 – Forms used to maintain reference in Narrative 1*

<table>
<thead>
<tr>
<th></th>
<th>NS data (n=5)(total = 228 references)</th>
<th>NNS data (n=10)(total 336 references)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Zero</td>
<td>16 (7.01%)</td>
<td>11 (3.27%)</td>
</tr>
<tr>
<td>Personal</td>
<td>116 (50.87%)</td>
<td>143 (42.55%)</td>
</tr>
<tr>
<td>Bare</td>
<td>0 (0%)</td>
<td>69 (20.53%)</td>
</tr>
<tr>
<td>Definite</td>
<td>95 (41.66%)</td>
<td>96 (28.57%)</td>
</tr>
<tr>
<td>Indefinite</td>
<td>0 (0%)</td>
<td>15 (4.46%)</td>
</tr>
</tbody>
</table>

*Table 3 – Forms used to maintain reference in Narrative 2*

For narrative 1, an ANOVA of the use of zero pronominal reference between NS and NNS groups for stories 1 and 2 showed significant differences (narrative 1
F=26.297, P<0.001, narrative 2 F= 9.512, P= 0.009), suggesting that NSs made much more use of zero pronominal reference than the NNSs during their retelling of the narratives, which was surprising considering the hypothesis made regarding the use of zero anaphora in Korean as described above, given that the pro-drop parameter is set to + in Korean and – in English. This may be explained by the finding of Kang (2009) as given above, that the Koreans even at lower levels of proficiency were aware of the typological differences between their L1 and the target language, and so were careful not to use zero anaphora even in contexts where it is acceptable (within clauses, as found in English).

For the use of personal pronominal reference, a significant difference between NS and NNS groups was also found for Narrative 1 (F=18.992, P<0.001) but not for narrative 2 (F=2.684, P=0.125). For the use of definite reference, a significant difference was found for narrative 1 (F=8.151, P<0.014) but again not for narrative 2 (F=3.645, P=0.79). For the use of bare forms, the NSs did not use any bare forms at all, and when the use of bare forms by NNSs is taken into account, the difference is still statistically significant for both stories (Narrative 1 F=8.208, P=0.013, narrative 2 F=4.744, P=0.048). As bare forms are not found in English, their use in NNS data is harmful to the coherence of the overall narrative due to the missing definiteness marking requirement typically needed by NSs.

The NSs used indefinite forms to refer to a previously given referent only once, with the NNSs using this form (as with the bare forms) in error more often. However, for narrative 1, the use of these forms between NS and NNS was not deemed to be significantly different (F=2.764, P=0.120), with a similar result for narrative 2 (F=3.421, P=0.087). This is interesting as it suggests that despite the relatively low English abilities of the NNSs that participated in this experiment, the incorrect use of the indefinite article to reference previously given referents did not happen as often as was previously expected (again, see Kang 2009).

2) What kind of variation can we find in the use of grammatical anaphoric cohesion between scaffolded and unscaffolded instances of narration within and between users?
The difference between the word counts for scaffolded and unscaffolded forms was not seen to be significantly different (F=0.320, P=0.578). Table 4 below shows data on the use of grammatical cohesive devices between scaffolded and unscaffolded instances of narratives:

<table>
<thead>
<tr>
<th></th>
<th>Scaffolded NNS data (N=10) (Total 324 references)</th>
<th>Un scaffolded NNS data (N=10) (Total 243 references)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Zero</td>
<td>10 (3.08%)</td>
<td>6 (2.46%)</td>
</tr>
<tr>
<td>Personal</td>
<td>135 (41.66%)</td>
<td>79 (32.51%)</td>
</tr>
<tr>
<td>Bare</td>
<td>42 (12.96%)</td>
<td>55 (22.63%)</td>
</tr>
<tr>
<td>Definite</td>
<td>120 (37.03%)</td>
<td>85 (34.97%)</td>
</tr>
<tr>
<td>Indefinite</td>
<td>17 (5.24%)</td>
<td>18 (7.40%)</td>
</tr>
</tbody>
</table>

Table 4 – Forms used to maintain reference to the animate objects between scaffolded and unscaffolded stories within groups.

The difference between scaffolded and unscaffolded stories in terms of zero anaphora, was not calculated as significant (F=1.455, P=0.262). The use of personal pronominal forms was also not calculated as significant (F=3.333, P =0.105). The use of bare forms was also not calculated as significant (F=2.664, P = 0.141). A similar pattern emerged for definite forms (F=3.634, P=0.093), as well as for indefinite forms (F=0.008, P=0.932).

Table 5 shows the data for the use of grammatical cohesive devices between those who took narrative 1 with scaffolding, against those who did not receive any scaffolding, while table 6 shows the data for the use of these devices between those who took narrative 2 with scaffolding against those who took narrative 2 unscaffolded.
Asian EFL Journal. Volume 13, Issue 4

<table>
<thead>
<tr>
<th></th>
<th>Scaffolded NNS narrative 1 (N=5) (Total 126 references)</th>
<th>Unscaffolded NNS narrative (N=5) (Total 107 references)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Zero</td>
<td>2 (1.58%)</td>
<td>3 (2.80%)</td>
</tr>
<tr>
<td>Personal</td>
<td>44 (34.92%)</td>
<td>27 (25.23%)</td>
</tr>
<tr>
<td>Bare</td>
<td>31 (24.60%)</td>
<td>27 (25.23%)</td>
</tr>
<tr>
<td>Definite</td>
<td>37 (29.36%)</td>
<td>42 (39.25%)</td>
</tr>
<tr>
<td>Indefinite</td>
<td>12 (9.52%)</td>
<td>8 (7.47%)</td>
</tr>
</tbody>
</table>

Table 5 – Forms used to maintain reference to animate objects in narrative 1 between scaffolded and unscaffolded groups

<table>
<thead>
<tr>
<th></th>
<th>Scaffolded NNS data (N=5) (Total 198 references)</th>
<th>Unscaffolded NNS data (N=5) (Total 136 references)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Zero</td>
<td>8 (4.04%)</td>
<td>3 (2.20%)</td>
</tr>
<tr>
<td>Personal</td>
<td>91 (45.95%)</td>
<td>52 (38.25%)</td>
</tr>
<tr>
<td>Bare</td>
<td>11 (5.55%)</td>
<td>58 (42.64%)</td>
</tr>
<tr>
<td>Definite</td>
<td>83 (41.91%)</td>
<td>13 (9.55%)</td>
</tr>
<tr>
<td>Indefinite</td>
<td>5 (2.52%)</td>
<td>10 (7.35%)</td>
</tr>
</tbody>
</table>

Table 6 – Forms used to maintain reference to animate objects in narrative 2 between scaffolded and unscaffolded groups

For narrative 1, the ANOVA between the zero nominal forms used showed that their use was not significantly different (F=0.182, P=0.681) between scaffolded or unscaffolded narratives. This was also the case for personal pronouns (F=3.729, P=0.090) and definites (F=0.059, P=0.815). For bare nominal forms, the difference of use between scaffolded and unscaffolded narratives was again seen as not significantly different (F=0.073, P=0.794) and this was also the case for indefinite forms (F=0.264, P=0.621).

For narrative 2, the ANOVA performed between the zero anaphora forms used again showed that their use was not significantly different between scaffolded and unscaffolded narratives (F=3.125, P=0.115), and this was also the case with personal pronouns (F=1.855, P=0.210) and indefinites (F=0.769, P=0.406).

However, significant differences were found in the narratives performed for narrative 2 between unscaffolded and scaffolded forms in participants’ use of definite articles (F=10.145, P=0.013). In addition, there was also a significant difference
found in participants’ use of inaccurate bare nominals between scaffolded and unscaffolded states (F=8.258, P=0.021). As explained above, bare forms are potentially the most damaging NNS contribution to the coherence of narratives for English NSs, so the impact of scaffolding (reducing the use of these forms by 37%) is particularly important for the maintenance of coherence here.

These findings would suggest that for narrative 2 at least, there is some relationship between the provision of scaffolding and the use of definites to create longer distance anaphora, and that the availability of scaffolding is useful in preventing inaccurate bare nominal forms being used to make reference. This hypothesis was supported by a repeated measures ANOVA using within group variables of definite and bare forms which was shown to be significant (F=13.5, P=<0.001). However, this result is perplexing in that differences were only found between scaffolded and unscaffolded narratives for narrative 2, but not narrative 1. Some possible explanations for this phenomenon are found in the discussion section of this paper.

3) What is the relationship between the use of co-referential reference maintenance and distant non-co-referential reference maintenance between native and non-native speaker groups? How does this relationship affect the coherence of the text?

During the coding of the transcriptions for referential form, additional coding was performed to ascertain the referential distance of each form used, as explained in the analysis section of this research. Coding was performed on each referential form used in the study, but the main focus of this research is on the use of personal pronominal co-referential anaphora, and the use of non-co-referential definite articles to signal longer distance anaphoric relationships typical of English NSs, within the narratives recorded. Table 7 below show the results for these forms provided correctly for the obligatory referents within each narrative.
There was no significant difference between the number of co-referential pronouns used between the NS’s performance of stories 1 or 2 (F=2.141, P=0.182). However, when comparing individual narrative cohesion in the form of co-referential pronouns between NS and NNS, the difference between the NS and NNS was highly significant for narrative 1 (F=14.862, P=0.002). A similar pattern is found for narrative 2 (F=4.870, P=0.046) although it should be noted that the P value here suggests that this effect is much greater for the narratives produced for narrative 1 than for narrative 2. For the use of longer distant non-co-referential definite articles for cohesion, the NS use of this form was significantly higher than NNS use (F=28.791, P=<0.001) in narrative 1 as well as narrative 2 (F=5.816, P=0.031). This data suggests that these aspects of cohesion appear to be the main difference between how NS and NNS provide overall coherence within the narratives that were elicited for narrative 1 and narrative 2.

As for the effect of scaffolding on co-referential pronominal forms for narrative 1, there was not a significant difference between scaffolded vs. unscaffolded performance, nor was there for definite article forms. The same was true for narrative 2 for co-referential forms. However, for narrative 2, there was a highly significant difference between the correct use of definite non-co-referential cohesive devices between scaffolded and unscaffolded states (F=27.842, P=<0.001).

Clearly for this aspect of cohesion, the influence of scaffolding helps the participants to create longer chains of accurate cohesive structure over longer distances of maintained narrative for this particular narrative. Interlocutor interactions such as ‘who’ questions to establish reference certainly helped this total (ex: I-who breaks the bottles? P-the dog [pause] or ex: I-whose cycle? Is that his

<table>
<thead>
<tr>
<th></th>
<th>NS narrative 1</th>
<th>NS narrative 2</th>
<th>NS stories combined</th>
<th>NNS scaffolded (narrative 1)</th>
<th>NNS scaffolded (narrative 2)</th>
<th>NNS scaffolded combined</th>
<th>NNS unscaffolded (narrative 1)</th>
<th>NNS unscaffolded (narrative 2)</th>
<th>NNS unscaffolded combined</th>
</tr>
</thead>
<tbody>
<tr>
<td>Use of co-referential personal pronoun</td>
<td>67</td>
<td>102</td>
<td>168</td>
<td>27</td>
<td>69</td>
<td>96</td>
<td>26</td>
<td>36</td>
<td>62</td>
</tr>
<tr>
<td>Use of non-co-referential definite article</td>
<td>46</td>
<td>59</td>
<td>105</td>
<td>20</td>
<td>50</td>
<td>70</td>
<td>16</td>
<td>4</td>
<td>20</td>
</tr>
</tbody>
</table>

Table 7 – participants’ use of co-referential personal pronoun and non-co-referential (distant) definite markers for anaphoric reference.
motorcycle? P-it’s the policeman’s motorcycle), but there were only a few (7) instances where this occurred, and when these were factored out of the analysis, there was still a significant statistical difference between their use in scaffolded and unscaffolded narratives (F=27.528, P=<0.001).

Discussion

From the results, clear differences between NS and NNS reference maintenance were found, in that NSs maintain discourse reference primarily with the accurate use of personal pronominals and definite articles between co-referential clauses, and non-co-referential clauses (over X-2 in anaphoric distance). This appears to be the main difference between NSs’ performance of coherent and cohesive narratives and those of the NNSs, who were not able to supply native-like use of these forms in the right contexts. There is a strong possibility that this difference in performance is not caused by differences in the pragmatic ability to mark new and given information between and across clauses (which is said to be universal), but by mismatches in the cohesive means to do so.

The other implication of these results is that scaffolding helps L2 learners to create and hold more accurate reference to discourse referents, and that instances of unscaffolded narrative discourse present unnatural difficulty for the L2 speaker. This appears to be characterized by an inversely proportional relationship between the accurate use of the definite article to maintain long distant anaphoric reference between animate referents (typical of NSs), and the incorrect, inaccurate use of bare nominal forms (typical of NNSs) to maintain reference, as found in narrative 2. This suggests that for participants with a lower L2 proficiency, a reliance on a limited lexical vocabulary to maintain reference is employed in the face of a lack of consistent grammatical means to do so.

Despite the significant difference found between scaffolded and unscaffolded performances of the narratives in terms of definite non-coreferential (distant) cohesion in narrative 2, the results were disappointing in that the difference in the use of other cohesive devices between these two groups for both narrative 1 and 2 was not statistically significant. I attribute the differences here post-hoc to the distinction between main and supporting characters and the distribution of referential devices to
maintain reference to them in discourse. Karmiloff-Smith (1985) showed an effect of referent type (main vs. supporting characters) with pronominal forms being used to introduce main characters and indefinites being used to introduce supporting characters, and any distinction between how referents are introduced in discourse may be also be evidenced in how those referents are maintained. In my opinion, while the number of referents in both stories was similar, it could be argued that the main/supporting status of some of the referents in narrative 1 is unclear – potentially Tintin, the dog, or the Mexican knife thrower may be considered as main characters, or Tintin may be seen as a secondary character, leaving the dog and the knife thrower as main characters. In narrative 2 however, Tintin and the dog are clearly main characters whose exploits continue throughout the events of the narrative, which I believe goes some way to explaining the differences in results between narratives 1 and 2.

In addition, the nature of the events that take place in a narrative is also a factor that may influence the coherence of any performance, as evidenced in Pu (1995), who in her research on Mandarin narratives claimed that ‘episodic organization of narrative production has psychological content: The story was hierarchically organized and remembered as a series of episodes. Subjects were highly sensitive to episode boundaries, regardless of how the picture sequence was segmented (Pu, 1995, p. 298).’ In my opinion, narrative 2 seems to follow a more linear sequence of events that allows for greater use of extended co-reference, while narrative 1 seems to be divided into two quite distinct sections (the section with Snowy in the house, followed by the scene with the knife thrower). This factor may go some way to explain the lack of a significant result in narrative 1 for extended co-referential forms, in that the ‘scene’ changes meant that characters were not kept in a topical position that would allow for extended co-reference. Sensitivity to such boundaries in narratives was evidenced in Lee (1981), who attributed better recall results for narrative sequences where co-referential clauses were contrastive rather than additive (as described in Halliday & Hasan (1976) in terms of cohesive conjunction), which suggests that language users were performing higher-order processing on unexpected co-referential clauses rather than predictable ones.

Of interest to this research was data that suggested that there was a significant difference between the use of accurate longer-distance non-co-referential definite reference between scaffolded and unscaffolded instances of narrative performance in
favour of scaffolded narratives, and that there was also a significant difference in the number of inaccurate bare nominal forms used in unscaffolded narratives. From this data we can conclude that the unscaffolded narratives made it harder for the participants to maintain accurate reference throughout the narratives, which made their narratives less coherent.

To test for this perceived lack of coherence, a correlation matrix was performed using SPSS to look at the interaction between the use of accurate co-referential personal pronominals, non-co-referential distant definite forms and the Likert scale scores of 5 NS participants who analyzed the 20 NNS narratives and rated them for coherence. The results are shown in table 8 below:

<table>
<thead>
<tr>
<th></th>
<th>co-ref</th>
<th>definite</th>
<th>Likert</th>
</tr>
</thead>
<tbody>
<tr>
<td>co-ref</td>
<td>1</td>
<td>0.877**</td>
<td>0.505</td>
</tr>
<tr>
<td>Sig. (2-tailed)</td>
<td>0.001</td>
<td>0.023</td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>20</td>
<td>20</td>
<td>20</td>
</tr>
<tr>
<td>definite</td>
<td>0.877**</td>
<td>1</td>
<td>0.505**</td>
</tr>
<tr>
<td>Sig. (2-tailed)</td>
<td>0.001</td>
<td>0.007</td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>20</td>
<td>20</td>
<td>20</td>
</tr>
<tr>
<td>Likert</td>
<td>0.505</td>
<td>0.505**</td>
<td>1</td>
</tr>
<tr>
<td>Sig. (2-tailed)</td>
<td>0.023</td>
<td>0.007</td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>20</td>
<td>20</td>
<td>20</td>
</tr>
</tbody>
</table>

**. Correlation is significant at the 0.01 level (2-tailed).
* . Correlation is significant at the 0.05 level (2-tailed).

Table 8: Results of correlation matrix between cohesive devices and coherency judgment.

As can be seen from the table, the accurate use of both co-referential and definite forms in the narratives correlate with the Likert-scale scores given for coherence by the NSs surveyed. This data strongly suggests that narratives with accurate co-referential and longer distance reference maintenance were rated as more coherent during the coherency judgment task. This correlation can be seen visually in the chart below:
Limitations of the Study and Opportunities for Further Research

Due to time/availability constraints, I was unable to have more than five participants per group, and I am confident that with an expansion of the number of participants, it is likely that significant differences would eventually be found between scaffolded and unscaffolded narratives in this regard.

I am also aware that the lack of a significant difference in this area may have been due to the low overall English proficiency of the participants (average 250 on TOEIC), in that there may have been situations where scaffolding may have helped participants of higher proficiency achieve coherence, even if it did not have the desired effect in participants with lower proficiency. These are issues that I hope will be reviewed in further studies on this topic. In addition, the status of main/supporting characters in discourse will necessarily be controlled for in future research, as well as the linearity of the events in the narratives to ensure that one sequence of events is not substantially more disjointed than the other.

A final issue for this research was raised in Reismann (1993), who states:

It is always possible to narrate the same events in radically different ways, depending on the values and the interests of the narrator. There is no
reason to assume that an individual’s narrative will or should be entirely consistent from one setting to the next (Reismann, 1993, p. 65).

I strongly agree that consistency within narratives would be something that is very hard to achieve, particularly if we consider the distinction between scaffolded and unscaffolded versions of the narratives used for this research. In addition, while the values and interests of the narrator certainly play a role in how the narrative is constructed and the path it takes, I feel that this effect goes both ways. I feel that while the role of the scaffold is primarily to help bring about the successful negotiation of meaning necessary for the narrative to be coherent, the aims of the scaffold (realized in the form of the scaffolding they choose to employ during the speakers’ performance) may play a large role in determining the course the speaker eventually takes. It would be potentially interesting to research the effect different interlocutors have on the performance of a single narrative speaker, as well as the effect of having two or more interlocutors interacting when scaffolding a narrative.

**Conclusion**

In terms of the hypotheses laid out in this research, differences between how NS and NNS maintain cohesive reference maintenance in English have been found, and the more accurately the NNS maintains co-referential and distant reference through the use of these cohesive devices, the more coherent the final discourse is likely to be. This theory was supported by a coherency judgment task performed by NSs on the transcripts of the NNSs’ narratives, which showed a strong correlation between NS interpretations of coherent discourse and the successful maintenance of accurate co-referential and non-co-referential reference. The accuracy of these kinds of reference can be supported by the presence of scaffolded collaborative input on the part of an interlocutor, which suggests that unscaffolded discourse is more challenging for the NNS when maintaining reference. However, the results gained from this study do not go far enough, and further study is required (with more participants) to investigate to what extent the influence of scaffolding has on accurate reference maintenance. Other potential sources of interest for further research include a possible variation in stimuli (picture sequences vs. open questions), the number of and intentions of the interlocutors during narrative performance, and the perspective that the speaker takes during the narrative production. Other source languages with
differing grammatical cohesive devices should also be investigated, as well as learners of different L2 proficiencies, and these variables will be tackled by this researcher at a later date.

References


Hergé. (1948). *Tintin and the Seven Crystal Balls*. Moulinsart, Paris


Appendix 1 - Description of Narrative 1 – Adapted from ‘Tintin and the Seven Crystal Balls’ (Herge, 1948).

Tintin approaches a mansion. He meets a butler at the door. Snowy the dog sees a cat and begins to chase it. Snowy slips on the floor as the butler is carrying some drinks on a tray, and the butler trips on the dog, falls, landing on one hand. Snowy jumps over the tray the butler is holding with his other hand, knocking the bottles off the tray. Snowy chases the cat upstairs, but is injured by the cat and comes back downstairs again. Tintin and Captain Haddock scold Snowy.

Tintin & the Captain head into a theatre, where a Mexican knife thrower and an Indian assistant are on stage. The Mexican throws a knife at the Indian, and the knife lands next to the Indian’s ear. More knives are thrown by the Mexican, all landing around the Indian. The Mexican gets a member of the audience to blindfold him. He throws a knife at the Indian, who is holding a melon that has been cut in half. The knife sinks into the melon. The Mexican turns to the crowd and bows. Tintin and the Captain applaud.

Appendix 2 – Description of Narrative 2 – Adapted from ‘Tintin in America’ (Herge, 1932).

Tintin and Snowy are on board a train. As they exit the train, a taxi and its driver are waiting for them. The taxi driver holds open the door for Tintin and Snowy. They get into the taxim, but as the taxi drives away, shutters come down over the windows, preventing any escape. As the car heads along the road, its tire bursts, and the taxi driver has to change the wheel. The taxi driver drives away, but Tintin and Snowy have escaped the taxi by cutting a hole in the floor of the taxi with a saw. As they walk along the road, they stop two police officers who are approaching on a motorbike and sidecar. They all drive after the taxi together, and when they catch it, one of the officers pulls out a gun and points it the taxi driver. As the police arrest the driver, a gangster appears from behind a tree, and throws a boomerang, knocking out the taxi driver. He catches the boomerang, and speeds off on the police officers’ bike. Tintin and Snowy and the police get into the taxi, and chase the gangster into a city. Tintin drives the taxi into another car, causing an accident. An ambulance arrives, and two doctors take Tintin into an ambulance, which drives away.
Appendix 3:
Instructions to Participants

You will see a picture sequence and I would like you to tell me the narrative of what happened.

You will have some time to look at the whole narrative before you begin, and you can look at the pictures while you give the narrative, so don’t worry about trying to remember what you saw in the pictures. Use the ← and → buttons on the keyboard to show the pictures.

시작하기 전에 전체 그림 볼 시간을 조금 드릴 테니, 그림들을 보시고 이야기를 해 주세요. 즉, 봐도 그림들을 기억 안 하셔도 됩니다.

키보드 ← 와 → 를 사용하여 그림을 보십시오.

There are blank speech bubbles in the pictures you see, but you do not have to make speech for the characters.

그림에 빈 대화창이 있으나 기기에 들어갈 대화를 꺼워 맞추도록 안 하셔도 됩니다.

You may begin telling the narrative whenever you are ready.

자 그림 준비되시면 이야기 할 준비를 해주세요.

Don’t worry about mentioning every picture in the sequence if you cannot do so.

각각의 사진을 모두 설명 안 하셔도 됩니다.

After you have completed the narrative, you will see another, similar picture sequence. We will do the second narrative in the same way as the first.

이야기가 끝나고 또 다른 이야기를 보시게 됩니다. 두 번째 그림 역시 첫 번째 스토리처럼 이야기해주세요.

In one of the stories, I might talk to you as you read the narrative. For the other narrative, I will be quiet and let you speak by yourself.

이 중 하나의 스토리는 이야기를 이야기하는 도중에 선생님이 옆에서 도와줄게 됩니다. 하지만 나머지 이야기는 전체 스토리를 흘러서 말씀해 주셔야 합니다.

If you feel you cannot continue the experiment for any reason, please let me know and we can pause and begin again at another time.

만약 이야기를 하는 도중에 장시 중단하고 싶으시면 선생님께 말씀하여 중지를 하시고 다시 시작하실 수 있습니다.

The experiment is being recorded on tape, and will only be used for the purposes of this experiment.

이 실험이 전부 녹음되어 녹음된 자료는 이 특정 리서치에만 이용됩니다.

Your names/identities will not be used in the publishing of this research.

당신의 이름이나 아이디는 절대 공개되지 않습니다.
Appendix 4 – Permissible Scaffolding for Interlocutor (Parts taken from Wilkes-Gibbs (1995), & Pellegrini and Galda (1990))

<table>
<thead>
<tr>
<th>Lexical Completion</th>
<th>It’s a…. – It’s a dog?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Phrasal Completion</td>
<td>He is…. – going into the house?</td>
</tr>
<tr>
<td>Sentence Completion</td>
<td>He is… - he is running away?</td>
</tr>
<tr>
<td>Extra move in Completion</td>
<td>It’s a… - it’s a dog – dog – right.</td>
</tr>
<tr>
<td>Continuation</td>
<td>He went up – and the dog went down.</td>
</tr>
<tr>
<td>Repetition</td>
<td>It’s a dog – a dog?</td>
</tr>
<tr>
<td>Ask for extension</td>
<td>Could you tell me more about that</td>
</tr>
<tr>
<td>Clarification question</td>
<td>What do you mean by that?</td>
</tr>
<tr>
<td>Descriptive Question</td>
<td>What does it look like?</td>
</tr>
<tr>
<td>Evaluation</td>
<td>That’s weird isn’t it?</td>
</tr>
<tr>
<td>Filler</td>
<td>Hmmmmmmmmm</td>
</tr>
<tr>
<td>Label</td>
<td>That’s a ________</td>
</tr>
<tr>
<td>Label Question</td>
<td>What’s that called?</td>
</tr>
<tr>
<td>Reinforcement</td>
<td>It’s a dog. – Good</td>
</tr>
<tr>
<td>Relate to experience</td>
<td>Have you seen that before?</td>
</tr>
<tr>
<td>Role Clarification Question</td>
<td>Who is that? Who is the Robber?</td>
</tr>
<tr>
<td>Additive slot</td>
<td>The robbers ran AND</td>
</tr>
<tr>
<td>Causal Slot</td>
<td>The robbers ran BECAUSE</td>
</tr>
<tr>
<td>Event Slot</td>
<td>What happens now?</td>
</tr>
<tr>
<td>Temporal Slot</td>
<td>The bridge fell and THEN</td>
</tr>
</tbody>
</table>
Socio-Economic Orientations in Foreign Language Learning Motivation: The Case of Yemen

Rais Ahmed Attamimi  
*Hadramout University of Science and Technology, Yemen*  
rais_attamimi@yahoo.com

Hajar Abdul Rahim  
*Universiti Sains Malaysia, Malaysia*  
hajar@usm.my

Bio Data

Mr. Rais Attamimi is an Assistant Professor in the English Department of the College of Arts and Education, Hadramout University, Seiyun, Yemen. He obtained his doctorate in Applied Linguistics from Universiti Sains Malaysia. He has taught in the Yemeni primary, secondary, and tertiary educational systems for 9 years. His research interests are socio-psychological variables in FL contexts, sociolinguistics, and Applied Linguistics.

Hajar Abdul Rahim is an Associate Professor in the department of English Language Studies of the School of Humanities, Universiti Sains Malaysia.

Abstract

This paper examines the relationship patterns between socio-economic factors, i.e. parental occupations, cultural capital, and motivation to learn English in the Yemeni context. Two survey questionnaires were used in this study, based on Gardner’s (2001a) framework and Bourdieu’s (1985, 1986, 1989) status-based approach to social stratification. Questionnaires were administered to 142 fourth-year students in the English Department of the College of Arts and Education, the Hadramout University of Science and Technology, Yemen. Besides questionnaires, individual semi-structured interviews were used to obtain supportive data. A range of quantitative and qualitative analyses were used to analyse the data of the study. The results of the study provide tangible evidence of the existence of a strong relationship between parental occupation, cultural capital, and the Yemeni students’ motivation to learn English.

Keywords: language learning motivation, integrativeness, parental occupation, and cultural capital
Introduction

The notion of ‘integrativeness’ is a crucial element in the Gardnerian’s socio-educational model of attitudes and motivation. The model proposes that learners who have integrative orientations towards learning the target language (henceforth TL) possess favourable attitudes towards the language community, and an inherent interest in foreign languages (Gardner, 1985; Masgoret et al., 2001). These “integratively motivated” learners also exhibit different aspects of motivated behaviour such as effort, an expressed desire, or enjoyment of the TL learning process (ibid). Such inclinations also suggest that these learners desire to “come closer to the other language community” (Gardner, 2001a, p. 5).

The interest among researchers and educators in finding the link between achievement behaviour and success in language learning has generated considerable research in many second language (henceforth SL) and foreign language (henceforth FL) contexts (Lamb, 2004). More importantly, the interest in this area has witnessed a kind of “motivational renaissance” (Gardner and Tremblay, 1994a) with new concepts from educational psychology widening the scope of the motivational research agenda, resulting in the emergence of new understandings and conceptualizations in the study of motivation (Lamb, 2004). One of the most crucial of these is the reformulation of the notion of integrativeness itself.

Gardner (2001b) considers the amount of interest in the area, evidenced by the numerous studies, as an endorsement of the significance of integrativeness. This assumption is further encouraged by the fact that only minor changes had been made to the socio-educational model that he originally proposed. Yet, recent studies in the FL contexts have urged researchers to revisit the definition of integrativeness. Studies by Dörnyei and Clément (2001), Dörnyei and Csizér (2002), Csizér and Dörnyei (2005a), and Lifrieri (2005) in particular have revealed a major shortcoming in the traditional definition of integrativeness. While its predictive value has been confirmed in many language learning contexts, the researchers find that the original understanding of integrativeness is lacking in its explanatory power in FL settings. The conventional interpretation of this concept presupposes the existence of a recognizable group of native speakers in the learner’s immediate environment with whom s/he wishes to interact and integrate in some way. This becomes problematic in the FL context because, as the researchers argue, the existence of native language speakers in the FL environment is not typical (Dörnyei and Csizér, 2002; Csizér and Dörnyei, 2005a).
Given this reality and the argument that motivation is found to reside in the interaction of the individual and environmental dimensions (Hickey, 1997; McGroarty, 2001), Dörnyei and Csizér propose a reformulation of the concept of integrativeness. They suggest that FL learners are motivated to learn a TL to enact “possible selves” which are congruent with some social traditions and/or habits in the learners’ environment.

The whole notion of FL learners enacting “possible selves” is intriguing and one that requires investigation in different FL learning environments. Indeed this is the motivation and stimulus for the current study which, based on the reformulation of integrativeness in FL settings as proposed by Dörnyei and Csizér (2002), and Csizér and Dörnyei (2005a), set out to investigate the nature of FL learners’ possible selves in the Yemeni context. Yemen comprises a highly stratified society (Al-Bana’a and Al-Jabli, 2002), so the focus of the study is on the relationship between Yemeni EFL students’ socio-economic backgrounds and their motivation to learn or use English to construct possible selves or an identity. In particular, the study is interested to find out if the subjects’ motivation to learn English is influenced by the existence of a TL group which comprises Yemeni speakers of English who have a good socio-economic standing. The group comprises a remarkably small number of individuals but who have achieved a social status that is very much envied by others in the society. The members of this group mainly include those who are employed in the oil project in the Hadramou Province in Yemen.

The Study

The current research replicates Lifrieri’s (2005) study which was conducted on a group of Argentinean EFL school boys. Lifrieri’s study investigated in general the impact of some sociological aspects on learners’ motivation towards learning English. The sample in Lifrieri’s study featured young school boys who were not able to express accurately their own motivations and they might “have just provided random answers” (Lifrieri, 2005, p.68). On the other hand, the current study focuses on the Yemeni context to find out how Yemeni university students’ socio-economic status correlates with their English language learning motivational patterns. Given the recent emphasis on expanding the traditional social scope of motivational models to include macro-social variables (McGroarty, 1998; 2001; Spolsky, 2000; Clément and Gardner, 2001; Dörnyei, 2001b; Gardner, 2002), this study is timely as Yemeni EFL students live in a highly stratified society where English is associated generally with groups possessing more social, cultural, or financial resources (Nielsen, 2003). As
established earlier, there exist a small number of socio-economically developed Yemeni English language speakers who have achieved an enviable social status in the Yemeni society and hold high-ranking, well-paid jobs in the oil industry as well as at foreign embassies and institutions. It is highly likely that Yemeni EFL students perceive practices such as speaking and using English by individuals in higher social positions as a symbol of status and prestige.

In this study, the socio-economic backgrounds of the sample are: the occupational status of their parents (i.e. economic capital) and the accumulation of cultural capital in their families. The present research thus aims to establish a link between students’ socio-economic backgrounds and motivation to learn English in the Yemeni EFL context, based on a methodology that is informed by Gardner’s (2001a) integrative motivation framework and Bourdieu’s (1985, 1986, 1989) reproduction theory. It is hoped that the link between Gardner’s framework of L2 motivation and Bourdieu’s approach in sociology will reveal an accurate picture of the relationships between the components and/or subcomponents of FL motivation and socio-economic factors. Based on Bourdieu’s status-based approach to social stratification, students’ internalised knowledge and perceived modes of status differences between them and the local TL group are shaped by their socialisation in given socio-economic positions (Bourdieu, 1985; 1986; 1989). On the basis of the students’ knowledge towards the meaning of the representation and the social practices’ symbols of the local TL group, their motivational patterns may then relate to their drive to identify or integrate with the social practices of that group.

The study set out to answer two research questions: 1) To what extent is there a reliable relationship between the economic capital (i.e. parental occupation) and the motivation of Yemeni university EFL students? and 2) To what extent is there a reliable relationship between the cultural capital (i.e. parents’ educational level) and the motivation of Yemeni university EFL students? The answers to these questions will contribute new insights into the ongoing research on FL learning motivation and have an impact on the future of FL instruction and educational planning and policy in Yemen.

Methodology

The current research is framed theoretically and conceptually within Gardner’s (2001a) integrative motivation framework and Bourdieu’s (1985; 1986; 1989) status-based
approach to social stratification. Also, the study is guided by the hypothesis that if there are individuals who are characteristically associated with English and who are also socio-economically established in the learners’ social environment, their integrative tendencies are most likely associated with those individuals, even if they are not native English speakers. Given the concept of integrativeness as proposed by Csizér and Dörnyei (2005a) and Dörnyei and Csizér (2002) and the perception of English in Yemen, that is, knowing English is regarded as a prestigious practice associated with the high society (Al-Quyadi, 2000), the present study assumes that the Yemeni EFL students’ integrative motives towards the local TL group are crucial to the relationship between their socio-economic status and motivation to learn English.

In order to address the objectives of the study, a survey was conducted at the English Department of the College of Arts and Education, the Hadramout University of Science and Technology, Yemen. Two sets of questionnaires were administered to 142 fourth-year Yemeni students. A correlational analysis of the mean scores from both questionnaires was carried out to ascertain the degree and direction of relations between motivation to learn English and socio-economic factors. Besides quantitative data, the research also obtained supportive data through qualitative research methods using interviews. The findings from the questionnaires and the interviews were combined to provide insight into the degree of relationships among the investigated variables as well as the students’ habitus. To explain, habitus is one’s view and place in the world and it serves to show how a student navigates his/her way through the educational system (Bourdieu, 1985, 1986). In this study, habitus is used as an analytical concept to explain the sample’s integrative motives towards the local TL group. The format of the questionnaires and the design of the interviews are described in the following sections.

**Quantitative Data Collection Instruments**

Questionnaires were adopted from Lifrieri’s (2005) research and used in this study, mainly because they facilitate the task of gathering data from a large number of participants (Judd, et al., 1991). The two questionnaires of the study measured and investigated the following:

1) The independent variables:
   a) Parental Occupations: (1) father’s Past/Present Occupation and (2) mother’s Past/Present Occupation.
b) Cultural Capital: (1) types of newspapers and (2) frequency of reading by the subjects’ parents.

2) The study attempts to determine how the independent variables influence the following four dependent motivational variables:
   a) Integrativeness.
   b) Motivation.
   c) Orientations.
   d) Overall Motivation.

The Motivation Questionnaire

The motivation questionnaire was used to collect data on the students’ motivation to learn English. Thirty statements were designed and used as stimulus items in the questionnaire. Subjects responded to each item based on an expanded Likert scale of seven points, ranging from Strongly Disagree to Strongly Agree. The original Attitude/Motivation Test Battery (AMTB) was modified to suit the objectives of the study and the context of EFL learning in Yemen (see Appendix A). Firstly, statements were rephrased to tap information on the students’ motivation, and their attitudes towards the TL group. Secondly, to obtain information on their general attitudes towards language learning, the original AMTB statements were adapted. Thirdly, certain constructs, such as the component in Integrative Motivation, which measure micro-contextual variables, were excluded to focus on the concern of the study that is the macro-contextual dimension of motivation. Following Gardner’s model, the items were grouped under the three major components of Integrativeness, Motivation and Orientations. The motivation questionnaire used in the study measured the following constructs:

1) **Integrativeness:**
   This component is measured by:
   a) Attitudes towards the TL Group, i.e. attitudes towards the qualities and habits of Yemeni speakers of English as a group with a better position, socially and economically;
   b) Interest in Foreign Languages, i.e. general preference for language learning; and
   c) Integrative Orientation, i.e. a desire to strengthen or build ties with in- or out-group members.
2) **Motivation:**

This component is measured by:

a) Motivational Intensity, i.e. the amount of effort given to learning the TL;

b) Attitudes towards Learning the TL, i.e. the affective response to learning English, and

c) Desire to Learn the TL, i.e. the degree of commitment to learning.

3) **Orientations:**

This component is measured by:

a) Integrative Orientation (see 1 (c) above), and

b) Instrumental Orientation, i.e. the inclination to learn EFL for pragmatic reasons.

The Socio-economic Questionnaire

The socio-economic questionnaire (see Appendix B) was used to obtain data on the socio-economic status of the sample. The metaphors of economic capital and cultural capital based on Bourdieu’s (1985, 1986) reproduction theory guided the design of the questionnaire. The two dimensions were measured using information on the participants’ parental occupation and parental educational level respectively. Thus, the socio-economic questionnaire comprises two sections, (1) parents’ occupation and (2) parents’ education. Each section contains 6 questions which are either yes-no, multiple-choice, or frequency-type questions while the rest is open-ended.

The motivation questionnaire and the socio-economic questionnaire were administered sequentially to 142 students who were randomly selected from a target population of 250 students. They were fourth-year students at the English Department of the College of Arts and Education, the Hadramout University of Science and Technology. The responses from the 142 questionnaires were tabulated and then calculated to obtain the average score for the motivational and socio-economic factors. This was followed by a statistical analysis using the SPSS 11.5 for Windows to find out the correlation result between the variables.

**Interviews**

The qualitative data for the study was obtained from interviews. To facilitate the process of eliciting supporting data, semi-structured interviews were carried out using six carefully constructed questions to address different concerns of the study. Each question contains a set of issues aimed at eliciting further information from the participants as follows:
i. What do you feel are the benefits of learning English?

ii. Are your parents supportive of your learning English? Why? How?

iii. How do you spend your time with your friends?

iv. How do you perceive a Yemeni who is also a good speaker of English?

v. Would you like to identify yourself with Yemeni speakers of English who occupy prestigious positions in the society? Why? Why not?

vi. Do you think that you will need English in the future? Why?

Whilst the questions were designed to be open-ended to allow for a wide range of possible answers and to enable correlation with data from the questionnaires, as can be seen above, some of the interview questions are yes-no questions. In these cases, more information was elicited from the participants by asking them to explain and elaborate on their responses.

The participants for the semi-structured interviews were 20 students who were selected from the pool of 142 participants. The interviewees were given the choice to respond in Arabic so they could express themselves better. The interviews were audio-recorded and conducted one week after the completion of the questionnaires. The data were transcribed using Arabic punctuation signs to support the meaning of the utterances, and then translated into English. Following this, the semi-structured interview data were examined and categorised based on the themes developed in the framework of study.

**Results and Discussion**

The presentation and discussion of the results are guided by the main issues in the study as reflected in the research questions, namely the relationship between Yemeni university EFL students’ motivation and economic capital, and between their motivation and cultural capital. The results on the students’ socio-economic status and motivation to learn English presented and discussed below are derived from the descriptive and correlation analyses of the data from the questionnaires as well as those obtained from the interviews.

**Basic Descriptive Analyses**

The first section provides an account of the students’ parental occupations and the accumulated volume of their family’s cultural capital derived from the socio-economic
questionnaire. Then, an overview of the motivational patterns of the sample that emerged from the results of the basic descriptive analyses conducted on the data of the motivation questionnaire is given. The mean score distribution (M), standard deviation (SD), and frequency and percentage distributions of sample’s responses from the two questionnaires are presented. The different trends in the responses of the sample with regard to the components and subcomponents in the questionnaires are then graphically represented.

**Parental Occupations**
Overall, the results show that the occupations of the parents are diverse, ranging from professionals such as doctors, lawyers, school teachers, university lecturers, government officials to casual workers and labourers. The results indicate that 75 (52.8%) of the students’ fathers are employed, while 67 (47.2%) of the sample confirm the unemployment status of their fathers. As regards mother’s occupation, the survey shows that 76 (53.5%) of the samples’ mothers are unemployed and 66 (46.5%) are employed. These results are displayed in the following two figures respectively.

![Figure 1: Fathers’ employment status.](image-url)
Father’s Past/Present Occupations are distributed within a range of 0-4 of Hollingshead’s (1975) occupational scale. The average mean score of the sample is $M=3.3\ (SD=1.23)$ which is notably higher than the one obtained for the component of Mother’s Past/Present Occupations (i.e. $M=1.90, SD=0.89$). After examining the frequencies and percentages of the participants’ responses, some interesting results have emerged. It is evident that out of the 142 valid responses, only 20 (14.1%) of the students came from families where the fathers have/had menial jobs, that is, they were/are employed at somebody else’s company, farm, or house as unskilled workers. A majority of the fathers (i.e. 48 (33.8%)) are semi-skilled workers: 25 (17.6%) of them are machine operators, barbers, bus drivers, childcare workers in non private household, cosmetologists, and file clerks, while 23 (16.2%) have/had jobs as guards, nursing aides, private housekeepers, seamstresses, service workers, taxi drivers, and truck drivers. Interestingly, only 7 (4.9%) of the participants’ fathers have/had more slightly high qualified jobs as teachers, salespeople, small owners, military personnel, and administrators. Lastly, 67 (47.2%) of the fathers are unemployed or retired. All these results are presented in Figure 3.
In contrast, the results on Mother’s Past/Present Occupations are distributed within a very limited range of scores of the Hollingshead occupational scale, i.e. 0-2. Mothers’ jobs are distributed across the two occupational scales of ‘extreme’ job categories of unskilled and menial service jobs. Of the total number of employed mothers, 57 (40.1%) of them are menial service workers including farm labourers, service workers, bellhops, maids, dishwashers, janitors, and ushers, while 9 (6.3 %) are unskilled workers such as cooks, food service workers, garbage collectors, gardeners/ground keepers, labourers, laundry/dry cleaning operators, school monitors and waiters. More than half of the mothers i.e. 76 (53.5%), are unemployed/retired, or homemakers. The percentage distribution of the mothers’ occupations is presented in Figure 4.
Cultural Capital

The cultural capital component focuses on the educational level of the parents. In the socio-economic questionnaire, this was measured based on two cultural aspects: (1) types of newspapers, and (2) frequency of reading. The type of newspapers is measured by a series of open-ended questions; whereas the frequency of reading newspapers is measured by frequency-type questions. The results of cultural capital are presented and discussed in terms of frequency and percentage distributions.

In this study, fathers in the sample’s families clearly read newspapers more frequently than mothers. It is also found that 107 students (75.4%) declare that their fathers are educated, while 35 (24.6) are uneducated. Only 46 (32.4%) of the students agree that their mothers are educated, while 96 (67.6%) stated otherwise. These results are presented in Figure 5.
The types of Yemeni newspapers frequently read by parents are: *Al-Ayyam* (75%), *Al-Sahwa* (52%), *Al-Thura* (47%), and *Yemen Times* (24%). Of the 107 students, 62 (i.e. 57.9%) said that their fathers read newspapers everyday, 20 (18.7%) read newspapers twice or thrice a week, and 25 (23.4%) read newspapers only on Fridays. These results are displayed in Figure 6.

![Figure 5: Parental education.](image)

![Figure 6: Fathers’ frequency of reading newspapers.](image)

The results show that mothers read much less compared to fathers. The results show that out of the 46 educated mothers, only 10 (21.7%) read newspapers everyday, 28 (60.9%) read newspapers twice or thrice a week, and 8 (17.4%) read newspapers only on
Fridays. The results of the frequency of reading newspapers by mothers are presented in Figure 7.

![Motivation to Learn English](image)

**Figure 7**: Mothers’ frequency of reading newspapers.

**Motivation to Learn English**

The results of the basic descriptive analyses indicate that the motivation of the sample is generally characterized by high levels of Integrativeness, Motivation, and Orientations. Table 1 presents the average mean scores and standard deviations of the four main dependent motivational variables.

<table>
<thead>
<tr>
<th>Motivation Variables</th>
<th>M</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Integrativeness</td>
<td>5.6</td>
<td>1.3</td>
</tr>
<tr>
<td>2. Motivation</td>
<td>5.6</td>
<td>1.3</td>
</tr>
<tr>
<td>3. Orientations</td>
<td>5.5</td>
<td>1.4</td>
</tr>
<tr>
<td>4. Overall Motivation</td>
<td>5.6</td>
<td>1.4</td>
</tr>
</tbody>
</table>

An examination of the results given in the table above shows that there are considerably high levels of Overall Motivation (M= 5.6) which exist among 113 (79.3%) students in the sample. Likewise, the average mean scores of the three main
motivational components are similarly high. The average mean scores of the three motivational components tend to be close to the highest score value in the motivation questionnaire, i.e. 7. This is evident from the two average mean scores of 5.6 equally obtained for Integrativeness and Motivation. In both cases 116 students (82%) positively rated all the items under these two motivational components. Also, the average mean score corresponding to Orientations is 5.5 with 108 students (76%) in the sample rating the items under this motivational component positively. The results in Table 1 are presented graphically in Figure 8.

![Figure 8: The distribution of mean scores of the motivational components.](image)

The high average mean scores and frequencies suggest that the majority of the Yemeni students in the sample are highly motivated to learn English. The students in this study are mature enough to be able to provide responses on their actual socio-economic backgrounds and are regularly engaged in the English language to reflect on their attitudes and motivation towards learning the TL. In the individual semi-structured interviews, a majority of them (i.e. 19 students) expressed that they had become more interested in learning English and believed that it would offer an enjoyable challenge for them. They also felt satisfied with the possible future opportunities which seemed to have motivated them to study the language. The students also stated that their increased interest in English was accompanied by favourable attitudes towards the TL group members. In other words, the students expressed that there is an inherent desire to learn English in order to form ties with the
Yemeni speakers of English. So, it can be said that the students in the sample are motivated and that they partake in language learning because of the satisfaction they feel and the interest they have gradually developed in the TL and towards the local TL group.

The positive integrative orientations among students in an EFL context like the Yemeni case must be given due importance. The inclination towards English is found in the responses from the semi-structured interviews. With reference to their fondness of learning English, 17 students stated that talking with other Yemenis in English was useful and helpful to practise expressing what they thought, and to enable them to know more about members in that group. Almost all the participants remarked that they were capable of communicating in English with their classmates and other friends in the surrounding area. This perhaps could be seen as a kind of confidence that generated the desire to be fluent in English and to integrate with the Yemeni speakers of English with good socio-economic positions in the society.

As English in Yemen is a compulsory foreign language subject and not a medium of instruction in schools or colleges, the high integrative tendencies shown by the sample indicate their inclination to be in the same socio-economic position as that of the local TL group. This group is the ideal example available in the students’ closer environment. Hence, the students construct identities that are mostly congruent with the practices and traditions of that group, which support and boost both the integrative and instrumental motivation towards learning English.

An interesting comment by one of the participants in the semi-structured interviews is worth mentioning here: “... I would like to speak English like some Yemenis so the opportunity of securing a good job in future is ensured”. This statement indicates the existence of high levels of both integrative and instrumental orientations in the sample. Interestingly, the findings from the individual semi-structured interviews are consistent with the descriptive results. Nineteen students in the interviews mentioned that they realised that English is more useful as a means of communication and they acknowledged that their growing interest in learning it would eventually help them to be hired in one of the oil companies. The students further expressed their wish and desire to be in the same social and economical position, with the Yemeni speakers of English, particularly the local prestigious TL group.

The students’ answers in the interviews concerning the three motivational subcomponents, i.e. Motivational Intensity, Attitudes towards Learning the TL, and Desire to learn the TL, can be summarised as follows: (1) their belief in succeeding in English is clearly reflected by the amount of effort exerted in learning it, (2) the degree
to which they aspire to achieve advanced proficiency in the language is associated with an ultimate interest and an actual desire to know the language, and (3) their admission that they think about and search for new ways and/or techniques to learn English better, and plan to continue to learn English.

All the above reasons may account for the increase in the average mean scores of the motivational components. The results reported accord with those of Chou’s (2005) study which found that English major students scored high in almost every aspect of the motivational factors. The participants of the interviews in the current study commented that learning English is a means to an end. It is a way for them to achieve pragmatic benefits, such as to be competent in English in order to get employed in the oil project and occupy high socio-economic positions just as the local TL group members has.

**Correlation Analyses**

This section discusses the relationships between the independent and dependent variables based primarily on the results of the correlation analyses. The analyses were conducted to determine the aspects of the sample’s motivated behaviour that might vary as a function of the students’ socio-economic factors. The relationships of the independent socio-economic variables and the components/subcomponents of the dependent motivational variables were assessed by correlating: (1) Cultural Capital, (2) Father’s Past/Present Occupation, and (3) Mother’s Past/Present Occupation against the constituent multi-item subcomponents in the motivation questionnaire, namely (1) Integrativeness (i.e. Attitudes towards the TL Group, Interest in Foreign Languages, Integrative Orientations), (2) Motivation (i.e. Motivational intensity, Attitudes towards Learning the TL, Desire to Learn the TL), and (3) Orientations (i.e. Instrumental Orientation and Integrative Orientation). The Pearson product-moment correlation was employed for this purpose.

It is important to add that all the significant correlations obtained in this study do not necessarily entail causal relationships. Indeed, correlations establish a relationship between two variables that, if strong enough, is inferred to be less attributable to chance. Any causal interpretations offered on the basis of significant correlations, throughout the presentation and discussion of the results in this paper, are based on the descriptive results presented earlier and the findings of the interviews. A summary of all the correlated motivational and socio-economic component variables is given in Table 2. The N is 142 for the whole data set. It should be noted that in the case of the negative correlations, as the coefficient value of one variable increases or
decreases, the value of the correlated variable moves in the opposite direction. In the case of a positive correlation, as the value of one variable increases or decreases, the value of the correlated variable moves in the same direction.

Table 2: Correlations for the motivational subcomponents and socio-economic variables (N= 142).

<table>
<thead>
<tr>
<th></th>
<th>Cultural Capital</th>
<th>Father’s Past/Present Occupations</th>
<th>Mother’s Past/Present Occupations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Attitudes towards the TL Group</td>
<td>r =0.84**</td>
<td>r = -0.71**</td>
<td>r = -0.48**</td>
</tr>
<tr>
<td>Interest in Foreign Languages</td>
<td>r =0.83**</td>
<td>r = -0.74**</td>
<td>r = -0.52**</td>
</tr>
<tr>
<td>Integrative Orientation</td>
<td>r =0.88**</td>
<td>r = -0.71**</td>
<td>r = -0.57**</td>
</tr>
<tr>
<td>Motivational Intensity</td>
<td>r =0.83**</td>
<td>r = -0.73**</td>
<td>r = -0.55**</td>
</tr>
<tr>
<td>Attitudes towards Learning the TL</td>
<td>r =0.82**</td>
<td>r = -0.71**</td>
<td>r = -0.51**</td>
</tr>
<tr>
<td>Desire to Learn the TL</td>
<td>r =0.84**</td>
<td>r = -0.71**</td>
<td>r = -0.50**</td>
</tr>
<tr>
<td>Instrumental Orientation</td>
<td>r =0.80**</td>
<td>r = -0.75**</td>
<td>r = -0.53**</td>
</tr>
</tbody>
</table>

** Significant at the 0.01 level.

As shown in Table 2, all of the correlations that included the motivational factors show interesting degrees of positive and negative relationships with the socio-economic variables. Clearly, the relationships between the motivational subcomponents and Mother’s Past/Present Occupations, as well as with Father’s Past/Present Occupations, are negative whilst Cultural Capital correlates positively with all motivational subcomponents. All of these relationships are significant at the 0.01 confidence level. Mother’s Past/Present Occupation has a moderately negative correlation with all of the motivational subcomponents. The coefficient values for: Attitudes towards the TL Group, Interest in Foreign Languages, Integrative Orientation, Motivational Intensity, Attitudes towards Learning the TL group, Desire to Learn the TL, and Instrumental Orientation with Mother’s Past/Present Occupation are:  r = -0.48;  r = -0.52;  r = -0.57;  r = -0.55;  r = -0.51;  r = -0.50; and  r = -0.53 respectively. All of these relationships are also moderate except for the one with Attitudes towards the TL Group, which has a weak correlation coefficient. All of these relationships might result from the absence of any influential role of the sample’s mothers, given that the majority are illiterate, and those whose work positions are located at the lower end of the occupational scale.
In contrast, Father’s Past/Present Occupations have a strong, significant, negative relationship with all of the motivational subscales. It is evident from the results that the role of fathers in the students’ families outperforms the mother’s role. The majority of students in the interviews expressed that their fathers played a more influential role in their families than the mothers. They stated that their fathers encouraged and supported their efforts to learn English language. Hence, the correlation coefficients of Father’s Past/Present Occupations with the motivational subcomponents are:

- $r = -0.71$ with Attitude towards the TL Group;
- $r = -0.74$ with Interest in Foreign Languages;
- $r = -0.71$ with Integrative Orientation;
- $r = -0.73$ with Motivational Intensity;
- $r = -0.71$ with Attitudes towards Learning the TL;
- $r = -0.71$ with Desire to Learn the TL; and
- $r = -0.75$ with Instrumental Orientation.

Finally, the Cultural Capital is strongly correlated with the motivational subcomponents. The correlation coefficients for this socio-economic component with the motivational subcomponent variables are:

- $r = 0.84$ with Attitudes towards the TL Group;
- $r = 0.83$ with Interest in Foreign Languages;
- $r = 0.88$ with Integrative Orientation;
- $r = 0.83$ with Motivational Intensity;
- $r = 0.82$ with Attitudes towards learning the TL;
- $r = 0.84$ with Desire to Learn the TL; and
- $r = 0.80$ Instrumental Orientations.

The results on the relationship between parental past/present occupations and all the motivational subcomponents are in keeping with the results presented in the previous section. It can generally be observed that the less skilled occupation a student’s father or mother had/has, the more integrative is the student attitudes towards the Yemeni speakers of English. This finding accords the point raised by Gardner (1985a). He draws our attention to some evidence of stronger parental influence on the learners’ opinions when the families tend to belong to the lower socio-economic stratum. The learners have a more pragmatic attitude towards English language learning.

Clearly, the parents’ occupations are strongly related to the formation of students’ attitudes towards the local TL group. Hence, it is suggested that the TL group for the sample of the present study is represented by Yemeni speakers of English, not native English speakers. This is evident from the students’ responses in the interviews with regard to their perception of the members of the local TL group. The group members’ distinguished social position creates within the students a desire to identify with their habits and practices. For students whose fathers or mothers have lower-ranked occupations, English learning tends to be associated with such characteristics. This supports the assumption of the study that English represents a
prestigious practice typical of high-status groups, at least in the minds of more deprived students. Whilst this is the case, there is nothing conclusive about how fathers or mothers’ involvements in the process of language learning affect students’ motivation, and how these parents with less prestigious jobs may impact the attitudes and motivation of these students. It is difficult to comment critically on the nature of the parental influence from the results of the study, but it seems that the fathers believe more strongly in the value of English for the attainment of high social status. It should also be noted that two of the twenty students interviewed, who were socio-economically better-off than the others, pointed out that English has been a much more widespread phenomenon in Yemen, and that being able to speak the language is not necessarily the privilege of only socially advantaged groups.

Although correlation is not causation, all of the aforementioned significant relationships between motivational component factors and the socio-economic factors are reflective of the students’ integrative tendencies towards the local TL group and their high level of motivation towards learning English in general. Interestingly, two significant correlations are negative, that is, when either one of the parents have or had lower-ranked occupations. In light of this, socio-economic factors are powerful predictors of the students’ integrative tendencies and motivation to learn English. The results also show that cultural capital has a strong, positive correlation with all of the motivational variables. This is despite the fact that the majority of the students in the sample are from families with a parent with a low-ranked occupation, and have smaller amounts of cultural capital. As expressed by the students in the interviews, this may be due to the sample’s knowledge of the importance of cultural capital, and the prevailing culture and linguistic characteristics of the parents.

**Conclusion**

The results of the study reveal that the Yemeni EFL students in the study developed strong integrative tendencies towards the local TL group and high level of motivation to learn English. Since language use is inextricably linked to social class and status, the students’ motivation to learn English is influenced by membership in a certain socio-economic strata (Lifrieri, 2005). It is evident from the study that that social stratification promotes a particular attitudinal identity construction/formation in the sample of this study. Social stratification actually influences all the sample of this study to learn English, observable via all their motivational levels. Socio-economic factors including parental occupation and cultural capital seem to be powerful factors
contributing to the high levels of the students’ motivation to learn English. The results of the study show the Yemeni EFL students with poorer economic backgrounds tend to be more optimistic and motivated. They exhibit consistent intra-group patterns of TL motivation, which points to the reproductive effects of Bourdieu’s habitus.

Csizér and Dörnyei (2005a, p. 29) place the source of learners’ identification/integration process in the internal representations of the self. Given this, the L2 learner devoid of his/her social contextualization, seems to be depicted as randomly desiring to seize certain attributes, which would make him/her agreeable or professionally successful in the society (ibid). This view of the L2 learner constructing an ideal image and dissociating him/her from the social reality (Lifrieri, 2005), has been challenged in the current study by the incorporation of some elements from the learners’ social space. In this study, the sample’s experiences in and with their social environment may have influenced their motivation, expectations and construction of new identities/images. The complex interactions between these dimensions are clearly reflected in the encompassing theoretical and conceptual framework of the current study. Unlike the concept of “possible selves” proposed by Csizér and Dörnyei (2005a), Bourdieu’s habitus, which is adopted in the current study, offers a psychosocially grounded construct to express the situation in which the sample internalizes ways of understanding their social positions and thinking about and reacting to their social world. Bourdieu’s theory contributes well to the current study, as it is based on a mental-social relational perspective (Lifrieri, 2005). Ultimately, the concept of social stratification clarifies aspects of the cognitive formation of Yemeni students’ attitudes towards the local TL group and their motivation to learn English.

It can be stated that socio-economic prestige is a crucial factor behind the strong integrative tendencies and the high motivation levels of the Yemeni EFL students in the sample. The students consider the use of English in their context as the privilege of members in high-status groups, where knowledge of English allows them the access to better employment opportunities and the potential to move up the social scale. They perceive the ability to use English language as a means to social distinction and closer to the TL group.

In conclusion, the results of the study provide further confirmation of the impact of learners’ socio-economic factors on their integrative tendencies in particular and on motivational patterns in general in FL learning. Integrativeness, the study shows, is particularly motivated by social aspirations. And the significant negative correlations
between parental occupations and integrativeness that emerged in the present study
deserve attention and exploration in other FL contexts.

References

Al-Quyadi, A.M. (2000). Psycho-sociological variables in the learning of English in
Yemen. Ph.D thesis, T. M. Bhagalpur University

Bourdieu, P. (1985). The social space and the genesis of groups. Theory and Society,
(14), pp.723-744.

theory and research for the sociology of education. New York, NY: Greenwood
Press.

14-25.

Clement, R., & Kruidenier, B.G. (1983). Orientations in second language acquisition:
The effects of ethnicity, milieu and target language on their emergence. Language
Learning, 41, pp.273-291.

group cohesion in the foreign language classroom. Language Learning, 44,
pp.417-448.

Robinson, (Eds.), The new handbook of language and social psychology. 2nd ed. New

Csizér, K., & Dörnyei, Z. (2005a). The internal structure of language learning
motivation and its relationship with language choice and learning effort. The Modern

British-Yemeni Society.

Language Learning, 40, pp.45-78.

Education Limited.


INTRODUCTION

Dear Students,

Kindly answer these questionnaires as instructed. Please Note that Your Accurate and personal Answers are very important. Further, please note that these questionnaires will not be seen by anyone else. The only one who will have an access to your answers is the researcher. Also, there is no wrong or right answer for the questionnaires. It depends on personal opinion. I cannot overstate the importance of your participation in this study.

Thank You Very Much,

Yours Faithfully,

Rais Attamimi

(1)-- Motivation Questionnaire

Instructions: Please read each of the statements given below carefully. Then for each statement, decide the extent to which you agree or disagree with it by circling the appropriate number based on the scale that is provided below:

Strongly Disagree
Moderate Disagree
Slightly Disagree
Neither Agree nor Disagree
Slightly Agree
Moderately Agree
Strongly Agree
Please, read the following statements and put a circle around only **ONE** (1) of the **SEVEN** options on the right.

1. Studying English is important for me in order to make new friends. 1 2 3 4 5 6 7

2. I would like to be able to talk with other people in English. 1 2 3 4 5 6 7

3. The Yemenis who speak English live in nicer neighbourhoods. 1 2 3 4 5 6 7

4. I would like to learn English well. 1 2 3 4 5 6 7

5. Studying English is important for me because with English I can do better in my other classes. 1 2 3 4 5 6 7

6. Studying English is important for me because my friends listen to songs in English. 1 2 3 4 5 6 7

7. I don’t have to work very hard to get my English assignments right. 1 2 3 4 5 6 7

8. The Yemenis who speak English have a lot of money. 1 2 3 4 5 6 7

9. I like to watch TV programmes in English to practise. 1 2 3 4 5 6 7

10. Studying English is important for me because I will be able to earn more money when I work. 1 2 3 4 5 6 7

11. If I could choose, I would not study English. I would choose another language. 1 2 3 4 5 6 7

12. When I finish university, I will stop studying English. 1 2 3 4 5 6 7

13. I study English at home to learn faster. 1 2 3 4 5 6 7

14. The Yemenis who speak English are very intelligent. 1 2 3 4 5 6 7

15. When my English lecturer returns my corrected assignments, I put them away without looking at the mistakes. 1 2 3 4 5 6 7
16. I want to have more hours of English in class to learn more.

17. Studying English is important for me to continue with my postgraduate studies later.

18. If there was not an English Department at my college, I would study English in another college anyway.

19. Of all the languages, English is the one I like the least.

20. I would like to read books in English.

21. Studying English is important for me because students from other colleges learn English.

22. I love to study English.

23. In class, I prefer that the teacher speaks only in English.

24. The Yemenis who speak English drive expensive cars.

25. Studying English is important for me because my friends in class also speak English.

26. If I have trouble understanding something in the English class, I seek help right away.

27. I hate English.

28. When I hear a song in English, I pay a lot of attention to try and understand the lyrics.

29. Studying English is important for me because it makes my parents happy.

30. I’m going to continue studying English.
Appendix B
The Socio-economic Questionnaire

The Socio-economic Questionnaire

Instructions: Thank you for agreeing to participate in this survey. Please take your time and read each question carefully. Think about what is being asked before you answer. Be sure to use your logic and choose and write the most appropriate answer that most corresponds to your perception of your current socio-economic status. Your responses will be anonymous. Findings of the survey will be reported as aggregate (as a group) only; there will be no way anyone can trace back a particular response to you.

Good luck!

SOCIO-ECONOMIC QUESTIONNAIRE

1. (a) Does your father work?

Circle the appropriate answer.

~ YES.

~ NO.
(b) What does your father do?

Write your answer in the space below.

(c) If he is unemployed, what was his last job?

Write your answer in the space below.

2.

(a) Does your mother work?

Circle the appropriate answer.

~ YES.

~ NO.

(b) What does your mother do?

Write your answer in the space below.
(c) If she is unemployed, what was her last job?

*Write your answer in the space below.*

<table>
<thead>
<tr>
<th>3.</th>
</tr>
</thead>
</table>

(a) Does your father read the newspaper/s?

*Circle the appropriate answer*

~ YES.
~ NO.

(b) If your answer was ‘YES’, how often does he read the newspaper?

*Mark the appropriate answer.*

~ Every day.
~ Twice or three times a week.
~ Only on Fridays.

(c) If your answer was ‘YES’, what newspaper/s does your father read?

*Write your answer in the space below.*
4.

(a) Does your mother read the newspaper/s?

*Mark the appropriate answer.*

~ YES.

~ NO.

(b) If your answer was ‘YES’, how often does she read the newspaper?

*Mark the appropriate answer.*

~ Every day.

~ Twice or three times a week.

~ Only on Fridays.

(c) If your answer was ‘YES’, what newspaper/s does she read?

*Write your answer in the space below.*

---

**YOU HAVE REACHED THE END.**

Thanks a lot for your cooperation!
An Analysis of L2 Motivation, Test Validity and Language Proficiency Identity (LPID): A Vygotskian Approach

John F. Haggerty
Seoul National University, Seoul, South Korea
john2teach@gmail.com

Bio Data
John F. Haggerty is currently a Visiting Professor and Teacher Trainer at Seoul National University, College of Education. For the last decade, he has been working in South Korea as an EFL/EAP instructor and teacher educator. He has worked at Sogang University, Yonsei University and Chonnam National University as well as a number of private language academies. In addition, he has been a speaking test rater and test writer for the last five years.

Abstract
This paper explores the potential impact of high-stakes English testing on young English language learners’ (ELL) attitudes, beliefs, and motivations. A more meaningful role for consequential validity in language testing is sought through engagement with sociocultural theory, specifically Vygotskian conceptions of identity formation, in order to more fully contextualize a high-stakes learning environment, its effects on younger ELLs, and its implications for the test validation process. This initial phase, to be followed by a larger study, consisted of a pilot questionnaire developed from a working model of Language Proficiency Identity (LPID), and subsequently administered to 202 ELLs of various education levels. An analysis of responses to 20 questionnaire items was conducted on the basis of a) education level, and b) the completion of one of two high-stakes English tests. There were significant correlations found for 15/20 responses from middle-school students, 4/20 from high school students, and 0/20 from university students. These preliminary results suggest that high-stakes English testing has a more dramatic effect on the language proficiency identity of younger ELLs. Some implications for test validity are discussed in light of current theory and research on identity formation.

Keywords: Language Testing; Test Validity; L2 Motivation; Identity; Sociocultural Theory

1. Introduction
My research interest stems directly from my experience preparing Korean L2 learners of various education levels for cognitively demanding high-stakes tests of English proficiency. Over the
years, I have become increasingly concerned about the effects of intensive test preparation on young learners, especially those who have begun preparing for language proficiency tests as early as elementary school. Throughout my teaching career in Korea, I have been enormously impressed by the willingness of most students to try to learn what is required of them, no matter how advanced it may be, and focus all their effort into achieving the highest standard possible. On the surface, they appeared to cope with their ubiquitous and demanding testing environment surprisingly well given the vital role it seems to play in their future success. When asked, most of my students have expressed general agreement with the benefits of testing, acknowledgement of its necessity, or a quiet acceptance of its reality. Some, frequently older learners, have expressed critical opinions or objections, but these have been rare despite my best efforts to uncover hidden sentiments. However, as a foreigner, my perceptions have inevitably been superficial at best and despite the dramatic social, political, and educational power of tests (Shohamy, 2001) evident in this learning environment, it has been difficult for this outsider to truly appreciate the impact of high-stakes testing on language learners. This study offered me an important opportunity to dig a little deeper into the attitudes, beliefs, and motivations of Korean language learners and attempt to better understand what effects high-stakes tests might be having on ELL socio-cognitive development.

In South Korea, English language learning and testing is pervasive and powerful. The implementation of the 7th National Curriculum in 2000 (which further expanded the role of English education in public schools), the breadth and proliferation of private language institutes, the high percentage of family income spent on English language education, and the continued rise of English language requirements as gatekeeping devices for education and employment, are all testaments to the significance and magnitude of English language learning in Korea (for a fuller discussion, see Nam, 2006 and Choi, 2008). For many, the results of language tests can dramatically influence the opportunities available to them and their ability to parlay these opportunities into success in a highly competitive education and employment market. As such, there is extreme pressure to perform on English tests even at quite young ages, some very young ELLs even taking cognitively-demanding English proficiency tests like the TOEFL as early as elementary school (Choi, 2008, p. 53).

The Test of English as a Foreign Language (TOEFL) created by the Educational Testing Service in the U.S has been used in South Korea for years as a benchmark (amongst other
factors) for acceptance to most universities and foreign language high schools (Choi, 2008, p. 45). More recently, the Test of English Proficiency (TEPS) created by Seoul National University in South Korea has risen in prominence to compete with the TOEFL in the language testing market. Both of these tests purport to measure the academic skills necessary to perform in English at the university level. Despite the reality that young English language learners (ELLs) are taking these cognitively demanding tests, designers have not adequately investigated test validity with specific reference to young ELLs. However, if modern notions of test validity are to be taken seriously, much more investigation into the nature and impact of high-stakes tests on younger ELLs is needed. While there are many facets of test validity that need to be addressed with respect to young ELLs, this study will focus on the social consequences of test use (Messick, 1989a, 1989b, 1995) by exploring the potential effects of cognitively demanding tests on younger ELLs. In attempting to address this issue, the following three research questions were considered:

- What differences exist in the attitudes, beliefs, and motivation of young ELLs who have completed high-stakes language tests compared to those who have not?
- Do these differ from ELLs at higher education levels?
- Can high-stakes test results be considered consequentially valid for young ELLs and if so, at what age?

This paper hopes to stimulate further discussion and research into these questions by considering the potential impact of high-stakes testing on ELLs at three different education levels: middle school, high school, and university. A great deal of research has been conducted on the role of ELL attitudes, beliefs and motivation and its impact on the achievement of language proficiency (for an extensive review of learner motivation studies, see Dörnyei, 2001). However the role of high-stakes testing in helping or hindering attitudes, beliefs, and motivations has received scant attention in motivation studies. Given the intensive role that high-stakes standardized testing plays in many meritocratic educational environments, particularly in Asia (Ross, 2008), and its growing importance resulting from educational initiatives such as ‘No Child Left Behind’ in the U.S., this is a consideration that has been overlooked for too long. This paper explores how ELL attitudes, beliefs and motivations may be influenced by the high-stakes testing environment in South Korea. Through engagement with Vygotskian conceptions of identity formation, the
potential socio-cognitive effects of high-stakes English tests on students of different education levels will be explored and some implications for the consequential validity of these tests will be offered. Admittedly, this will be a mere scratch of the surface; however, it is hoped that this might spur further investigation and, hopefully, a greater understanding of how high-stakes language testing influences young language learners around the world.

2. Modern Concepts of Test Validity

There are two concepts that together form the major impetus for this paper: validity in language testing and the formation of language proficiency identity (LPID). To be sure, this is an ambitious and challenging alliance to make and one that undoubtedly warrants some explanation. To begin with, it is important clarify the nature of test validity being considered in this paper. Not to be confused with reliability (primarily concerned with the ability of assessments to accurately and consistently reproduce test scores), validity requires theoretical rationale and empirical evidence to establish the appropriateness of test interpretation and use (Messick, 1989a). Three major aspects of validity (content, criterion-related, and construct) have most often been investigated in traditional test validation studies (Crocker & Algina, 1986). In general terms, these concepts explore: 1) the concordance of test content with the subject being assessed (e.g., university-level academic text); 2) the ability of test scores to concur or predict performance on some external criterion (e.g., success in an academic reading course); and 3) the strength of the theoretical constructs (on which the test is hopefully based) to measure what they purport to measure (e.g., the ability to read academic text at a university level). While these remain indispensible concepts in the test validation process, few modern test theorists would argue that they are sufficient to properly validate a test alone.

Perhaps the most influential modern concept of validity in language testing has been put forward by test theorist and psychometrician Samuel Messick (1989a, 1989b, 1995), in which he argued for a significant expansion of the concept of validity. His progressive ‘validity matrix’ (see Figure 1 below) merges four major facets of testing into an inseparable symbiotic relationship, comprised of construct validity, test relevance and utility, the value implications involved in test design and interpretation, and the social consequences of test use.
Some have incorrectly assumed that this matrix permits the isolation of each cell to the exclusion of others, suggesting that construct validity research can somehow occur in a “social and value vacuum” (McNamara, 2001, p. 336). However, this is clearly not what Messick had intended. As a ‘unified concept’, he was arguing that validity in language testing involves value judgements from the very first decision made, the implication being that “both meaning and values, as well as both test interpretation and use, are intertwined in the validation process” (Messick, 1995, p. 749). These value judgements have real-world social consequences that need to be properly considered in order for a unified vision of test validity to be properly satisfied. Messick’s unified vision of validity, placing the social consequences of testing at the apex, represented a revolution in thinking about test validity and still continues to stir up a great deal of debate about the practical limits of test designers’ social responsibility.

Another influential reformulation of the test validation process has been put forth by Bachman and Palmer (1996) in their proposed model of ‘test usefulness’. Six test qualities, comprised of reliability, construct validity, authenticity, interactiveness, impact, and practicality, were proposed as a pragmatic method for evaluating the strengths and weaknesses of a test. While impossible to maximize all qualities, this framework could enable test designers to formulate an argument that, on the whole, supported the use a test for a specific purpose. Within the framework, construct validity was separated from the other more social factors involved in test interpretation and use (i.e. authenticity, impact, and practicality) in order to provide a more
pragmatic test validation process. However, Bachman himself later conceded that taxonomies of test qualities such as this one essentially provide only a “list of independent qualities” to consider and do not “bring us any closer to a coherent theory of test use” (2005, p. 4). Instead, borrowing from Kane (1992, 2002), he later suggested the utilization of an ‘assessment use argument’ in the format of a Toulman argument (consisting of claims, warrants, and rebuttals) in order to more explicitly address the connection between validity and test use (Bachman, 2005). This approach acknowledged the need for test designers to continually understand and respond to the social effects of test use as part of their validation process.

Other modern theorists (for example, McNamara, 2001, 2005; Shohamy, 2001; Kunnan, 2005; McNamara & Roever, 2006; Fulcher, 2009; and others) have also challenged the field to move one critical and controversial step forward by more thoughtfully engaging advancements in social theory and/or more actively investigating the wider social context surrounding the design and use of language tests. They have warned of the dangers of claiming impracticality, avoiding responsibility, and of relying on complex statistical techniques alone to justify the interpretation and use of tests. While there is still a great deal of debate about the theoretical applicability of modern social theory to language testing and the practical limitations involved in properly accounting for the wider social context, there appears to be more agreement than ever with the sentiment expressed by McNamara and Roever (2006, p.1) that “… through marrying itself to psychometrics, language testing has obscured, perhaps deliberately, its social dimension” and that this has now become an indefensible position.

Some significant inroads have been made to help reveal the social dimension of language testing. A few notable studies include investigations into differences in language test perception amongst L1 and L2 learners (Fox & Cheng, 2007), the impact (or washback) of language testing in the educational environment (Cheng, 2005), tests as indicators of linguistic identity (Eades et al, 2003), the role of language tests in situations of intergroup competition and conflict (McNamara, 2005), and the potential human rights violations that can result from the use of language tests (McNamara & Shohamy, 2008). These, and other studies of the sort, represent valiant attempts to more fully contextualize the wider social environment in which language tests are deployed. They represent the reality that tests do not, indeed cannot, exist in a theoretical vacuum. It is a reminder that, as always, “validity judgements are value judgements” (Messick, 1995, p.748). However, in order to realize the full implications of Messick's matrix, the field
must continue to build on past efforts. It must continue to push the boundaries of what has been considered practical or appropriate in the past. One way this can be done is to more meaningfully engage the developments that have occurred in social theory, and more specifically, postmodern and poststructuralist concepts of self and society (McNamara, 2001). Insight into the potential effects of high-stakes language testing on the identity formation of young ELLs offers a unique opportunity to more fully engage social theory and further contextualize the social dimension of language testing (McNamara & Roever, 2006).


Identity is indeed an elusive and highly controversial concept. This should come as no surprise given that conceptions of identity attempt to address two of the most difficult questions we face as human beings: who we are and how we come to be. In the area of SLA, dissatisfaction with traditional concepts of identity prevalent in early research began in earnest in the 90s (e.g., Norton Pierce, 1995; Lantolf, 1996) and culminated in an oft-cited article by Firth and Wagner in 1997. While this article was perhaps more symptomatic of developments already underway in the field (Block, 2007, p. 872), it certainly stirred up a great deal of debate and challenged SLA researchers, inter alia, to move beyond overly simplistic notions of language acquisition and identity. It was argued that the traditional bifurcation of language into that produced by/in the individual and that produced by/in society, and the subsequent apathy expressed towards the latter, had lead to a myopic and stagnant theoretical state of affairs as well as an unwarranted distortion of the way language actually operates in the real world.

Indeed, studies into individual language cognition up to that time had shed considerable light on crucial aspects of SLA; however, there had been an unfortunate and dangerous tendency to ascribe error patterns, motivations, attitudes, and ultimately identities to the individual while ignoring the social environment that plays an integral part in its formation. For many researchers in this tradition, identity was a “taken-for granted resource, rather than, or as well as, a topic of investigation” (Firth & Wagner, 1997, p. 288), whereby essentially only “one identity really matter[ed], and it matter[ed] constantly and in equal measure throughout the duration of the encounter being studied” (Firth & Wagner, 1997, p. 292). This simplistic notion of identity as a priori, asocial, and ahistorical, while convenient for the researcher, drastically misrepresented the nature of identity and identity formation. What was needed, according to Firth and Wagner (1997) was a more holistic approach that “problematizes and explores the conventional binary
distinction between “social” and “individual” (or cognitive) approaches to language use and language learning” (p. 296).

A fuller explanation of SLA would need to encompass the individual cognition of language learners, the social environment in which language is exchanged, and the interaction between these two inevitably symbiotic realms. What was required was a “reconceptualization of SLA as a more theoretically and methodologically balanced enterprise that endeavours to attend to, explicate, and explore, in more equal measures and, where possible, in integrated ways, both the social and cognitive dimensions of S/FL use and acquisition” (Firth & Wagner, 1997, p. 286). At the very least, this would require greater attention be paid to the social environment in which language lives and breathes. For conceptions of identity, conveniently simplistic and static assumptions would need to become considerably more complex.

In the decade that transpired after the publication of Firth and Wagner’s (1997) article, the authors argue that for a great deal of mainstream SLA research, it has largely been business as usual. Yet, the momentum which initially began in the 80s and 90s has continued to build steam and has pushed SLA research into new and revealing areas of inquiry, including social-interactional approaches, constructivism, and sociocultural theory. One unfortunate result of this expansion has been the further bifurcation of the field into those investigating SLA from a strictly cognitive perspective and those employing from more varied sociocultural/socio-interactional approach (Firth & Wagner, 2007, p. 804). While the field continues to struggle with these important epistemological concerns, the future direction of SLA remains to be seen and it is currently unclear whether it will be able to “withstand the current bifurcations, competing methods, critiques, and internal tensions, and remain generally cohesive” (Firth & Wagner, 2007, p. 813).

While a unified SLA theory that satisfies both these splintered camps may be next to impossible, there is certainly much to be gained by adopting elements from both whenever and wherever possible. One theoretical approach which attempts to incorporate aspects from both sides of this cognitive/social divide was developed by Lev Vygotsky’s through his socio-cognitive conception of learning and identity formation (Vygotsky, 1978). The following model of language proficiency identity (LPID) was largely inspired by these concepts of identity formation. By incorporating this approach, it is hoped a greater harmony can be achieved.
between the individual cognition of language learners and the social environment that impinges upon them.

4. A Working Model for LPID

In a recent attempt to more meaningfully engage social theory, Lazaraton and Davis (2008) examined the potential role that conceptions of identity play in the oral performance of L2 speakers in paired speaking assessment. They advanced the concept of ‘language proficiency identity’ (or LPID) and by utilizing discourse analytic techniques, they explored the notion that speakers bring a conception of themselves as more or less proficient, and that this plays a significant role in the nature of the interaction that occurs in paired assessment. They argue that the “test taker comes to a paired task (if not all speaking tests) with a language proficiency identity, which is constructed, developed, and maintained in and through the discourse in which they are taking part” (Lazaraton & Davis, 2008, p. 324). Greater insight into the nature, strength, and effect of LPID, they argue, will help to shed light on the potential threat this poses to test validity in paired assessment.

The analysis of LPID in Lazaraton and Davis' article was restricted to a specific performance, the paired speaking assessment. However, if LPID is indeed a useful construct, it is reasonable to suppose that its development has a much longer and more complex history within the individual. In the socio-cognitive approach to cognitive development advanced by Lev Vygotsky, this would involve moving from a microgenetic perspective, or the analysis of a specific process occurring during an individual’s development, to an ontogenetic one involving the development of an individual over time (Lantolf & Thorne, 2006, p. 29). Nowadays, many theorists speak of multiple identities that are always in a state of flux, often in conflict with each other, and continually changing in response to new social and political forces (Norton, 2000). These are enormously complex relationships that are nicely captured by Vygotskian conceptions of socio-cognitive learning and identity formation. Unfortunately, space does not permit a thorough examination of Vygotskian theory, but for those who wish further information, Thorne (2005) puts forth a persuasive case for the “methodologically, politically, and ethically vigorous tools” that are made available with this approach (p. 402) and Lantolf and Thorne (2006) have provided a thorough explication of the major tenets of Vygotskian theory and their practical applications to the field of SLA.
Figure 2 below is a concept map for LPID that attempts to incorporate Vygotskian concepts of socio-cognitive development and will hopefully assist in the development of a more robust ontogenetic LPID framework in the future. This is a first stab at peeling back a few layers of LPID and will undoubtedly continue to adapt as the concept is explored further. It is hypothesized that two major factors, *assessment experiences* and *learning experiences* (represented as large ovals) will figure prominently in the nature and development of LPID.
The main categories, assessment and learning experiences, are further broken down into eight sub-factors that are theorized to have a significant influence (represented as small ovals). These factors were conceptualized along two social-psychological planes that exist within the individual language learner. Vygotsky proposed that both an external (intrapsychological) and internal (interpsychological) plane was “necessary for human thinking to emerge and develop” (Lantolf & Thorne, 2006, p. 153). Conceptualizing LPID in this way, as a kind of ‘mindscape’ that is permeated by both social and individual cognitive factors, permits us to expand the cognitive lens in significant ways. While still largely an individualistic conception, it does acknowledge and attempt to incorporate the social environment impinging upon belief, attitudes, and motivations, and is, at the very least, a step in the right direction.

This model is theorized to operate within a much larger framework of identities that is, as Bonny Norton (2000) and others have observed, always in a state of flux as an individual traverses time and space. This larger framework is also couched within an external world of activity that differentially impinges upon each individual for various economic, social, political and historical reasons. Given this complexity, one can begin to understand the difficulty modern theorists face in specifying a precise model. The artificial separation created here is not intended to re-bifurcate, but to better visualize the interconnectedness that exists in each individual, between how we perceive the world around us and how we perceive ourselves. A brief explanation of each subcategory follows.

4.1 Assessment Experiences

While there has been much written about the importance of language learning experiences in the attainment of L2 proficiency, the role of assessment in this process has largely been subsumed and given a subsidiary role or excluded from consideration altogether. This is somewhat surprising given the traumatic and lasting experiences that are often reported by ELLs based on assessment episodes (for some particularly poignant examples, see Shohamy, 2001, Ch. 2). For LPID, especially in this testing intensive environment, assessment experience is theorized to play a much larger role, and as such, it has been disentangled from learning experiences and given an equally pivotal role in the development of ELL proficiency identity. Four main inter/intra sub-factors guided the development of this exploratory construct and will be explained in more detail below.
4.1.1 Performance/Signification

*Performance* is the external behaviour of a language learner that is viewed (and reacted to) by others and can occur on a variety of social stages, including the completion of proficiency or achievement tests, class presentations, or even classroom conversations with teachers or colleagues. These external performances, and the assessments they engender, send vital signals to language learners about how ‘successful’ ELLs have been in their performance. *Signification* is the extent to which these external performances and their assessments are internalized, and to varying degrees, accepted as meaningful and valid. ELLs may place great significance on their performance based on a strong belief in the value of external opinion. Alternatively, they may de-signify their performance for a variety of reasons including, but not limited to, distrust of others, a lack of belief in the value of someone’s opinion, or the desire to avoid activity they are not very good at doing or they feel is too stressful. Although performance and its signification includes a wide variety of external behaviours, performed for a variety of reasons, it was decided to focus on English test performance for the purposes of this study.

4.1.2 Other Evaluation/Self Evaluation

For some performances, there are higher stakes involved as they are judged, ranked and labelled for the learner. These are often dramatic moments in one’s language learning trajectory and can place an indelible stamp on a learner. Despite their significance, they have often been relegated to the sidelines within motivation studies. *Other evaluation* is comprised of the external formalized judgments bestowed upon us by various organizations assigned with the task. Formal performances (for example in standardized testing) are often objectified in order to be compared and ranked. These performances are normalized or set against an established benchmark often to grant access to educational or occupational resources. These evaluations are likely to play a significant role in the assessment experiences of many ELLs. High-stakes testing, due to its formalized and standardized nature, may be viewed, and is invariably presented, as an authoritative and objective measure of proficiency or achievement. Stakeholders often place great importance on these evaluations. On the other hand, *self evaluation* is an internal, subjective capacity that, although no doubt influenced by external evaluation, reflects the evaluations we make about ourselves. As such, there can be varying degrees of resistance to formal evaluations. This resistance can take the form of distrust in test format, scoring,
interpretations of scores and/or the uses to which they are put. Some ELLs may even feel they are able to evaluate their own performance far better than any external test.

4.2 Learning Experiences

The learning environment of ELLs has been given a great of attention in studies of learner attitudes, beliefs, and motivations, especially since an ‘educational shift’ in L2 motivational studies occurred in the early 90s that encouraged researchers to emigrate away from traditional psychologically-inspired motivational concepts to more ‘education-friendly’ ones (Dörnyei, 2001, p. 104-105). This wealth of research has helped to reveal the wider social context of learning and the vital role it plays in motivation. For the concept of LPID, it is theorized that the learning experiences of ELLs, comprised of (at least) the four main inter/intra sub-factors described below, will be highly influential in the learning process.

4.2.1 Social Value/Comfort

The social value of any learning endeavour will likely be influenced by an abundance of external semiotic and linguistic resources. These external cues influence our conception of social value and, in turn, the degree of external reward available. These could vary from academic and employment requirements, media reports, advertisements, images, government programs and publications, or familial and collegial conversations to name a few. Comfort reflects the degree of stress and anxiety experienced while learning a language. Undoubtedly, the degree of comfort felt will be influenced by a number of factors such as our acceptance of external evaluations, our confidence in performing, our reasons for learning, our learning environment, and even our biological makeup. However, it is hypothesized that for LPID, the level of comfort in learning will be heavily influenced by one’s conception of social value combined with one’s perceived ability to obtain it. Generally speaking, we are likely to feel more anxious when the social stakes involved are high. As our perception of social value increases, a lack of faith in our performance level will likely increase this discomfort as well.

4.2.2 Aspiration/Inspiration

Aspiration is the external objective for any learning endeavour and can be influenced by, amongst other things, a desire for employment or promotion, acceptance to an organization or academic establishment, or a need to please one’s parents. This has an inter-dependent and
symbiotic relationship with our *inspiration* for learning. The reasons we are inspired to learn are not always easy to define and often become intertwined with our aspirations. Our inspiration for learning might involve what would remain if our external goals could be factored out, which of course they cannot. Inspirations could include a desire to do what one gains pleasure in doing, to explore one’s creative ability, or fulfill one’s adventurous desires. Within motivation theory, this relationship is perhaps more adequately captured by the “L2 motivational self-system” developed by Dörnyei (2009) which argues for a dialectical relationship between our ‘ought to L2 self’ and our ‘ideal L2 self’. These conceptions of our ‘future self’ affect attitude, motivation, and behaviour in interconnected ways that may hold more explanatory power in EFL/ESL contexts than the more widely known concepts of instrumental and integrative motivation developed by Gardner and Lambert (1972) that seem to fall short in describing the complex relationships that exist between power, identity, and language learning (Pierce, 1995, p. 17).

Lastly, it is worth mentioning that the factor pathways for this LPID concept map, including linkages which are not explicitly shown, are intended to be *omnidirectional* and, as such, can influence any other sub-factor at any time in one’s cognitive development. The utility of these sub-factors, and the relative strength of their relationships, requires much more investigation involving far more sophisticated analyses than were possible for this study. There are undoubtedly additional sub-factors that influence LPID as well, but these hopefully serve as a jumping off point towards a more complete model. For this exploration, these sub-factors served as an exploratory theoretical framework for the development and analysis of questionnaire items. While the LPID sub-factors were helpful in the development of potential questions, the main categories of ‘assessment experiences’ and ‘learning experiences’ are the basis by which the overall scales were developed and for which internal consistency was measured.

5. Method

Using the LPID concept map as a guide, a list of over 40 potential questions was initially created. The questions were designed to tap into different aspects of each LPID inter/intra sub-factor being explored and to hopefully contribute to the larger constructs of ‘assessment experience’ and ‘learning experience’. While some questions may be said to focus more on one psychological plane than the other, no attempt was made to specify questions for each plane; instead, inter/intra sub-factors were treated as unified. The questions were first translated into Korean and then ‘back-translated’ (Brislin in Dörnyei, 2002, p.51) into English using two
separate translators. Then, a team-based approach was utilized to assess the suitability of the translated questions (Dörnyei, 2002, p.50). A small group including two Korean EFL teachers and two Canadian EFL teachers (including myself) reviewed the translated questions and identified and discussed any questions we felt were awkward or redundant. Based on these results, some questions were eliminated and some re-worded and re-translated. A total of 32 questions remained after this process was completed.

The next stage involved pre-testing the questions with two classes of university students (N=62) in order to assess the internal consistency of the two main scales (assessment and learning experiences). The ‘reliability’ feature of SPSS (v. 16), with the ‘scale if item deleted’ option checked, was utilized for this purpose. Questionnaire items that did not display enough variability or did not contribute significantly to their scale were eliminated. However, this was also balanced against the need to include questions inspired by all four sub-factors. As a result of this process, the questionnaire was narrowed down to 20 questions (10 per category) with ‘assessment experiences’ achieving a Cronbach Alpha score of .76 and ‘learning experiences’ achieving a score of .72. This indicated reasonable internal consistency for the main categories. This process helped to shorten the questionnaire for young learners who, it was felt, would have trouble maintaining their focus. Due to time and budgetary considerations, this pre-test sample was fairly small; however, it provided support for the intelligibility of the translated questions and for the potential utility of the main categories.

The 20 selected questionnaire items, along with questions about the respondents age, education level, English language proficiency, and TOEFL/TEPS completion, comprised the final pilot questionnaire and a four-point Likert scale (1=strongly disagree to 4=strongly agree) was utilized in order to avoid neutral responses. The questionnaire was administered over a 5-day period in Seoul, South Korea. Permission was sought and given to administer the questionnaire on-site in two middle schools and two high schools in the same socio-economic area. As this was part of a larger research project to be completed at a later date, an initial target of 200 total participants, a minimum of 50 per education level, was set for this stage of the study. Middle and high school students were randomly asked to complete the questionnaire as their classes were dismissed and they were leaving the school. This more personal ‘one-to-one administration’ (Dörnyei, 2002, p. 67) was chosen over the selection of a few large classes in order to make a more personal connection to each student and explain the purpose of the study,
but also to obtain a greater cross-section of the students in the school and reduce the potential for classroom bias (e.g. the principal choosing only highly motivated classes to participate). University students from two locations were also randomly approached on campus, given an explanation of the research purpose, and asked to complete the questionnaire. These universities were located in the same socio-economic area as the middle and high schools. No rewards were promised or given to any of the participants to promote voluntary participation. Respondents were not asked if they had completed any specific English tests prior to completing the questionnaire nor were any questionnaires rejected on this basis. Some students, largely male, declined to participate, but in general the response rate was quite positive at over 80% for all education groups. The data collected were inputted into SPSS software (v.16) and subsequently analyzed.

6. Participants

There were a total of 202 respondents. The youngest respondent was 12 and the oldest was 38 (the mean age was 17.5). Of these, 65 were in middle school, 83 in high school, and 54 in university. An overwhelming majority in each education group (76 % or above) reported their English proficiency to be low to high intermediate with the remainder fairly evenly split between beginner and advanced. Although there were more than two times as many female respondents as male, they were represented fairly equally (approx. 80%) in each education group. Table 1 below lists the number of students in each education level who had and had not taken at least one TOEFL or TEPS test. As a group, university students were more likely (48%) to report having completed one of these tests than those in high school (29%) or middle school (32%).
Table 1: TOEFL/TEPS Completion by Education Level

<table>
<thead>
<tr>
<th>EDUCATION LEVEL</th>
<th>TOEFL/TEPS TAKEN</th>
<th>TOTAL</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>NO</td>
<td>YES</td>
</tr>
<tr>
<td>Middle School</td>
<td>(44)</td>
<td>(21)</td>
</tr>
<tr>
<td></td>
<td>68%</td>
<td>32%</td>
</tr>
<tr>
<td>High School</td>
<td>(59)</td>
<td>(24)</td>
</tr>
<tr>
<td></td>
<td>71%</td>
<td>29%</td>
</tr>
<tr>
<td>University</td>
<td>(28)</td>
<td>(26)</td>
</tr>
<tr>
<td></td>
<td>52%</td>
<td>48%</td>
</tr>
<tr>
<td>TOTAL</td>
<td>(131)</td>
<td>(71)</td>
</tr>
<tr>
<td></td>
<td>65%</td>
<td>35%</td>
</tr>
</tbody>
</table>

7. Results

The primary variables under consideration for this paper were education level, TOEFL/TEPS completion and the 20 questionnaire items based on the LPID categories. When a post-hoc reliability analysis (including all education groups) was performed on the main categories, both achieved reasonable internal consistency with Alpha scores of .75 (assessment experiences, 10 items) and .71 (learning experiences, 10 items). However, the results were mixed when each education group was considered separately. Internal consistency was quite strong for the MS group data with categories achieving Alpha scores of .81 (assessment experiences) and .83 (learning experiences). However, internal consistency for the HS (.69/.66) and UNI (.67/.61) groups was less impressive, with Alpha scores failing to surpass the .70 threshold (as recommended by Dörnyei, 2002). This indicated that further modification would be necessary for the categories to attain reasonable internal consistency across all three education groups.

Since this analysis also required splitting the data into education levels, and then further into TOEFL/TEPS completion, the resultant sample sizes became too small for more advanced statistical techniques to be utilized (e.g. factor analysis or SEM). However, an analysis of the correlations between each education/test group and the questionnaire items would provide valuable exploratory information that could better inform the design of the larger study. Therefore, point-biserial correlation coefficients were calculated for each of the 20 questionnaire items and whether respondents had completed one of the two high-stakes English tests (the TOEFL or TEPS). Table 2 below lists the correlations found for each education/test grouping.
Overall, 15 significant correlations were found for the middle school group, \( p < .05 \), while only 4 were found for the high school group and none for the university group.

The following analysis will examine questions inspired by each LPID subcategory separately. Selected items (2 per category) based on the strength of their associations and will be presented visually in bar charts. To more fully describe these relationships, the correlation coefficients for each education group will be included directly above their respective relationships in the bar chart. These results will then be further analyzed and some initial conclusions offered.

**TABLE 2: Point-biserial Correlations for LPID Questions and TOEFL/TEPS Completion**

<table>
<thead>
<tr>
<th>QUESTIONS</th>
<th>TOEFL / TEPS TAKEN (No/Yes)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Middle School</td>
</tr>
<tr>
<td><strong>Assessment Experiences</strong></td>
<td></td>
</tr>
<tr>
<td>I value the results of my English tests</td>
<td>.533**</td>
</tr>
<tr>
<td><strong>English tests have made me feel discouraged about learning</strong></td>
<td>.517**</td>
</tr>
<tr>
<td>My score on English tests reflects my English ability</td>
<td>.501**</td>
</tr>
<tr>
<td>I think the English tests I have taken so far are fair</td>
<td>.453**</td>
</tr>
<tr>
<td>English tests allow students to see how well they are doing in English</td>
<td>.348**</td>
</tr>
<tr>
<td>Students deserve the score they receive on English tests</td>
<td>.294*</td>
</tr>
<tr>
<td>Students should trust English test scores</td>
<td>.286*</td>
</tr>
<tr>
<td><strong>English tests do not predict English language ability in the future</strong></td>
<td>.262*</td>
</tr>
<tr>
<td>I must study harder in order to do well on English tests</td>
<td>.195</td>
</tr>
<tr>
<td>I am the best judge of my English ability</td>
<td>-.042</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Learning Experiences</strong></td>
<td></td>
</tr>
<tr>
<td>I learn English because I enjoy it</td>
<td>.579**</td>
</tr>
<tr>
<td><strong>I feel stressed when using English</strong></td>
<td>.512**</td>
</tr>
<tr>
<td>Students should be required to learn English</td>
<td>.411**</td>
</tr>
<tr>
<td>Studying English has had a positive effect on my learning experience</td>
<td>.387**</td>
</tr>
<tr>
<td>Universities and companies should value English ability</td>
<td>.363**</td>
</tr>
<tr>
<td>English tests are an important reason I study English</td>
<td>.302*</td>
</tr>
<tr>
<td>It is important for Korean students to use English well</td>
<td>.299*</td>
</tr>
<tr>
<td>Students must study English if they want to be successful</td>
<td>.217</td>
</tr>
<tr>
<td>English is important in Korea</td>
<td>.041</td>
</tr>
<tr>
<td><strong>I worry about my English test scores</strong></td>
<td>.030</td>
</tr>
</tbody>
</table>

Note: Negatively worded items (in italics) were transformed before analysis

\*p<.05, two-tailed; 
\**p<.01, two-tailed
7.1 Performance/Signification

**Performance Value of Tests**

- *p<.05, **p<.01

**Encouragement of Tests**

Figure 3 above displays the bar charts and correlation coefficients for the questionnaire items ‘I value the results of my English tests’ and ‘English tests have encouraged me to learn English’, grouped by education level. Middle school (MS) students who had completed a TOEFL or TEPS (hereafter indicated with +) expressed the most agreement amongst all groups with these statements. MS students who had not completed one of these tests (hereafter indicated with -) expressed the strongest disagreement amongst all groups. These two MS groups also significantly disagreed with one another on these statements (r=.533, p<.01; r=.517, p<.01). There was general agreement between the high school (HS) and university (UNI) groups as no statistically significant differences were found. The MS+ group was also much more likely than the MS- group (see Table 2) to agree with the statements ‘Students deserve the score they receive on English tests’ (r=.294, p<.05) and ‘English tests predict language ability in the future’ (r=.262, p<.05). MS+ students were much more likely to value their test performance, be encouraged by their test scores, and hold a stronger belief that these scores will predict future success in the L2. Conversely, the MS students who had not completed a high-stakes test appeared to negatively signify English tests and place far less importance on this kind of performance. These results

*p<.05, **p<.01

**Figure 3: Selected Questionnaire Items for Performance/Signification**
suggest that those MS students who had completed a TOEFL or TEPS had somewhat more positive assessment experiences and signified these experiences in more positive ways.

7.2 Other Evaluation/Self Evaluation

Figure 4 above displays the bar charts and correlation coefficients for TOEFL/TEPS completion and the questionnaire items ‘I think the English tests I have taken so far are fair’ and ‘My score on English tests reflects my English ability’. The MS+ group reported a stronger belief than the MS- group in the fairness of their English test scores (r=.453, p<.01) and that these scores were good indications of their ability (r=.501, p<.01). The relationships were far weaker for the HS groups and virtually disappeared for the UNI groups. It is interesting to note that the strength of belief in the fairness of test scores reversed somewhat for the university students (although the correlation was not statistically significant). A similar relationship was found (see Table 2) for the MS groups for the questionnaire items ‘English tests allow students to see how well they are doing’ (r=.348, p<.01) and ‘Students should trust English test scores’ (r=.286, p<.05). These results suggest that those middle school students who had completed a TOEFL or TEPS
possessed more trust and acceptance of the fairness and accuracy of English testing and increased reliance on test scores as a measurement of progress. The strength of the relationship decreased for the HS groups and disappeared for the UNI groups.

### 7.3 Social Value / Comfort

**Figure 5: Questionnaire Items for Social Value/Comfort**

Figure 5 above displays bar charts and correlation coefficients for the questionnaire items ‘It is important for Korean students to use English well’ and ‘I feel comfortable when using English’. The greatest disparity once again occurred within the MS group. While not as dramatic as elsewhere, The MS- group reported the weakest belief in the importance of using English well amongst all groups while the MS+ group reported the strongest, a result which significantly differed between the two groups (r=.299, p<.05). More strikingly, the MS- group reported the lowest comfort in using English amongst all groups, a result which significantly differed from the MS+ group (r=.512, p<.01). The MS+ group was also more likely than their cohorts (see Table 2) to agree with the statement ‘Studying English has had a positive effect on my learning experience’ (r=.387, p<.01). These results suggest that MS students who had completed a TOEFL or TEPS placed more social value on learning English than their cohorts, and perhaps
due to their positive attitudes towards performance, had a more positive English learning experience. The opposite appears to be the case for MS students who did not complete one of these tests. Interestingly, on whether it was important to study English to be successful (see Table 2), the UNI+ group reported a weaker belief in this statement than their cohorts, although the results were not statistically significant.

7.4 Aspiration/Inspiration

Figure 6 above displays bar charts and correlation coefficients for the questionnaire items ‘Universities and companies should value English test scores’ and ‘I learn English because I enjoy it’. Both the MS+ (r=.363, p>.01) and HS+ groups (r=.371, p>.01) were more likely to desire English test scores to be highly valued for educational and employment purposes. There was little difference found between the university groups. The MS+ group also reported far more enjoyment in learning English than their respective cohorts (r=.579, p<.01). This relationship was also found between the HS groups (r=.305, p<.01) but was not statistically significant for the UNI groups. The MS+ group was also more likely (see Table 2) to feel that students should be
required to study English (r=.411, p>.01) and that English tests were important reasons for studying English (r=.302, p<.05). It appears that, perhaps due to a confluence of positive attitudes and beliefs, that the MS+ group possessed the most positive harmony between their aspiration and inspiration for learning English. Interestingly, there was little difference in the cohort groups on how much they worried about their English test scores. After a discussion of the limitations of this study, it is to the polarized beliefs, attitudes, and motivations discovered overwhelmingly in the MS cohort groups that this paper will turn.

8. Limitations

There are some important limitations to this study that need to be mentioned. First, it must be noted that middle school students may be more likely to answer questionnaire items according to what they feel is the ‘right’ answer and not necessarily report what they genuinely feel. Given that students were being requested to complete a survey by a foreigner, some may have felt uneasy about expressing negative beliefs and attitudes about English. This tendency may have been stronger for those who had been exposed to English learning and testing to a greater extent. As always, it is difficult to investigate younger ELL beliefs and attitudes and future research should incorporate far more sophisticated questionnaire designs in addition to qualitative methods to help minimize this potential and reveal more robust information.

Additionally, this research only considered the completion of one of two high-stakes English tests in Korea and it did not consider the number of times these tests were completed. While these are two very common high-stakes English tests in Korea, there are a number of additional tests which play a role in the Korean educational market (see Choi, 2008, for a fairly exhaustive list). It is highly likely that older students have completed a greater number and variety of high-stakes language tests, and may have even become somewhat numb to their effects. This could help explain the reduced effects witnessed in the HS groups and the lack of any significant differences found for the UNI groups. A larger study incorporating a greater number of tests, including the number of times completed, would help to better understand these relationships.

Finally, this is also a rather small convenience sample and as such, the data cannot adequately represent the population from which it was drawn nor can it be extended to other
populations. These are limitations that can hopefully be improved upon in future studies. The proceeding observations are made with these important caveats in mind.

9. Discussion and Conclusion

A somewhat confusing picture emerges from the data collected for this study. While there were no significant relationships found for any of the questionnaire items for the university (UNI) groups, and only a few found for the high school (HS) groups, quite dramatic differences were found for the middle school (MS) groups, statistically significant differences being discovered on 15 of the 20 questionnaire items (10 of which were significant at the .01 level). On the original two research questions, the data collected strongly indicated that there was a significant difference in the assessment and learning experiences of young ELLs who had taken a high-stakes test and that this difference was far less dramatic for higher education groups.

Focusing on the assessment experiences of MS students, the data collected for this study indicated that those who had completed a TOEFL or TEPS reported far greater trust in their English test scores and had a stronger belief in the accuracy and fairness of English tests than their cohorts who had not completed one of these high-stakes tests. They also placed more significance on these external performances, were more likely to view them as measurements of their achievement and progress, and were less discouraged as a result of test scores. Alternatively, MS students who had not completed one of these tests were more likely to distrust their English test scores, de-signify their performances and feel discouraged as a result of English tests. There was a similar, but far weaker relationship found for HS students and this relationship largely disappeared (or reversed in some cases) for UNI students. The assessment attitudes and beliefs of MS students polarized markedly around high-stakes test completion.

With regards to the learning experiences of ELLs, the data collected indicated that MS students who had completed a TOEFL or TEPS tended to place more value on their English test scores and wanted their scores to be highly valued by universities and companies. They also reported less stress while using English, felt that English had a positive influence on their learning, and were more likely to learn English because they enjoyed it. By comparison, MS students who had not completed one of these tests tended to devalue their English tests (both personally and socially), experienced more discomfort while using English, and felt far less enjoyment. These relationships again broke down for HS students and largely disappeared for
those in university. These results suggest that MS students are more intensely influenced by their assessment and learning experiences and, as a result, become more polarized in their attitudes, beliefs and motivations towards it.

Turning to the third research question posed, these results have a number of important implications for a unified vision of high-stakes testing validity for younger ELLs and the appropriate age at which such testing should occur. The data collected support the contention that ELLs at the middle school level who are able to do the test (for a variety of reasons) may actually benefit from the experience. It may motivate their learning and promote feelings of accomplishment. This may engender more positive perceptions of proficiency and improve their overall assessment and learning experience. Not being able to do one of these tests may engender negative perceptions of assessment and a more negative language learning experience for MS students. Somewhat surprisingly, the social consequences of high-stakes testing for young ELLs in this context appear to be less about what it does to young learners who are able to take the test, but what it does to those who are not. Some may question whether even a unified concept of test validity can or should incorporate these kinds of social consequences, preferring instead to rely on other social agencies to address the issue. For the remainder of this paper, I will put forth a case for why test designers themselves are best suited for this role and why an “effect-driven” test validation process (Fulcher, 2009) must be encouraged in order for consequences such as these to be properly acknowledged and minimized.

This study has explored this issue from a socio-cognitive perspective, and has attempted to build a more meaningful link between the attitudes, beliefs and motivations of young ELLs, their evolving proficiency identity, and the social consequences of test use. However, as Block (2007) has pointed out, there is also a need to rebalance the individualistic perspective that has dominated identity work for the last decade and begin to examine issues, especially in FL contexts, from a social class perspective as well (p. 872). From this angle one may wonder, as Bernard Spolsky did as early as the 60’s, if this is primarily a class issue that limits the study of English “to the children of parents well enough established financially or politically” (Spolsky in Fox et.al, 2007). Indeed, preparation for high-stakes tests in Korea as elsewhere usually requires significant economic resources in order to attend private language schools. More prestigious private language institutes, arguably more effective at test preparation, can be out of reach for many students. Considering the costs of taking standardized tests, and the importance of test
scores, it is plausible that young students who are able to gain access to these educational resources feel more privileged because they are able to do something that others their age are not. The converse could be the case for those who are not financially able to prepare for and take these tests.

However, it is important to point out that these samples were selected from two schools located in the same relatively high socioeconomic area in downtown Seoul. For the purposes of this study, it was decided not to collect specific socioeconomic data from young ELLs, and therefore the relative degree to which these relationships are the result of socioeconomic class or the high-stakes testing environment remains unclear. In all probability, both factors are playing some role. However, it is also probable that the high-stakes testing environment created by test use magnifies any socioeconomic effects; a relationship which underscores and emphasizes, rather than minimizes, the social consequences of high-stakes test use for young ELLs. To be sure, these are contextual issues which need to be more fully explored; however, the data from this study strongly suggest that the attitudes, beliefs, and motivations of young ELLs are strongly associated with high-stakes testing experience, and that these unintended social consequences of test use are dramatically influencing the assessment and learning experience of young ELLs in this context. We are then left with a decision of who to make accountable, an opinion that will no doubt be influenced by one’s conception of test validity.

If the cognitive level required to understand and successfully complete a high-stakes test is beyond what a typical young ELL is able to achieve, it is at the very least worthwhile to investigate the consequences this may have on their short-term and long-term attitudes, beliefs, and motivations. If it can be shown that this has a positive impact on conceptions of proficiency and the learning trajectory of young ELLs who are somehow able to do it (as suggested by this study), while negatively impacting those who are not, it is an area where proponents of unified concept of validity should be gravely concerned. It is precisely these kinds of unfair and unjust consequences that modern language testing should be striving to avoid. At minimum, these results call for further investigation into the effects of high-stakes language testing on LPID, especially for young learners. It is necessary for test designers and administrators to better understand the dramatic effects of language testing practices on young learners, particularly when high-stakes are involved in test use. The concept of “effect-driven testing” advanced by Fulcher (2009) places social responsibility squarely with test designers and challenges them to
align “explicit statements of intended test-effect to test design decisions” and if these intended
effects are not being realized, “retrofit their validity argument” (p. 13). This approach carries
great potential to minimize negative test impact. The results of this study strongly suggest that
test designers make more explicit statements concerning the age-appropriateness of their test,
and the potential negative impact (on individual test takers and the larger educational context)
that could result if the test is not used as intended.

Considerable effort has been expended to improve the construct validity of high-stakes
language tests, improve their relevance and utility, and explore the value implications inherent in
their design and deployment. However, as Messick envisioned, the social consequences of test
use is situated at the apex of a complex and unified validity argument. The effects of high-stakes
language testing on young ELLs’ learning and assessment experience, and on their
socio-cognitive identity formation, may be viewed by some as pushing the limits of the test
validation process to the extreme. However, by incorporating these kinds of social consequences
into the test validation process, test designers can more fully acknowledge the wider social
context in which their tests live and breathe, actively engage developments in social theory, and
pursue the highest standards of test validity. It is a tall task to be sure, but one that needs to occur
at the source in order to limit, as much as possible, the damaging social consequences that can
result from the deployment of a high-stakes test.

References

Bachman, L. (2005). Building and supporting a case for test use. Language Assessment
Quarterly, 2(1), 1-34.

Press.


Cambridge: Cambridge University Press.


Input Enhancement, Noticing, and Incidental Vocabulary Acquisition

Katerina Petchko

National Graduate Institute for Policy Studies, Tokyo, Japan

kpetchko@grips.ac.jp

Bio Data

Dr. Katerina Petchko is Assistant Professor at the National Graduate Institute for Policy Studies (GRIPS) in Tokyo, Japan. She teaches graduate courses in thesis and dissertation writing and conducts research on literacy development.

Author Note

I thank Paul Nation, who made valuable suggestions on an earlier draft of this paper, the anonymous reviewers, who made constructive comments on this manuscript, and the learners in Japan and Russia who kindly volunteered to participate in the study.

Abstract

This study investigated the effect of input enhancement on vocabulary acquisition from reading at 98% known-word coverage. 47 intermediate EFL learners from 11 language backgrounds read a level-appropriate English story containing 12 nonwords under one of two conditions—with or without textual enhancement of the target words. The participants were tested on word noticing, word meaning recognition, and word meaning recall. Both groups showed large gains on all the tests: The enhanced reading group scored 58% correct on noticing, 43% correct on meaning recognition, and 24% correct on meaning recall, whereas the unenhanced reading group scored 65% correct on noticing, 39% correct on meaning recognition, and 25% correct on meaning recall. However, there were no significant differences between the groups on any of the tests, indicating that textual enhancement did not have any effect on either noticing or learning. Interviews with the participants confirmed that the learners in both groups had noticed many of the new words. Large variation was observed among the learners on vocabulary gains and among the words on pick-up frequency.

Keywords: incidental vocabulary learning, input enhancement, noticing
Introduction

Incidental vocabulary learning from reading is considered to be a major mode of vocabulary acquisition in a first language (Nagy, Herman, & Anderson, 1985; Sternberg, 1987). Similar claims have been made for second language (L2) vocabulary acquisition (Krashen, 1989; Nation, 2001). In the vocabulary acquisition literature, *incidental learning* is often defined as the accidental learning of vocabulary without an intention to learn (Hulstijn, Hollander, & Greidanus, 1996), or as a “by-product, not the target, of the main cognitive activity, reading” (Huckin & Coady, 1999, p. 182). In other words, it is a process that occurs naturally, during reading, while the learner’s attention is focused on overall text comprehension.

The L2 literature is replete with reports showing that reading for meaning leads to small but statistically significant vocabulary gains both in ESL (Cho & Krashen, 1994; Dupuy & Krashen, 1993; Horst, Cobb, & Meara, 1998; Knight, 1994; Pitts, White, & Krashen, 1989) and EFL (Day, Omura, & Hiramatsu, 1991; Webb, 2008) contexts. On the other hand, there is reason to believe that L2 reading for meaning does not automatically translate into vocabulary acquisition and that, at best, only *some* learners acquire *some* of the words *some* of the time. This is reflected primarily in the small pick-up rates associated with incidental learning of vocabulary. For example, in Cho and Krashen’s (1994) study, the two participants who did not use a dictionary acquired only seven and eight words per 70 pages read. Horst et al. (1998) reported a gain of five words for the 109 pages their learners had read. In Day et al.’s (1991) experiment with high school students, learners were able to recognize the meaning of only 0.9 words after reading for 30 minutes. Moreover, learners often fail to acquire new words from reading even after numerous encounters with these words (Horst et al., 1998; Waring & Takaki, 2003).

Could the outcome of incidental vocabulary learning be improved? For example, would enhancing vocabulary items in context—by bolding or underlining them—make them more noticeable to learners? More important, would this enhancement lead to greater vocabulary pick-up rates? This study is an attempt to answer these questions.
What Variables Affect Incidental Vocabulary Learning?

Previous research has uncovered several factors that may influence the success of incidental vocabulary acquisition from reading. One such factor is the proportion of unknown words in the text because in order to learn new words from context, learners must be able to comprehend what they are reading. Several studies have investigated the relationship between known-word coverage and text comprehension. Laufer and Sim (1985), for example, estimated that learners need to have at least 95% coverage of the running words (i.e., one unknown word in every 20) to gain reasonable comprehension, which in their study was operationalized as a score of at least 55% on a reading comprehension test. Nation (2001), however, argues that one unknown word in every 20 is “still a heavy load of unknown vocabulary” (p. 233). Hsueh-chao and Nation (2000) found that even with a reasonably easy text, learners needed 98% coverage (i.e., one unknown word in every 50) to gain “adequate unassisted comprehension” (p. 422) and have reasonable success at guessing correctly from context. In their study, few learners reading at 95% coverage could fully comprehend the reading.

Another crucial factor in incidental vocabulary learning is word guessability. Because incidental learning does not assume any explicit instruction, words can only be learned through lexical inferencing, or by guessing their meaning from context. Previous research has shown that the presence in the context of sufficient linguistic and semantic clues is one of the most important determinants of word guessability (e.g., de Bot, Paribakht, & Wesche, 1997; Hulstijn, 1992) and learning (Webb, 2008). For example, in a recent study of the effects of context on incidental vocabulary learning, Webb found that learners who had seen target words in more informative contexts (i.e., in which few meanings were possible apart from the correct meaning, as in *He was not ill, and of course the beds in the ancon are for ill people*, where *ancon* means *hospital*) scored significantly higher on vocabulary recognition tests than did learners who had seen target words in less informative contexts. This finding implies that studies using natural readings, particularly in EFL contexts, may have underestimated potential vocabulary gains from reading because in natural prose, context often does not provide enough clues or provides misleading information about word meanings (Kelly, 1990; Schatz & Baldwin, 1986). Ehri (2002), for example, argues that on average, only 25-30% of words can be guessed correctly in natural readings; Nation (2001) gives a
Lexical inferencing may also be disproportionately difficult for EFL (as opposed to ESL) learners because of their small vocabularies (Hunt & Beglar, 2005).

Perhaps the most important variable in incidental vocabulary learning is reading comprehension, or the ability to extract meaning—both literal and interpretive—from printed text, which requires that the reader be able to decode and retrieve the meaning of each individual word in the text, put the meanings of these words together, and then interpret the whole phrase within the context of the sentence, paragraph, and the overall understanding of the text (Adams, 1990). Reading comprehension has been shown to be a significant determinant of the amount of incidental vocabulary learning from reading, at least in a first language (Swanborn & de Glopper, 1999, 2002). The relationship between reading comprehension and incidental vocabulary learning may be even more important in a second language because of a greater variability among learners, especially at lower levels of proficiency. Therefore, ascertaining reading comprehension in a study of reading in a second language is critical because learners who have failed to achieve successful reading comprehension—both at the level of literal comprehension and at the level of interpretation—can hardly be expected to demonstrate significant vocabulary growth. Unfortunately, the dimension of reading comprehension has been neglected in many second-language studies and it has often been assumed, rather than confirmed empirically, that learners have read and understood the text.

**The Role of Noticing in Learning**

Despite the importance of these conditions—adequate known-word coverage, word guessability, and text comprehension—to incidental vocabulary learning, they are clearly not sufficient. Previous research has shown that even when the known-word coverage is high and the target words are guessable from context, learners still often fail to learn them—in some cases, despite numerous encounters. For example, in Webb’s (2008) experiment, although learners who met target words in more informative contexts outperformed those who saw the same words in less informative contexts, on the cognitively more demanding meaning recall test, the experimental group demonstrated an average gain of just one word. In Waring and Takaki’s (2003) study, which used a story with 96% known-word coverage, at the immediate posttest, the scores for words occurring four or five times were near zero on all the tests.
One reason discussed in the literature (e.g., Hulstijn et al., 1996) for the modest vocabulary gains from reading for meaning is the apparent failure on the part of the learner to notice new words. This is an important observation because although the exact mechanism of incidental vocabulary learning is still not fully understood, many researchers (e.g., de Bot, Paribakht, & Wesche, 1997; Gass, 1988; Hatch & Brown, 1995) agree that input processing in incidental learning must involve an initial stage of noticing a new word.

The crucial role of noticing in language learning has been noted by many researchers. Schmidt (1990), for example, views noticing as “the necessary and sufficient condition” (p. 129) for input to become intake. Referring to vocabulary acquisition, he defines noticing as “conscious registration of the form…of a word” (Schmidt, 1995, p. 29) and argues that learning without noticing is impossible. Ellis (1995) maintains that noticing is important for the acquisition of both form and meaning of a new word and that it strongly facilitates the acquisition of new vocabulary.

Support for the facilitative effect of noticing on language learning comes primarily from the grammar acquisition literature. Leow (2000) found that noticing significantly facilitated L2 learners’ intake and written production of morphological forms: Learners who had noticed target forms were able to take in and produce significantly more of these forms than learners who had not noticed the target forms. Rosa and O’Neil (1999) exposed 67 learners to a Spanish conditional form and found that noticing had a significant effect on the learners’ ability to recognize the target structure. These results lend strong empirical support to a positive association between noticing and L2 learning. They may also at least partly explain the low acquisition rates often reported in incidental vocabulary research: If learning is contingent on noticing, then learners who fail to notice the presence of unfamiliar words in the input can hardly be expected to acquire them from the input.

**Input Enhancement and Noticing**

If noticing is an important—in fact, a first—condition for learning, then techniques that promote noticing may also improve learning. One such technique is textual enhancement, or the visual enhancement of items by bolding or underlining them in order to increase their perceptual salience. Several studies have found such techniques to be more effective for learning than purely communicative instruction or even
memorization. For example, Leeman, Arteagoitia, Fridman, and Doughty (1995) investigated the effect of a variety of focus-on-form techniques on language learning and found that enhancement techniques designed to draw learners’ attention to the input were superior to purely communicative instruction for both accuracy and frequency of target-form production. Robinson (1997) compared grammar acquisition under different conditions including textual enhancement to promote focus on form and found that textual enhancement was superior to memorization and led to “generalizable, non-item-specific learning” (p. 239). Jourdenais et al. (1995) reported that input enhancement significantly improved the noticing of target forms and that noticing facilitated acquisition.

Recently, some researchers (e.g., Waring & Takaki, 2003) discussed input enhancement as a possible way to help learners notice and learn new words in reading for meaning. However, whether highlighting new words in a text would improve their noticing and, ultimately, their acquisition is still an open question. On the one hand, studies from grammar acquisition suggest that such techniques may be effective for both noticing and acquisition. On the other hand, it can be argued that the nature of attention as a limited capacity resource (Kahneman, 1973) may preclude learners from simultaneously carrying out two tasks—i.e., attending to form and processing content for meaning—if these tasks draw on the same pool of available resources. In fact, the idea that one’s ability to perform two tasks concurrently depends on the kind and amount of resources required by each task (Wickens, 1980) was used by Bill Van Patten (1996) to argue that when processing foreign-language input, learners are often unable to attend simultaneously to form and meaning and, furthermore, that they are likely to allocate more attentional resources to meaning than to form. This argument, however, rests on an assumption that the processing of input for form and for meaning represents two different processes drawing on the same pool of cognitive resources. Some researchers (e.g., DeKeyser, et al. 2002), however, have argued that attending to form and meaning in communicative interactions should be conceived of as “a single task drawing on the verbal encoding resource pool” (p. 809). These researchers cite evidence from experimental studies of incidental learning showing that learners can, and do, acquire new forms during processing for meaning. In reading a text that is readily comprehensible and that contains only a small amount of unknown vocabulary, attending simultaneously to form and content may be particularly unproblematic. Under these conditions, textual enhancement may not provide any additional benefit because the task of reading for meaning will likely direct learners’ attention to
unfamiliar words (Schmidt, 1990), particularly if they are deemed important for comprehension.

It has also been shown (e.g., Alanen, 1995) that textual enhancement does not always have an effect on performance. Arguably, this is because noticing of the input is not sufficient for learning: For new information to be learned, it must be processed sufficiently deeply. In fact, some learning models in cognitive psychology (e.g., Craik & Lockhart, 1972) and L2 acquisition (Laufer & Hulstijn, 2001) view the quality of information processing as a key factor that determines whether or not new information will be learned. From this perspective, the effectiveness of textual enhancement will depend on whether reading a text with highlighted words leads to more elaborate processing than does reading an unmarked text.

The Notion of Vocabulary Acquisition

A key issue in vocabulary research is what constitutes acquisition. Many earlier studies of incidental vocabulary learning have relied on a single, typically multiple-choice, test to establish vocabulary acquisition. However, as Waring and Takaki (2003) rightly point out, such tests may be far from an ideal measure of vocabulary knowledge because they measure only prompted recognition, which does not necessarily reflect the type of word knowledge that is needed for natural reading. Their own data suggest that less than half the words identified correctly on a multiple-choice test are available for unprompted recall.

An alternative to the multiple-choice format is a test of unprompted meaning recall, such as one asking learners to supply a translation or a definition of a target word. Although nowhere near as common as the multiple-choice test, tests of meaning recall have also been used in incidental vocabulary research (e.g., Hulstijn, 1992; Knight, 1994). However, such tests have been criticized for lacking sensitivity and for a potential underestimation of the actual vocabulary gains (Nation, 2001) because vocabulary gains from reading are often partial (Hunt & Beglar, 2005) and are made in stages (Curtis, 1987; Hendriksen, 1999; Paribakht & Wesche, 1997), ranging from word form recognition to prompted meaning recognition to unprompted meaning recall to production.

A better approach is to use several tests to allow learners to demonstrate partial acquisition of new vocabulary. The use of several tests also provides a better means of capturing potential vocabulary gains because different tests may be sensitive to
different aspects of word knowledge. On the other hand, the use of multiple measures must be tempered with the need to include dependent measures that are uncorrelated with one another, i.e., that measure separate, unrelated aspects of behavior (Tabachnick & Fidell, 2001). In this study, two types of vocabulary test were used, a multiple-choice test to assess prompted meaning recognition, and a word meaning test to assess meaning recall. It was hypothesized that these two tests measured somewhat separate facets of learning and were therefore unlikely to correlate highly. At the same time, the use of these two different tests improved the chance of capturing partial vocabulary gains.

Research Question

The purpose of this study was to determine the effect of input enhancement in reading for meaning on the noticing and subsequent acquisition of vocabulary. The following research question was formulated: What is the effect of the textual enhancement of unfamiliar words in a text on their (a) noticing (i.e., word form recognition), (b) prompted meaning recognition, and (c) unprompted meaning recall?

Method

Participants

Forty-seven EFL learners from three educational institutions—a university in Tokyo (n = 21), a graduate institute in Tokyo (n = 18), and a university in St. Petersburg, Russia, (n = 8)—participated in the study. The learners came from 11 language backgrounds: Japanese (n = 20), Russian (n = 10), Indonesian (n = 4), Chinese (n = 4), Korean (n = 1), Sinhala (n = 2), Uzbek (n = 2), Kyrgyz (n =1), Azerbaijani (n = 1), Thai (n = 1), and Vietnamese (n = 1). Most of the learners were participants in a larger study to investigate cross-linguistic reading development and all volunteered to participate in the experiment described in this study. All had studied English for at least nine years and had intermediate to high-intermediate proficiency. For the non-Japanese students, this was determined on the basis of their TOEFL scores (i.e., all had scores above 500 on the pencil-and-paper TOEFL). Unfortunately, such data were not available for the Japanese students, and their level was determined on the basis of their language learning history and classroom observation. In addition, prior to the experiment, the participants were tested on receptive vocabulary knowledge using the Vocabulary Levels Test (Nation, 1983). All scored 30 out of 30 at both the
1st and 2nd 1000 word level and were judged to have adequate vocabulary to understand a text written at the 2000 word level.

Study Design
Participants were randomly assigned to one of two groups, Enhanced Reading (ER, \( n = 24 \)) and Unenhanced Reading (UR, \( n = 23 \)). Care was taken to ensure that the groups contained similar numbers of speakers of different languages. Learners in the ER group read a story with 12 nonsense words textually enhanced by bolding and underlining; learners in the UR group read the same story with unfamiliar words not marked in any way. The participants were then tested on reading comprehension and on three types of test: word form recognition, unprompted meaning recall, and prompted meaning recognition.

Target Words
In a study of incidental vocabulary learning, it is important to base the estimates of vocabulary acquisition from context on the actual words not known by the particular learners rather than on some hypothetical estimates of what the learners may or may not know (Nation, 2001). To ensure that the target words would not be familiar to learners before the reading, a common approach used in incidental vocabulary research is to substitute these words with nonsense words (also called nonwords). This approach guarantees that the demonstrated vocabulary gains can come only from the reading.

Twenty-two words were selected from the text as potential target words. The main criterion for selection was word guessability from the immediate and/or overall context of the story. These words were replaced with English-looking nonwords and pilot-tested on a group of advanced EFL learners (\( n = 12 \)), who were asked to read the story and to judge the guessability of the nonwords and their plausibility and pronounceability in terms of English spelling conventions. Based on the results, 12 target words were selected. All represented concepts that were thought to be familiar to the participants. Two were nouns, one occurring twice and one occurring three times; three were adjectives, one occurring twice and two occurring once; one was an adverb, occurring once; and six were verbs, all occurring once. There were more verbs than other parts of speech because the pilot test had shown that they were easier to guess and were more important to the overall understanding of the story. Two of
the verbs were used in the Simple Past tense and were inflected as regular English verbs; one was used with the inflection –s (present tense, third person singular). These verbs appeared in the tests in the same form as in the text. Participants were not penalized for supplying only the base form of these verbs. The target words are shown in Appendix A.

Reading Materials and Reading Comprehension

The text used in the study was a narrative by William Caine about an artist who agrees to discredit the work of a younger artist in exchange for money but who secretly sends the money to the young artist to help him become independent. The text was shortened to ensure that it could be read in approximately 10 minutes and its vocabulary was screened to ensure that all of the words excluding proper names were within the most frequent 2000 words of English. All words in the text outside the 2000 word level were replaced with words or phrases within the 2000 level except one. The exception was the word published. This word was not replaced because it was not crucial to the overall understanding of the story and because many of the participants were likely to know it. The final version was 775 words long: 90% at the 1st 1000 word level, 3.7% at the 2nd 1000 word level, and 4.1% proper names, which were counted as familiar words. Thus, there was an expected known-word coverage of 97.8% of the running words. The 12 target words occurred in the text 16 times and accounted for 2% of the running words. The text is shown in Appendix B.

A reading comprehension test was used to measure comprehension of both main ideas and additional information in the story. The test consisted of three open-ended questions, five yes/no questions, and seven multiple-choice questions. The multiple-choice questions had two distractors, one correct answer, and an I don’t know option. The answer choices were randomly placed in the first, second, or third position; the I don’t know option was always in the final position. Six of the questions covered main ideas, and the remaining nine covered additional information. Fourteen questions dealt with explicit information and one required an overall understanding of the story. None of the questions contained any of the target words. Two points were awarded for each correct answer to a question covering a main idea and one point was awarded for each correct answer to a question covering additional information. The maximum number of points was 21. An arbitrary level of all correct answers on the main-idea questions and at least seven correct answers on the questions dealing with
additional information was set as a criterion for adequate text comprehension. This rather high level of comprehension was chosen to ensure that the participants had read the story and understood most of it.

**Vocabulary Tests**

On the multiple-choice test, each target word had one correct answer, four distractors, and an *I don’t know* option. The distractors were syntactically the same as the target words but semantically different in order to allow the participants to demonstrate even modest vocabulary learning; all were plausible choices based on the context of the story. Participants were asked to circle the words with the meaning closest to that of the target words. They were told not to guess and to choose *I don’t know* if they did not know the answer. The test was scored as the number of correctly identified target words.

On the word meaning test, participants were given a list of the 12 target words and were asked to supply a definition or provide a synonym for each word in English or in their native language. They were told that they could give several alternatives and that their answers could be as long as they liked. Correct answers were defined as all contextually acceptable answers even if they were not identical to the original words. Contextual acceptability was determined prior to scoring using the following procedure.

First, the story containing the nonsense words was given to a panel of eight judges (three native English speakers, two advanced-level Russian English speakers, and three intermediate-level Japanese English speakers), who were asked to come up, individually, with as many contextually possible alternatives for the nonsense words as they could. The alternatives produced by the panelists were then typed up on a list and the list was again presented to the panelists, who were asked to evaluate how closely each word fit the context. The words that were selected unanimously by the panelists as acceptable alternatives were then used to create a scoring sheet, which was used to score participants’ responses on this test.

In the scoring procedure, all non-English answers were first translated into English. Only three participants chose to write their answers in their native language: Two wrote their answers in Japanese and one did so in Russian. The two Japanese learners defined one word each—the word *moop* as *sainoo*, which is the main dictionary definition of the word *talent*. The Russian learner defined three words: the
word *moop* as *talant* (*talent*), the word *quented* as *skazala “da”* (*said “yes”*), and the word *heefy* as *ochen’ horoshii* (*very good*). The translations together with the English-language answers were checked against the scoring sheet and those that were deemed correct were awarded one point each. The participants were not penalized for supplying an incorrect part of speech (e.g., *talented* for *talent*). This rather loose definition of correct answers was used to give the participants credit for demonstrating even small gains in word knowledge. Previous research has shown that learners need to meet an unknown word many times before it can be learned (Horst et al., 1998; Waring & Takaki, 2003). In this study, most words occurred only once; it was therefore unrealistic to expect perfect responses on a meaning recall test.

*Operationalization of Noticing*

In studies conducted under the attentional framework, measures of noticing generally fall into two categories: online measures (i.e., think-aloud protocols) and offline measures (i.e., post-exposure tasks). Although online measures are considered more direct and higher in internal validity, offline measures are also appropriate if the goal is to “make inferences as to whether learners either paid attention to or became aware of targeted forms in the input” (Leow, 2000, p. 570), rather than to differentiate between degrees of awareness. Because the goal of this study was to establish noticing, post-exposure tasks were considered appropriate. Thus, noticing was equated with post-exposure word form recognition and operationalized as a score on a word-form recognition test administered immediately after exposure to input. The test consisted of the 12 target words and 12 distractors, which were also English-looking nonwords. The distractors did not look similar to the target words. The order of the words was randomized. Participants were instructed to circle the words that they had seen in the text. This test was scored using a procedure from Waring and Takaki (2003): one point was awarded for each correct answer, one point was awarded for each incorrect answer, and then adjusted means were calculated by subtracting the number of incorrectly recognized words from the number of correctly recognized words.
Procedure

Data were collected on multiple occasions during a 2-year period as part of a larger project to investigate reading-related abilities. All participants were tested individually or in small groups in a quiet room on school premises. The learners were told that the purpose of the study was to investigate reading-related abilities and that they would read a short story and answer some questions. They were not told about any vocabulary tests.

Two versions of the text were prepared, one with the target words bolded and underlined and one with the target words left unmarked. The groups were given the same instructions: To read the story for meaning and enjoyment. The participants were not told that there would be unfamiliar words in the story. Dictionaries were not allowed. As soon as the participants finished reading, the text was taken away from them and they were given the reading comprehension and vocabulary tests in this order: The reading comprehension test was given first and it was followed by the word form recognition, word meaning, and multiple-choice test. The word form recognition test required the least amount of word knowledge and the word meaning test required the greatest amount. The multiple-choice test was given last to ensure that the participants could not have remembered some of the word meanings from the word meaning test. As soon as a participant finished one test, the test was collected and another one was given. It took the participants approximately 10-20 minutes to finish the story and the tests.

After the completion of the tests, 20 participants (ten from each group) were interviewed. They were asked to indicate (a) the difficulty level and the extent of comprehension of the story, (b) whether they noticed any unfamiliar words, (c) whether they tried to infer the meaning of these words and how easy it was, and (d) how difficult it was to recall the target words on the tests. The interviews took about 5-10 minutes.

Results

Statistical Analyses

Analyses were carried out with SPSS version 10.0. Prior to analyses, reading comprehension, word form recognition, meaning recall, and meaning recognition were examined for accuracy of data entry, missing values, and the fit between the distributions of these variables and the assumptions of multivariate analysis of
The descriptive statistics for the three dependent variables are shown in Table 1. To determine whether there were any overall differences between the groups, a between-subjects MANOVA was performed on the three dependent variables: word form recognition, meaning recognition, and meaning recall. The independent variable was text condition (a text with target words textually enhanced vs. a text with target words unmarked). Hotelling’s $\mathbf{T^2}$ was used to compare group means, as recommended by Tabachnick and Fidell (2001, p. 20). This statistic is calculated from Hotelling’s Trace coefficient provided by SPSS using the following formula: Hotelling’s $\mathbf{T^2} = \text{Hotelling’s Trace} \times (N – m)$, where $N$ is the sample size across the groups and $m$ is the number of groups. The $\mathbf{T^2}$ has the same $F$ value, degrees of freedom, and significance level as the Trace statistic. No significant differences were found between the groups on any of the dependent measures, Hotelling’s $\mathbf{T^2} = 3.5, F(3,40) = 1.12, p = .352$. Thus, the dependent variables were not significantly affected by textual enhancement.
**Interview Data**

In the interviews, participants from both groups responded in a similar way. All indicated that the text was “easy” or “very easy,” that they were able to understand all or most of it, and that they had noticed unfamiliar words. Several participants said that the text contained “many” unfamiliar words, but most said that there were “a few.” Many said they thought the nonwords were real English words, just unfamiliar. Participants in both groups also said that they had tried to guess the meaning of the unknown words, and the majority noted that it was “not so difficult.” Finally, many learners pointed out that although they could recognize many of the target words on the tests, by then, they had simply forgotten what they meant. This is how one participant (from the UR group) described her experiences with the reading:

> Сам текст простой для понимания. Слова, которые не знаешь, сразу цепляются глазу, но смысл их понятен из контекста. Но вот когда начинаешь потом пытаться из памяти выловить что они значат... Тут-то вся загвоздка и таится. [The text itself is easy to understand; the words you don’t know catch the eye immediately, but their meaning is clear from the context. But when later you try to remember what they mean... That’s where the snag is. (Translated from Russian by the author.)]

Another participant (from the ER group) described her experiences in a similar way: “When I saw some of the words on the [vocabulary] tests, I could clearly remember seeing them in the text; I just couldn’t remember what they meant.” These two examples represent a typical response given by those interviewed. Overall, the similarities in the learners’ responses provide converging evidence supporting the lack of differences between the groups. These responses also suggest that the main hurdle facing the participants was not noticing new words, but rather, storing them in long-term memory.

**Discussion**

*Effect of Textual Enhancement*

The research question asked if textual enhancement of unfamiliar words in reading for meaning improved their noticing and subsequent meaning recognition and recall. The results show that the treatment had no effect on either noticing or acquisition:
Participants in both groups noticed and were able to recognize and recall approximately the same number of words. The lack of any significant differences between the groups on the word form recognition, meaning recognition, and meaning recall tests suggests that both groups must have attended to the target words with a similar amount of mental effort. These findings were confirmed in the interviews: The majority of the participants in both groups indicated that they had noticed the target words and that they had tried to guess their meaning.

The absence of statistically significant differences on the vocabulary measures between the learners is consistent with the noticing hypothesis (Schmidt, 1990), which postulates a crucial role for noticing in learning. If noticing precedes vocabulary learning, then the groups should be expected to differ on the vocabulary measures only if they differed on the measure of noticing. In other words, the treatment should improve acquisition only if it improved noticing. However, because the treatment in this study did not result in better noticing, it did not lead to increased vocabulary gains. Why did textual enhancement have no effect on noticing and learning? Although the experimental design used in this study precludes making causal inferences, several possibilities can be entertained. First, it has been argued that because of a limited cognitive capacity for processing information, L2 learners tend to process input for meaning before they process it for form (Van Patten, 1996). It is therefore possible that under limited processing conditions, some learners in the ER group did not even notice the highlighted words. This scenario, however, appears to be unlikely because participants in this group scored close to 60% on the word form recognition test. More important, an examination of their responses on this test revealed no erroneously selected items, suggesting that the participants in this group did process the target words for form as well as for meaning. The interview data also confirmed that learners in the ER group had noticed the highlighted words and attended to them.

The second explanation assumes the validity of the levels-of-processing theory (Craik & Lockhart, 1972), which stipulates that recall is constrained by the quality of information processing. According to this theory, textual enhancement would be expected to improve learning only if it led to deeper levels of information processing. This, in turn, appears to be related to task demands (e.g., Daneman & Carpenter, 1980): More cognitively challenging tasks require more elaborate processing, which leads to more stable learning. Some theories of L2 vocabulary acquisition also point to the crucial role of tasks in L2 input processing. For example, Laufer and Hulstijn (2001) introduced the notion of task-induced involvement and argued that the quality
of mental processing is determined by the demands of the particular task, possibly because task demands focus learners’ attention “on relevant features of the input” (Schmidt, 1990, p. 149). In this experiment, the learners in both groups received the same instructions—to read a story for meaning and enjoyment, and they were not told that it would contain any unfamiliar words. It is possible that these instructions imposed similar cognitive demands on the learners and led to the use of similar processes during reading and learning. Thus, the lack of any significant differences between the groups may indicate that textual enhancement does not automatically improve the quality of input processing because the task of reading a text with highlighted words is cognitively no more demanding than the task of reading an unmarked text. Although this hypothesis was not tested directly, the null result obtained in the study suggests the use of similar cognitive processes by both groups.

It is also possible that highlighting may not provide any additional benefit when conditions crucial to incidental learning (i.e., adequate known-word coverage and word guessability) are met. Logically, this is because the intended goal in reading (i.e., meaning comprehension) would serve to focus the reader’s attention on unfamiliar words. If the number of these words is small and the surrounding context is understandable, these words will probably “stand out,” so the reader is likely to notice them even if they are not marked in any way. Indirect support for this claim comes, again, from studies of grammar acquisition. For example, in Robinson’s (1995) study of grammar learning under four conditions—implicit learning (i.e., students memorized sentences), incidental learning (i.e., students read sentences to answer comprehension questions), rule-search (i.e., students searched for rules exemplified by presented sentences), and instruction (i.e., students viewed rule explanations and applied the rules to new sentences)—there was no significant difference between the conditions on the extent of noticing: An overwhelming majority of learners in all conditions claimed to have noticed rules in the presented sentences even though the presented input was not enhanced in any way. Arguably, in reading, where the goal is to extract meaning, readers would be expected to pay even more attention to unfamiliar words than Robinson’s students paid to grammatical elements.
Response Patterns

An examination of learning data for each participant revealed several patterns. First, the participants in this study noticed many more words than they learned. Clearly, noticing alone is not sufficient for vocabulary acquisition. More research is needed to clarify the relationship between the noticing of unfamiliar words in context and their acquisition from context. Second, words that occurred 2-3 times in the text were recalled much better on both vocabulary tests than were words that occurred only once. For example, moop, which was used three times, was correctly recognized more than 86% of the time and was correctly defined almost 70% of the time; lantic, which occurred twice, was recognized more than 70% of the time and correctly defined 60% of the time. For words that were used only once, the percentages were substantially smaller: For example, for the word tranch, which was recognized most frequently in this word group, the pick-up rates were 45% for meaning recognition and 22% for meaning recall; for the word blunded, the least frequently recognized word, the pick-up rates were 20% for recognition and 7% for recall. The relationship between the number of times a word is seen in context and its acquisition from context is well supported in the incidental vocabulary literature (e.g., Waring & Takaki, 2003).

Third, although the learners differed widely in the number of words they were able to identify on the vocabulary tests (from 0 to 10 on the meaning recognition test; from 0 to 11 on the meaning recall test), a predictable pattern was observed, with performance on the meaning recognition test being much better than on the meaning recall test. In other words, the majority of the participants could define far fewer words than they could recognize. However, there were several exceptions: Four participants defined more words than they were able to recognize, and three participants defined and recognized the same number of words. There was also a small group of learners who showed no overlap between the words they could define and those they recognized. For example, one participant correctly defined moop but was then unable to recognize the meaning of this word on the multiple-choice test; another correctly defined lantic and tranch but recognized the meaning of different words. These results suggest a considerable degree of variation among learners not only in the particular words they pick up but also, possibly, in how they process these words and how they store them in long-term memory.
Limitations and Directions for Future Research

The present study is limited by the small number of participants tested and the relatively small number of words used. It could be replicated with larger samples, longer readings, and a greater number of target words to clarify the relationship between noticing and vocabulary learning under various conditions. Think-aloud protocols could be used to differentiate between degrees of awareness, as has often been done in studies of grammar acquisition.

More important, because learners in this experiment read an easy text with 98% known-word coverage in which unfamiliar words were easily guessable, these results cannot and should not be extrapolated to reading in more natural settings, in which learners have little control over unknown word density and word guessability. Whether textual enhancement might work in such settings is a question for future research. On the one hand, it is possible that under more realistic reading conditions, textual enhancement would have a positive effect on noticing by directing learners’ attention to unfamiliar words. On the other hand, it is hard to imagine substantial vocabulary gains from reading a text that is beyond learners’ current level of reading comprehension ability or in which words are not easily guessable from context. If the goal in L2 reading is learning as much as enjoyment, perhaps learners should be encouraged to read level-appropriate material that provides sufficient contextual support for potentially difficult vocabulary.

Conclusion and Pedagogical Implications

The present study investigated the role of textual enhancement of vocabulary on its noticing and subsequent up-take in reading for meaning at 98% known-word coverage. No significant differences were found between the learners who had read a text with the target items textually enhanced and those who had read an unmarked text on any of the measures, indicating that textual enhancement did not have any effect on either noticing or learning. Learners in both groups noticed more than half of the target words and picked up roughly one third of them.

The findings have several implications for vocabulary research and teaching. First, highlighting words in a text does not appear to be an effective way to promote their acquisition, at least for intermediate learners reading a level-appropriate text. In fact, there appears to be no need to encourage learners to pay special attention to vocabulary in level-appropriate readings, because learners reading such texts are able to notice
many new words and learn some of them even when these words are unmarked. Furthermore, both quantitative and qualitative data suggest that noticing new words in reading for meaning may not be the main problem facing the learner, at least when the text is comprehensible and the proportion of unknown vocabulary is small. The lack of significant differences between the groups on any of the measures and the similarities in the learners’ comments about the relative ease of noticing the target words in the text and the relative difficulty of remembering them on the tests indicate that the problem in incidental vocabulary learning may lie in how unfamiliar words are processed, stored, and retrieved from long-term memory after they have been noticed as well as in how—and how fast—they are forgotten.

Second, the lack of any effect of textual enhancement on vocabulary noticing and learning in the presence of a large number of positive findings from grammar acquisition studies suggests that grammar and lexis may be processed in qualitatively different ways. More research is needed to clarify possible differences.

Finally, there appears to be a considerable variation among learners in the particular words they pick up as well as in the way they process these words and store them in long-term memory. This variation may be attributable to learners’ individual differences in vocabulary processing and storage and it suggests an important role that these differences may play in determining the learning styles and patterns of strengths and weaknesses that individual learners bring to the task of reading. Teachers and course developers need to take into account the existence of these differences in the language classroom when teaching reading classes and preparing reading materials.
References


Appendix A

Original words, target words, and the number of occurrences in the text

<table>
<thead>
<tr>
<th>Original words</th>
<th>Target words (number of occurrences)</th>
</tr>
</thead>
<tbody>
<tr>
<td>discourage</td>
<td>tranch (1)</td>
</tr>
<tr>
<td>talent</td>
<td>moop (3)</td>
</tr>
<tr>
<td>admires</td>
<td>wodes (1)</td>
</tr>
<tr>
<td>criticize</td>
<td>incholate (1)</td>
</tr>
<tr>
<td>gasped</td>
<td>blunded (1)</td>
</tr>
<tr>
<td>agreed</td>
<td>quented (1)</td>
</tr>
<tr>
<td>worthless</td>
<td>lantic (2)</td>
</tr>
<tr>
<td>give up</td>
<td>speat (1)</td>
</tr>
<tr>
<td>unhappy</td>
<td>sminted (1)</td>
</tr>
<tr>
<td>carefully</td>
<td>tauciously (1)</td>
</tr>
<tr>
<td>nice</td>
<td>heefy (1)</td>
</tr>
<tr>
<td>benefactor</td>
<td>archentor (2)</td>
</tr>
</tbody>
</table>
Appendix B

MENDOZA SELLS HIMSELF

It was ten in the morning when Mendoza’s servant came into the living room with a card. The painter took the card and read, ‘Mrs. Burtenshaw.’ The name meant nothing to him.

The servant said, ‘The lady looks rich.’

‘All right,’ said Mendoza. ‘Ask her in.’

Mendoza disliked the woman at once, though she seemed eager to please him.

‘I hope you will forgive my early visit,’ she began. ‘I want to speak to you about my son Charlie. I want your advice, and I want your help. He is my only child. He grew up a good boy until he decided to become a painter.’ The woman stopped for a moment.

‘Do you want me to give him lessons? If so I regret to say that I won’t be able to…’

‘Oh, no, my dear Mr. Mendoza,’ she said, “quite the opposite, I want you to let him come and show you his drawings, and then I want you to tranch him, to tell him that he has no moop. You are the only man who can do me this favor. Charlie wodes you, he has collected hundreds of your drawings. Just tell him that he is no good. He should go into business.’

‘But what if your son refuses to change his mind? What will you do then?’ asked Mendoza.

The woman’s voice became firm. “In that case he will never get any money from me as long as I live.’

Mendoza realized that she would do exactly as she said. He suddenly wanted to tell her to get out, but he checked himself and went on. ‘I’m sorry, Mrs. Burtenshaw,’ he said, ‘But I can’t promise anything until I see his work. He may be really gifted’.

She took out a sheet of paper which she handed to Mendoza.

Mendoza looked at the drawing in silence. It was very good.

He said, ‘I must say the drawing shows a lot of moop. You must know that moop cannot be taught. A painter either has it or he doesn’t. Don’t you understand that I find that drawing of your son most promising? Why should I tell him that he is no good?’

She looked at him. ‘You don’t think that I am asking you to incholate my son’s work for nothing? I am prepared to pay you for that.’

‘May I ask how much?’ Mendoza asked sweetly.

“Well,’ she said, ‘it will be one hundred dollars.’

Mendoza shook his head. ‘I will not do this for less than a thousand dollars.’

‘A thousand dollars!’ she blunded. ‘You must be joking.’

‘Not at all,’ Mendoza was firm. ‘A thousand dollars and not a penny less.’

To his great surprise she quented.

The money came the next day. And two days later Mrs. Burtenshaw with her son came to see Mendoza. Charles, very excited, brought some more of his drawings which he showed to Mendoza. No word said Mendoza while he looked through them. No word said Charles.

Then Mendoza put down the drawings and said: ‘Mr. Burtenshaw, I am sorry to say this, but your drawings are lantic. I think you should speat painting and go into business as your mother wants. Goodbye to you both, goodbye.’

He opened the door and Charles ran out of the room.
Four years had passed. The same servant showed Mr. Charles Burtenshaw into Mendoza’s living room. ‘You probably don’t remember me,’ said Charles, ‘but I was here four years ago with my mother and you said my drawings were lantic. Could you look at my work now?’

‘You haven’t stopped painting then?’

‘No, I haven’t. When I left your house, I was so sminted. I wanted to cry. On the street, a young man came up to me and handed me an envelope. When I opened it, to my surprise, I found one thousand dollars. That money made me independent of my mother. I left her house and went to Paris where I lived hard and worked hard. Now I’m back. This is my first published drawing.’

Mendoza examined the drawing tauciously. It was really good.

‘I think,’ he said, ‘it’s a heefy piece of work. Congratulations! The thing is really good.’

Charles smiled with pleasure. ‘Thank you very much,’ he said, ‘If I could only thank my archentor… But I don’t know who he is, I don’t know his name.’

‘In your place I wouldn’t worry! Just go ahead and make a big success of yourself. I think your unknown archentor will be quite happy.’
Appendix C

Reading Comprehension Test

1. Mendoza was  
   a. a musician   b. a writer   c. a painter   d. I don’t know

2. Before meeting the woman, did Mendoza know her?  
   a. Yes   b. No   c. I don’t know

3. When Mendoza saw the woman, did he like her?  
   a. Yes   b. No   c. I don’t know

4. What did the woman want Mendoza to do?  
   a. tell her son that his drawings were bad  
   b. give her some money  
   c. give her son drawing lessons  
   d. I don’t know

5. What did the woman want her son, Charlie, to become?  
   a. a businessman   b. a painter   c. a politician   d. I don’t know

6. The woman told Mendoza that her son, Charlie,  
   a. had never seen Mendoza’s paintings  
   b. loved Mendoza’s paintings  
   c. did not like Mendoza’s paintings  
   d. I don’t know

7. When the woman showed Mendoza her son’s drawing, Mendoza thought it was  
   a. bad   b. average   c. good   d. I don’t know

8. How much money did the woman finally pay Mendoza?  
   a. $100   b. $1,000   c. $10,000   d. I don’t know

9. What did Mendoza tell Charlie when Charlie and his mother came to see him?  
   a. that his drawings were no good and he should not become a painter  
   b. that his drawings were good and he should become a painter  
   c. that he should take drawing lessons  
   d. I don’t know

10. What happened to Charlie after he left Mendoza’s house?  

11. Did Charlie continue to live with his mother?  
   a. Yes   b. No   c. I don’t know

12. Did Charlie become a painter?  
   a. Yes   b. No   c. I don’t know

13. Why did Charlie come to see Mendoza the second time?  

14. Did Charlie know who gave him the money?  
   a. Yes   b. No   c. I don’t know

15. Where do you think the money that Charlie was given on the street came from?
Shattering the hierarchical education system:
The creation of a poststructural feminist English classroom

Ya-huei Wang
Department of Applied Foreign Languages, Chung-Shan Medical University, Taiwan; Department of Medical Education, Chung Shan Medical University Hospital, Taiwan
yhuei@csmu.edu.tw

Bio Data
Ya-huei Wang is an Associate Professor at Chung-Shan Medical University. She received her PhD in education from National Changhwa University of Education, Taiwan. Her research interests include learning strategies and applying effective motivational strategies to class instruction. Her papers have been published by Expert Systems with Applications, the British Journal of Educational Technology, the Asian EFL Journal Quarterly, etc.

Abstract
The study aims to examine the effectiveness of the poststructural feminist pedagogical model developed by the author in 2006 and apply the model to the English classroom to investigate whether it has any potential to increase Taiwanese students’ English learning achievement, critical thinking ability, and satisfaction with their class. The quantitative methods used in the study are an English achievement test, a critical thinking ability test, and a student satisfaction questionnaire. The qualitative methods are a semi-structured questionnaire and interviews. An independent sample t-test was used to determine if there were any statistically significant differences in the means between the students in the traditional classroom and the students in the poststructural feminist classroom. The research results show that the poststructural feminist pedagogical model has positive effects upon the participants in the experimental group. Several conclusions are elicited from the study. First, in the English language achievement, the students instructed according to the poststructural feminist pedagogical model significantly outperform those receiving the traditional banking instruction. Second, they are equipped with significantly better critical thinking ability. Third, they express significantly greater satisfaction than those receiving traditional banking instruction.

Keywords: Poststructural feminist pedagogy, Banking education, Resistance, Silence

Introduction
The traditional Chinese class is characterized as one in which students memorize lessons from textbooks and are generally passive receivers of information (Krause &
In this kind of class, students are instructed with the traditional teaching methods, mainly grammar translation and lecturing. Teachers analyze the grammatical construction of each single sentence or single phrase, expecting students to know the meaning of every single word. In this kind of teaching, students have been trained to seek correct answers and interpretations, which are good for equipping students with language knowledge necessary for passing entrance examination in Taiwan (Chen, 2002). However, in some way, it undermines the attempt to generate an atmosphere of inquiry and genuine dialogue, which is essential to help students become critical and reflective thinkers (Apol, 1998; Shannon, 1995). Moreover, having been rooted in a collective-oriented culture—unlike Western students, who are adventurous in nature and actively confronted with their external environment—Chinese students care too much about “losing face” and dishonoring the family name (Bond & Hwang, 1987). Hence, this “face” problem causes students to not dare to express their opinions in class and instead passively accept the teacher’s instruction. In addition, a student talking back to a teacher or questioning the content of the lesson would be interpreted as an insult to the elders (Hong, Lawrenz, & Veach, 2005). Also, to maintain a harmonious relationship, people in collectivistic cultures are inclined to protect the face of others as well as their own face (Hofstede, 1980; Ting-Toomey, 1988). Hence, high context communication occurs in collectivistic cultures, such as China, Japan, and other countries in East Asia, in which most information is expressed indirectly and implicitly in contexts to promote interpersonal harmony (Hofstede, 2001). In consideration of mutual face, people of collectivistic, high-context cultures are inclined to manage conflicts with others by using avoiding, obliging, integrating, or compromising styles in order not to cause disrespect and disrupt harmony (Ting-Toomey, 1988).

This is especially true for female students in Taiwan, who are strongly deterred from airing their opinions, as to do so would be a sign of poor manners (Canada & Pringle, 1995). Therefore, in order to be considered polite, respectful, and well-raised, females in Taiwan are forced to remain silent, thereby creating the required ‘good impression’. In the dominant hierarchical and androcentric classroom structure, students, especially female students, are discouraged from giving voice to their thoughts and opinion; thus Chinese students do not participate in class discussion (Krause & O’Brien, 2001; Salili, 1995). They are therefore, to an extent, marginalized and exist outside their knowledge construction (Maher & Tetreault, 1994).
The authoritarian Chinese classroom structure is teacher-centered. The knowledge transmission is a one-way flow, from professor-to-student. This one-way instruction process is what Freire (1970) called the “banking” education system. In this banking education system, the teacher is regarded as the sole authority in terms of their professional knowledge and expertise, thereby creating a hierarchical relationship between teachers and students.

Unlike the “banking” education, poststructural feminist educators attempt to develop students’ authority and view knowledge as socially constructed and culturally bound. The promotion of multiple authorities allows different classroom dynamics and voices to emerge (Bakhtin, 1981; English, 2005; Maher & Tetreault, 1994; Tisdell, 1998).

The purpose of this research is to verify the poststructural feminist pedagogical model and then to explore the effects of the application of this model on student achievement in learning English, on student critical thinking ability, and on student satisfaction levels.

**Poststructural feminist pedagogy**

Poststructural feminist pedagogy assumes that reality is socially constructed through language and narratives (Butler, 1990). Berger and Luckmann (1967) claim that the social construction of reality refers to the process that people actively get involved in creating the reality and knowledge through social and intellectual interaction. Language is the essential vehicle to construct knowledge and reality, such as conversations and communication in learning activities (Gergen, 1985). Through the ongoing language interactions, meaning is continuously constructed, deconstructed, and revised. As Mead (1934) says on symbolic interactionism, meanings are created by human beings through their ability to use symbols and to socially interact with others. Therefore, it is necessary to challenge and deconstruct the “dominant discourse” (Hartsock, 1987). Poststructural feminist pedagogy aims to challenge the assumptions of binaried categories—not only the teacher-student category but also the male-female category—embedded in this traditional view of authority and power. Poststructural feminist pedagogues believe that communication and learning undertaken in the classroom should be the responsibility of both teachers and students. By blurring the teacher-student responsibility and the power relations in learning and communicating, the notion that students should be silent recipients of knowledge
while teachers the powerful speakers is removed (Ropers-Huilman, 1996).

Poststructural feminist pedagogy questions the role and authority of teachers. Teachers’ authority stems mainly from two sources: their knowledge and the power invested in the position of being teachers. This being the case, teachers, engaging in banking concept practices, thus focus students’ attention on the formal elements of writing, so there is little room left for critical content. To a certain extent, therefore, teachers regress to the position of oppressive masters, and students regress to passives slaves (Adorno & Horkheimer, 1995). Poststructural feminist pedagogy intends to not only deconstruct the patriarchal (subject/object, active/passive) education structure between teachers and students, but also to deconstruct the unbalanced (center-margin) class structure relationship between male students and female students (Flax, 1990; Martin, 1988; Orner, 1992).

The socially constructed inequity in the distribution of power lies in a discourse based on hierarchical and androcentric principles in which teachers and male students are always the subjects in control of classroom interaction, while female students are always silenced and trivialized in class. Moreover, both male and female teachers have paid more attention to male students than female students. Hence, female students are doubly oppressed, both by teachers and male students (Hopf & Hatzichristous, 1999; Smith, 1992). Thus being marginalized in the masculine and hierarchical orders, female students become not a subject but a passive object of the education systems. The poststructural feminist pedagogy indicates that all the students, no matter whether they are males or females, should be treated and paid attention equally to deconstruct the hierarchical and androcentric dominance in the classroom. Hence, poststructural feminist pedagogy intends to empower female students or marginalized students and give them a voice and status they have lacked in the traditional power structures.

According to Tisdell (1998), a poststructural feminist education consists of four elements. First, poststructural feminist discourse calls for the recognition of the significance of gender with other structural systems of privilege and oppression. In other words, the intersection of gender with other systems of oppression and privilege, such as race and class, is the key to the construction of the self in feminist poststructuralism. Second, all poststructural feminist theories question the notion of a single “truth,” declaring that there are many different “truths” in varying contexts. However, in order to control the oppressed groups, the dominant social system
controls what has counted as “knowledge” and “truth” when determining the official curriculum through the politics of the knowledge production process. Third, poststructural feminist education believes that identity is in a state of flux and therefore always shifting. Hence, while one consciously scrutinizes the impact of social systems of privilege and oppression on one’s identity, the “discourse” is constantly disrupted and shifting. Fourth, the poststructural feminist practice deconstructs categories and binary opposites such as teacher-student, male-female, subject-object, speaking-silence, etc.

While deconstructing the notion of a single “truth” and the binary opposites, poststructural feminist classrooms also celebrate and encourage difference and diversity, aiming to establish a classroom atmosphere in which all participants—teachers and students—are treated with equality and respect. By blurring and equalizing the power relations between teachers and students, students in some way take over their learning process. The notion that students should be silent recipients of knowledge and teachers powerful speakers is therefore removed (Ropers-Huilman, 1996). Hence, both teachers and students undertake and share the responsibility to be knowledge negotiators, rather than solely knowledge “providers” or “receivers.”

In the classroom, knowledge negotiation and communication can proceed through speech or silence. Influenced by Chinese culture, people in Taiwan have been described as “reticent, emotionally restrained, and harmony-oriented” (Kuo, 1992). Hence they would rather silence themselves than express their disagreement with others to avoid interaction conflicts (Takahashi & Beebe, 1993). However, although a voice can be associated with empowerment and knowledge construction in the process of knowledge negotiation and communication, in the poststructural feminist class, silence could also be interpreted as resistance, power, and the active construction of knowledge (Belenky, Cliachy, Goldberger, & Tarule, 1986). Therefore, silence and speech have different meanings and implications, depending on certain situational factors and the identities of the participants within the classroom instructions (Ropers-Huilman, 1996). According to O’Barr and Wyer (1992), both speech and silence are meaningful and powerful, for they have the potential to be liberatory or repressive. Poststructural feminist practice grapples with these tenuous and shifting relations of language in educational environments.

Through a series of literature reviews (Bakhtin, 1981; Butler, 1990; De Lauretis,
1984; Foucault, 1982; hooks, 1989; Maher & Tetreault, 1994; Orner, 1992; Weiler, 1991), a poststructural feminist pedagogical dynamic model was initially developed in 2006 and revised by the researcher. The slightly revised model is shown in Figure 1 and described as follows.

**Creating a Positive Participatory Learning Environment**

Poststructural feminist pedagogy aims to foster a secure and comfortable non-competitive learning environment, ensuring that all students have not only equal opportunity to speak out about their experiences but also have equal powers in influencing decision-making both inside and outside of the classroom, in which high levels of trust, personal commitment, gender equality, and democratic dialogue are possible.

**Initiating Empowerment and Subjectivity**

Instead of seeking a reversal of the patriarchal power structure, poststructural feminist pedagogy seeks to empower or give students a voice. “Empowerment” is a key concept in the poststructural feminist class, which seeks to interrupt the reinforced patriarchal dominance in the classroom and give power equally to all students, including female and marginalized students.
**Incorporating Life Experiences into Teaching**

In order to empower students, poststructural feminist pedagogues declare that students find their voice most naturally when the material they are studying is relevant and connected to their lives. Hence, the teaching practices should incorporate life experiences from students. Inviting student experiences into academic discussion not only makes intellectual issues come to life but also keeps students engaged in the process of learning interactions.

**Listening to Student Voices**

Poststructural feminist pedagogues engage in a purposeful process of listening to student voices, be that a speaking voice or a silent voice. In being encouraged to raise their “authentic voice” in the public space, students make themselves visible and have a chance to define themselves as authors of their own world.

**Fostering Dialogical Interactions**

Poststructural feminist pedagogues believe that the development of an egalitarian classroom atmosphere can best be accomplished through the use of dialogical interactions. Through dialogue, both students and teachers are given an opportunity to have their voices heard. Multiple viewpoints can also be aired and considered through this form of communication and interaction.

**Embracing Diversity**

Poststructural feminist classrooms encourage both teachers and students to appreciate the differences among students, such as racial difference, gender difference, skill difference, etc. Both teachers and students should recognize and respect the reality that students enter the classroom at different levels of power. These differences in power exist between students as well as between students and the teacher.

**Understanding Student Resistance/Silence**

Poststructural feminist pedagogy declares that student silence can be a political act. Hence, teachers should be attentive to those who have no voice and to those who are unable or unwilling to give voice to their thoughts. The voice and utterance are implicitly present in the silence, which could be used as a weapon of resistance toward dominant classroom interactions.
Incorporating gender education into English classroom

Gender issues are incorporated into the poststructural feminist class, in which both teachers and students, especially male students, can realize that females, in their socially constructed roles, are deliberately marginalized in the masculine order and subjugated into the position of “other.”

Focusing on constant self-reflection on instructional process and identity development

Both teachers and students should undertake constant self-reflection in order to realize how unjust knowledge and identity are built up in an authoritative society. For students, self-reflection helps deconstruct the dominant education system. For teachers, constant self-reflection could make them consciously aware of whether they are reproducing the dominant social order in their class.

Hypotheses

This study will achieve these aims by testing the following hypotheses:

Hypothesis 1: College students in the poststructural feminist English class will have better English learning achievements than those in the traditional English class.

English learning achievement tests are directly related to certain English courses with the purpose of assessing and evaluating how successful students have been in achieving their learning objectives (Hughes, 2003). There are two kinds of achievement tests: final achievement tests and progress achievement tests. In this study, final achievement tests are taken to measure the progress students have made.

Hypothesis 2: College students in the poststructural feminist English class will have superior critical thinking ability to those in the traditional class.

Critical thinking ability is an ability which students can use to improve their thinking quality by skillfully managing their thought structures and the surrounding intellectual stimuli (Paul & Elder, 2001). Through reasonable and reflective thinking focusing on deciding what to believe or what to do (Ennis, 1984), students are offered a chance to freely express their own ideas, demonstrate the interrelationships among their ideas, and generate a higher level of critical thinking using their own ideas (Walstad & Becker, 1994).

Hypothesis 3: College students in the poststructural feminist English class will attain a greater level of satisfaction in their English class than those in the
traditional class.

Evaluation of coursework is crucially important to colleges, for students’ educational satisfaction is defined as students’ willingness to recommend a class or program to others (Chadwick & Ward, 1987). To enhance a mutual understanding of expectations between students and faculty (Cross & Angelo, 1988), the use of student evaluation survey after the completion of instruction is necessary, for the collected data and comments can provide a systematic feedback on the effectiveness and the quality assurance of class instruction.

**Research method**

**Participants**

After the pretest results of the English achievement test and the critical thinking ability test, two homogeneous and normally distributed classes were chosen as the experimental group and the control group for this study. Also, all incoming freshmen in this university were required to take the English placement test at the beginning of the academic year to demonstrate students’ general level of English. Based on the test results, the students in the experimental group and control group were placed in the intermediate level class, which indicated that they were at about the same proficiency level. By the flipping of a coin, the class of B1 was chosen as the experimental group, with class C1 the control group. The participants were freshmen, who had studied English for at least six years since junior high school. There were a total of 69 participants in the study. The experimental group consisted of 34 students—18 female students and 16 male students. The control group consisted of 35 students—20 female students and 15 male students. Since the number of participants in the control group and the experimental group was unequal, the unequal pretest-posttest design was adopted in this study.

**Experimental design and procedure**

A quasi-experimental design was used in this study because random assignment of students to classes was not possible. The experiment took place during the 2008 fall semester and was implemented for eight weeks, four periods a week. The researcher taught both the control group and the experimental group. Both groups received the same teaching materials, homework, and evaluation procedure, but are situated in different learning environments. The pretest-posttest experimental design is illustrated in the diagram below.
In the experimental learning environment, as mentioned previously, the teacher’s authority was downplayed by creating an egalitarian and positive learning climate, shown in Figure 1. The students in the control group were instructed with the traditional teaching methods, mainly the grammar translation and lecturing methods. In this classroom, the teacher tended to analyze the grammatical construction of every single sentence or single phrase, expecting the students to know the meaning of every single word. Hence, the entirety of the instruction was limited to an activity performed by an authority serving as information provider, who stood in the front of a classroom and verbally distributed factual information to students. However, in the experimental group, the teacher in this class served as a facilitator as well as a monitor, working around the groups to offer guidance and assistance. The seats in the control group were arranged in traditional straight rows, with the instructor standing on the platform at the front of the classroom, alienating the students and adopting the role of the authority in the provision of information. Differing from the seat arrangement in the control group, the seats in the experimental group were arranged in circles, consisting of four people, so that the instructor could stay within the group and with students. In the learning circles, students could have immediate interactions with peers and the teacher.

Although, the teaching materials had communicative activities, as the traditional approaches used in Taiwan, the students in the control group had a chance to speak English through dialogue practices in the textbook. However, an atmosphere of inquiry and genuine dialogical interaction was not encouraged. Nevertheless, the
experimental students bore the responsibility to constantly reflect upon the proposed speaking topics, being empowered to share their opinions.

Prior to the experiment, the English learning achievement test and critical thinking ability tests were administered to both the experimental group and the control group to determine students’ English language ability and critical thinking ability. The results showed that both groups were at about the same proficiency level. During this study, a series of interviews were also conducted, allowing the students to detail their experiences of the poststructural feminist English class interactions and to openly share their experiences, thoughts, and feelings with the teacher. The initial and subsequent interviews were open-ended, semi-structured interviews. To foster a comfortable interviewing environment, all interviews were conducted in Mandarin Chinese and were not audio taped, but only memo-noted. Moreover, to generate convincing interpretations, after translating the interviews, the researcher asked a bilingual teacher to examine the translated data. After the eight-week experimental intervention, the English learning achievement post-test and critical thinking ability post-test were administered to both the experimental group and the control group. Additionally, both groups were required to fill out a post-experimental student satisfaction questionnaire relating to this English course. Because this questionnaire was anonymous, the participants were free to express their true feelings about the class instruction.

**Teaching procedure and content**

The teaching material used in these two classes was from an intermediate level teaching material—*World Link: Developing English Fluency* (Stemppleski, Morgan, & Douglas, 1995), published by Thomson ELT, a leading provider of materials for English language teaching and learning throughout the world. The *World Link* textbook, combining dynamic vocabulary with essential grammar and universal topics, is a core series for young adults learning to speak and communicate English confidently and fluently. The teaching content is as follows (Table 1):
### Table 1
The teaching content

<table>
<thead>
<tr>
<th>Week</th>
<th>Unit &amp; Lesson</th>
<th>Vocabulary</th>
<th>Listening</th>
<th>Speaking</th>
<th>Grammar</th>
<th>Reading &amp; Writing</th>
</tr>
</thead>
</table>
| 1    | Unit 1: New Friends, New Faces  
Lesson A: Meeting new people | Online pen pals: age, city, e-mail address, interest... | “Hi, I’m Fiona.” Listening for interests & details | “Nice to meet you.” Introducing yourself      | Review of the simple present                   | “Celebrity doubles” Reading about people who look like celebrities |
|      |                        |                                                                            |                                                                             |                                               |                                               |                                                 |
| 2    | Unit 1: New Friends, New Faces  
Lesson B: What does he look like? | He’s in his fifties: age, eye color, hairstyle, e-mail address, interest... | “What does he look like?” Listening for details of appearance | Question intonation review                     | Describing people                              | “Guess who?” Writing about a classmate         |
| 3    | Unit 2: Express Yourself!  
Lesson A: Feelings | “How do they feel?” angry, bored, nervous, scared... | “Why are you smiling?” Listening for gist and feelings | “How’s it going?” Asking how someone is       | Review of the present continuous               | “World Greetings” Reading about gestures in three countries |
|      |                        |                                                                            |                                                                             |                                               |                                               |                                                 |
| 4    | Unit 2: Express Yourself!  
Lesson B: Body language and gestures | Gestures: shake hands, bow, kiss, point... | “What are they doing?” Listening for gestures and details | Linking sounds with ‘s’                       | Object pronouns                               | “Smiling in e-mail” Writing an e-mail using smileys and abbreviations |
| 5    | Unit 3: What Do We Need?  
Lesson A: At the supermarket | At the supermarket: apples, chicken, fish, lettuce... | “Shopping List” Listening for food items | “We need potatoes.” Expressing need           | Count and noncount nouns with some and any     | “Garage sale bargains” Reading about bargain hunting at garage sales |
| 6    | Unit 3: What Do We Need?  
Lesson B: Let’s go shopping! | At the mall: drugstore, toy store, stationery store... | “Flea markets” Listening for business hours and merchandise | Weak vowel sounds | Some/any; much/many; a lot of                | “My favorite place to shop” Writing about a favorite place to shop |
| 7    | Unit 4: Around the World  
Lesson A: Places in my city | In the neighborhood: bank, gym, Internet, café, library... | “Where are they?” Listening for location and details | “Is there a theater near here?” Asking for and giving directions | Prepositions of place                          | “The best cities to live in” Reading an article about two great cities to live in |
| 8    | Unit 4: Around the World  
Lesson B: Cities around the world | Two cities: cost of living, crime, pollution... | “In the suburbs” Listening for cities and description | Sentence stress | How much/how many?                          | “My city” Writing about a city you know        |

### Instrumentation

**English learning achievement test**

All participants had to take the pre-test and post-test based on the database of the Longman English Interactive (LEI) (Bakin, 2004), before and after the experiment. The LEI is a comprehensive software program covering grammatical content, reading skills, listening comprehension, vocabulary, etc., and was reviewed by many experienced English teachers and experts, including Prof. Bakin, Prof. Biache, Prof.
Rost, Prof. Chapelle, and Prof. Jamieson, in 2004, who have years of teaching experience at the universities in the United States, for instance, the University of California, Berkeley. Therefore, the LEI testing program can be considered a valid testing program that is presented in well-organized sections.

**Critical thinking ability test**

Critical thinking ability pre- and post-tests were used to measure the students’ reflective thinking and their capacity to organize, synthesize, and express knowledge. The critical thinking ability test was initially reviewed by three experienced English teachers. After the test, two raters carried out the grading of the critical thinking ability test papers based on the criterion developed by the researcher. The measures of the Pearson product-moment correlations between the first and second raters of the critical thinking ability test in the experimental group and the control group are reported in Tables 2 and 3. The resulting correlation coefficients reflect the overall agreement of the two raters. Based on these inter-rater comparisons, the reliability estimates are between 0.70 and 0.86, and all the p-values are less than 0.01. Hence, it can be concluded that the critical thinking ability test yielded consistent and reliable results.

**Table 2**

<table>
<thead>
<tr>
<th>Test</th>
<th>Correlation Coefficient</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Length</td>
<td>0.86</td>
<td>0.00**</td>
</tr>
<tr>
<td>Focus</td>
<td>0.77</td>
<td>0.00**</td>
</tr>
<tr>
<td>Content</td>
<td>0.78</td>
<td>0.00**</td>
</tr>
<tr>
<td>Organization</td>
<td>0.79</td>
<td>0.00**</td>
</tr>
<tr>
<td>Style</td>
<td>0.71</td>
<td>0.00**</td>
</tr>
<tr>
<td>Overall</td>
<td>0.85</td>
<td>0.00**</td>
</tr>
</tbody>
</table>

**Table 3**

<table>
<thead>
<tr>
<th>Test</th>
<th>Correlation Coefficient</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Length</td>
<td>0.84</td>
<td>0.00**</td>
</tr>
<tr>
<td>Focus</td>
<td>0.78</td>
<td>0.00**</td>
</tr>
<tr>
<td>Content</td>
<td>0.76</td>
<td>0.00**</td>
</tr>
<tr>
<td>Organization</td>
<td>0.75</td>
<td>0.00**</td>
</tr>
<tr>
<td>Style</td>
<td>0.70</td>
<td>0.00**</td>
</tr>
<tr>
<td>Overall</td>
<td>0.83</td>
<td>0.00**</td>
</tr>
</tbody>
</table>

**Note:**

**p < 0.01**
Student satisfactory questionnaire

The student satisfaction questionnaire was made up of 34 multiple-choice questions with a 5-point Likert scale varying from “strongly disagree” to “strongly agree” (Wang & Liao, 2008). The study used Cronbach’s Alpha to test the internal reliability of each category of the questionnaire. Generally speaking, the minimal acceptable level of reliability is 0.7, though 0.8 or greater is preferable (Churchill, 1979). That is, the nearer the result to 0.8, the more internally reliable the scale. The reliability coefficients for the categories in the student satisfaction questionnaire are shown in Table 4.

Table 4
Reliability coefficients/Cronbach’s Alpha for the categories in the student satisfaction questionnaire

<table>
<thead>
<tr>
<th>Category</th>
<th>Control Group</th>
<th>Experimental Group</th>
</tr>
</thead>
<tbody>
<tr>
<td>Instructional Objective</td>
<td>0.7576</td>
<td>0.7669</td>
</tr>
<tr>
<td>Instructional Material/Method</td>
<td>0.8398</td>
<td>0.8772</td>
</tr>
<tr>
<td>Teacher’s Qualities</td>
<td>0.9092</td>
<td>0.8819</td>
</tr>
<tr>
<td>Class Climate/Environment</td>
<td>0.8601</td>
<td>0.9139</td>
</tr>
<tr>
<td>Assessment</td>
<td>0.8247</td>
<td>0.8424</td>
</tr>
<tr>
<td>Overall</td>
<td>0.9479</td>
<td>0.9454</td>
</tr>
</tbody>
</table>

In addition, the Pearson correlation coefficient between any two categories in the student satisfaction questionnaire was calculated. The resulting correlation coefficients were between 0.397 and 0.760. The p-value was less than 0.05, which indicates that there is a significant correlation between any two categories in the student satisfaction questionnaire. The reliability coefficients for the categories in the student satisfaction questionnaire are shown in Table 5.

Table 5
Pearson correlation coefficient analysis of the experimental group in the student satisfaction questionnaire

<table>
<thead>
<tr>
<th></th>
<th>Instructional Objective</th>
<th>Teaching Material/Method</th>
<th>Teacher’s Qualities</th>
<th>Class Climate/Environment</th>
<th>Assessment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Instructional Objective</td>
<td>1</td>
<td>0.643**</td>
<td>0.407*</td>
<td>0.397*</td>
<td>0.604**</td>
</tr>
<tr>
<td>Teaching Material/Method</td>
<td>0.000**</td>
<td>1</td>
<td>0.626**</td>
<td>0.625**</td>
<td>0.557**</td>
</tr>
<tr>
<td>Teacher’s Qualities</td>
<td>0.000**</td>
<td>0.000**</td>
<td>1</td>
<td>0.760**</td>
<td>0.398*</td>
</tr>
<tr>
<td>Class Climate/Environment</td>
<td>0.000**</td>
<td>0.000**</td>
<td>0.000**</td>
<td>1</td>
<td>0.543**</td>
</tr>
</tbody>
</table>
Initial and subsequent in-depth student interviews with individuals or groups

With the aim of eliciting collective qualitative data on the students’ difficulties and anxieties in regard to class interaction, the student interviews were conducted in a semi-structured format, following the techniques of semi-structured interviews. By conducting initial semi-structured interviews with participants, the historical background of the participants was collected. Also, the initial interviews allowed the researcher to establish a rapport and intimacy with the participants. Subsequent interviews were less structured to allow participants to elaborate on their experiences in the poststructural feminist English class.

Data Analysis

The collected data were analyzed quantitatively and qualitatively. The statistical package used to analyze the quantitative data was SPSS (Statistical Packages for the Social Science). Independent sample t-tests were used to determine if there were any statistically significant differences in the means between the students in the traditional classroom and the students in the poststructural feminist classroom. Additionally, based on building a holistic and complex understanding of students’ reactions and classroom interactions, qualitative data analysis was used.

Results

Results of Hypothesis Testing

Hypothesis 1: College students in the poststructural feminist English class will have better English learning achievements than those in the traditional English class.

To test Hypothesis 1, the results of the English achievement pre- and post-test of both groups were examined via t-tests and compared as follows.

By comparing the two groups in terms of percentage rates, based on the scores on the English achievement pre- and post-test, it is shown that the poststructural feminist pedagogical model has significant effects on the students’ English learning achievements. As shown in Table 6, in the English achievement pre-test there were no
significant differences between the control group and the experimental group in terms of listening, vocabulary, grammar, reading, and the overall section. After the treatment, the scores in the post-tests for both groups were examined, revealing significant differences between these two groups in English learning achievement in terms of listening, vocabulary, grammar, reading, and the overall section.

After comparing the English proficiency scores of the male students in these two groups in terms of percentage rates, it was found that in the pre-test, as shown in Table 6, there were no significant differences between the male students in these two groups in terms of listening, vocabulary, grammar, reading, and the overall section. After the treatment, there are still no significant differences in the vocabulary and reading sections. However, the means of the male students in the experimental group in the listening, grammar, and overall sections are higher than the means of the control group.

As for the female students, there were initially no significant differences between the females in both groups in all of the sections. After the treatment, all the means of the females in the experimental group are significantly higher than the means of the females in the control group.

Table 6. Comparison between the both groups based on the scores on the English achievement pre- and post-test

<table>
<thead>
<tr>
<th></th>
<th>Listening</th>
<th>Vocabulary</th>
<th>Grammar</th>
<th>Reading</th>
<th>Overall</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pre-test</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Post-test</td>
<td>*</td>
<td>*</td>
<td>**</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>x%</td>
<td>10.45%</td>
<td>8.75%</td>
<td>21.58%</td>
<td>8.81%</td>
<td>10.69%</td>
</tr>
<tr>
<td>Male Pre-test</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Male Post-test</td>
<td>*</td>
<td>--</td>
<td>**</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>y%</td>
<td>9.28%</td>
<td>30.15%</td>
<td>23.25%</td>
<td>10.78%</td>
<td>13.22%</td>
</tr>
<tr>
<td>Female Pre-test</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Female Post-test</td>
<td>*</td>
<td>**</td>
<td>**</td>
<td>*</td>
<td>*</td>
</tr>
</tbody>
</table>

Experimental Group: N=34; Male: N=16; Female: N=18
Control Group: N=35; Male: N=15; Female: N=20

\[ x\% = \frac{\text{The mean of experiment group} - \text{The mean of control group}}{\text{The mean of control group}} \]

\[ y\% = \frac{\text{The mean of the males in the experiment group} - \text{The mean of the males in the control group}}{\text{The mean of the males in the control group}} \]

\[ z\% = \frac{\text{The mean of the females in the experiment group} - \text{The mean of the females in the control group}}{\text{The mean of the females in the control group}} \]

*\( P < 0.05; **P < 0.01 \)
A comparison between the male students and female students in the experimental group in terms of percentage rates, as shown in Table 7, revealed that in the English achievement pre-test there were no significant differences between the male students and the female students in the experimental group in terms of listening, vocabulary, grammar, reading, and overall performance. After the treatment, there are still no significant differences between the males and females in the listening and vocabulary sections. However, in the grammar, reading, and overall sections, the means of the females in the experimental group are significantly higher than the means of the males in the experimental group.

Table 7
Comparison between the male and female students in the experimental group based on the scores of the English achievement pre- and post-test

<table>
<thead>
<tr>
<th></th>
<th>Listening</th>
<th>Vocabulary</th>
<th>Grammar</th>
<th>Reading</th>
<th>Overall</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pre-test</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Post-test</td>
<td>--</td>
<td>--</td>
<td>*</td>
<td>*</td>
<td>**</td>
</tr>
</tbody>
</table>

In conclusion, from the data shown in Table 6, it is clearly demonstrated that those instructed according to the poststructural feminist pedagogical model generally outperform those receiving the traditional English instruction. Additionally, as shown in Table 7, in the grammar and overall performance, the female students in the experimental group outperform the males in the same group. Therefore, it can be inferred that this form of pedagogy gives female students more confidence and has a more positive effect in improving their English language achievement.

Hypothesis 2: College students in the poststructural feminist English class will have superior critical thinking ability to those in the traditional class.

To test Hypothesis, the results of the critical thinking ability test in the experimental group and the control group were examined via t-test. The results are illustrated in Table 8.

As shown in Table 8, there were initially no significant differences between the two groups in the critical thinking ability test, which indicates that the experimental group and control group were homogeneous in the critical thinking ability test in
terms of length, focus, content, organization, style, and overall rating. After the treatment, there is still no significant difference in the style rating. However, the means of the experimental group in the length, focus, content, organization, and overall ratings are higher than the means of the control group.

When comparing the ratings of the critical thinking ability between the male students in these two groups in terms of percentage rates, as shown in Table 8, it was found that in the pre-test there were initially no significant differences between the male students in these two groups in all categories. After the treatment, the means of the male students in the experimental group are higher than the means of the control group in the length, focus, content, organization, and overall ratings.

As for the female students, as shown in Table 8, there were initially no significant differences between the females in both groups in the critical thinking ability test. After the treatment, excluding the means of the style section, all the means of the females in the experimental group are significantly higher than the means of the females in the control group.

Table 8
Comparison between the two groups based on the scores of the critical thinking ability pre- and post-test

<table>
<thead>
<tr>
<th></th>
<th>Length</th>
<th>Focus</th>
<th>Content</th>
<th>Organization</th>
<th>Style</th>
<th>Overall</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pre-test</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Post-test x%</td>
<td>225.68%</td>
<td>32.28%</td>
<td>18.81%</td>
<td>19.03%</td>
<td>--</td>
<td>**</td>
</tr>
<tr>
<td>Male</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Pre-test</td>
<td>**</td>
<td>*</td>
<td>*</td>
<td>**</td>
<td>--</td>
<td>**</td>
</tr>
<tr>
<td>Male Post-test y%</td>
<td>227.81%</td>
<td>33.15%</td>
<td>15.19%</td>
<td>15.05%</td>
<td>--</td>
<td>**</td>
</tr>
<tr>
<td>Female</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Pre-test</td>
<td>**</td>
<td>*</td>
<td>*</td>
<td>**</td>
<td>--</td>
<td>**</td>
</tr>
<tr>
<td>Female Post-test</td>
<td>222.22%</td>
<td>31.98%</td>
<td>20.13%</td>
<td>12.43%</td>
<td>--</td>
<td>**</td>
</tr>
</tbody>
</table>

Experimental Group: N=34; Male: N=16; Female: N=18
Control Group: N=35; Male: N=15; Female: N=20

\[ x\% = \frac{\text{The mean of experiment group} - \text{The mean of control group}}{\text{The mean of control group}} \]
\[ y\% = \frac{\text{The mean of the males in the experiment group} - \text{The mean of the males in the control group}}{\text{The mean of the males in the control group}} \]
\[ z\% = \frac{\text{The mean of the females in the experiment group} - \text{The mean of the females in the control group}}{\text{The mean of the females in the control group}} \]

\*P < 0.05; **P < 0.01
A comparison between the male students and female students in the experimental group in terms of percentage rates, as shown in Table 9, indicated that in the critical thinking ability pre-test there were initially no significant differences between the male students and the female students in all categories. After the treatment, there are still no significant differences between the males and the females in terms of focus, content, and style ratings. However, in the length, organization, and overall sections, the means of the females in the experimental group are significantly higher than the means of the control group.

Table 9
Comparison between the male and female students of the experimental group based on the scores in the critical thinking ability pre- and post-test

<table>
<thead>
<tr>
<th></th>
<th>Length</th>
<th>Focus</th>
<th>Content</th>
<th>Organization</th>
<th>Style</th>
<th>Overall</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pre-test</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Post-test</td>
<td>**</td>
<td>--</td>
<td>--</td>
<td>**</td>
<td>--</td>
<td>*</td>
</tr>
<tr>
<td>y%</td>
<td>5.51%</td>
<td>--</td>
<td>--</td>
<td>4.98%</td>
<td>--</td>
<td>3.15%</td>
</tr>
</tbody>
</table>

Experimental Group: N=34; Male: N=16; Female: N=18

\[ y\% = \frac{\text{The mean of the females in the experiment group} - \text{The mean of the males in the experiment group}}{\text{The mean of the males in the experiment group}} \]

\*P < 0.05; \*\*P < 0.01

In conclusion, from the data shown in Table 8, it is clearly demonstrated that those instructed according to the poststructural feminist pedagogical model outperform those receiving the traditional English instruction. Additionally, as shown in Table 9, in the length, organization, and overall ratings, the female students in the experimental group outperform the males in the same group. Therefore, it can be inferred that this form of pedagogy gives female students more confidence and has a more positive effect on the development of their critical thinking ability.

Hypothesis 3: College students in the poststructural feminist English class will attain a greater level of satisfaction in their English class than those in the traditional class.

To test Hypothesis 3, a post-experimental questionnaire, employing a 5-point Likert scale, was given to both the experimental group and the control group. Examined using t-tests, the results of both groups’ overall satisfaction are illustrated in Table 10. The participants instructed according to the poststructural feminist pedagogical model
clearly have significantly stronger satisfaction levels than those receiving the traditional English instruction.

Analysis of the results of the student satisfaction questionnaire in terms of gender, as shown in Table 10, shows that male participants instructed according to the poststructural feminist pedagogical model clearly have significantly higher satisfaction levels than those males receiving traditional English instruction in all categories except assessment.

As for the female students, as shown in Table 10, the means of the experimental group are all higher than the means of the females in the control group.

Table 10
Comparison between the two groups based on the results of the student satisfaction questionnaire

<table>
<thead>
<tr>
<th>Instructional objective</th>
<th>Instructional Material/Method</th>
<th>Teacher’s Qualities</th>
<th>Class Climate/Environment</th>
<th>Assessment</th>
<th>Overall</th>
</tr>
</thead>
<tbody>
<tr>
<td>x%</td>
<td>**</td>
<td>**</td>
<td>**</td>
<td>*</td>
<td>**</td>
</tr>
<tr>
<td>Male</td>
<td>14.55%</td>
<td>15.45%</td>
<td>16.00%</td>
<td>15.14%</td>
<td>16.40%</td>
</tr>
<tr>
<td>y%</td>
<td>*</td>
<td>**</td>
<td>*</td>
<td>*</td>
<td>--</td>
</tr>
<tr>
<td>Female</td>
<td>14.80%</td>
<td>16.78%</td>
<td>17.06%</td>
<td>19.53%</td>
<td></td>
</tr>
</tbody>
</table>

x% = \( \frac{\text{The mean of experiment group} - \text{The mean of control group}}{\text{The mean of control group}} \)

y% = \( \frac{\text{The mean of the males in the experiment group} - \text{The mean of the males in the control group}}{\text{The mean of the males in the control group}} \)

z% = \( \frac{\text{The mean of the females in the experiment group} - \text{The mean of the females in the control group}}{\text{The mean of the females in the control group}} \)

Experimental Group: N=34; Male: N=16; Female: N=18
Control Group: N=35; Male: N=15; Female: N=20

Furthermore, in the experimental group, as shown in Table 11, there are no significant differences between the male students and the female students in all categories. Table 10, above, shows that the participants instructed according to the poststructural feminist pedagogical model have a greater level of satisfaction than those receiving traditional English instruction. Therefore, it can be inferred that both male and female students attain a greater level of satisfaction when instructed according to the
poststructural feminist pedagogical model.

In conclusion, from the data displayed in Tables 10 and 11, it is clearly shown that the students instructed according to the poststructural feminist pedagogical model, no matter whether they are male or female, attain significantly higher levels of satisfaction than those receiving the traditional English instruction.

Table 11
Comparison between the male and female students of the experimental group based on the results of the student satisfaction questionnaire

<table>
<thead>
<tr>
<th>Instructional objective</th>
<th>Instructional Material/Method</th>
<th>Teacher’s Qualities</th>
<th>Class Climate/Environment</th>
<th>Assessment</th>
<th>Overall</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experimental Group %y</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
</tbody>
</table>

Experimental Group: N=34; Male: N=16; Female: N=18

\[ y\% = \frac{\text{The mean of the females in the experiment group} - \text{The mean of the males in the experiment group}}{\text{The mean of the males in the experiment group}} \]

Results of the interviews

The results of student interviews reveal that the participants in the poststructural feminist class enjoyed this English class because they were given more time and opportunities to express their opinions and to interact with their classmates as well as the teacher, Christine.

- I feel that staying at Christine’s English class is so comfortable that I feel so free to speak up in class and to express my opinions. I am not afraid of being criticized. (M2)
- It’s a nice experience to share experiences with my peers and to listen to others’ opinions. (F4)
- Before, while sitting in rows, we didn’t have many chances to interact with our classmates. Now, the frequency of face-to-face interactions is increasing. We do not fall into sleep in class anymore. (M6)
- I can freely speak my opinions even though sometimes my opinions are against the teacher’s. I am not worried about being criticized or refused. The teacher is not used to interrupting students’ opinions. She shows respect to everybody. (F8)

Some interviewees revealed how, within this form of class instruction, they felt quite safe and free to express their ideas, even though these might be different to those of
the teacher. They felt the teacher in the poststructural feminist class was willing to pay attention to students’ opinions and offer assistance to help students overcome their shyness.

- When it was my turn to speak English, I was always anxious and afraid of losing face in public. However, in this class, I have less pressure and anxiety while interacting with my classmates and the teacher. (F3)
- With Christine’s timely support and suggestions, I feel not so embarrassed to express my opinions in class. I enjoy having my voice heard in class. (F6)
- We all feel that the class is so secure and safe. While talking with each other or speaking up in class, we are not afraid of being mocked or criticized. We always get positive feedback from our classmates and our teacher. (M3)
- Every time we express our opinions in class, Christine listens to them attentively. We feel respected, and that’s why we like to talk a lot in class. (M4)

Additionally, students in the poststructural feminist classroom thought that the teacher, Christine, was willing to offer adequate and additional academic aid to them. The teacher was usually available to assist them after class.

- Whenever I have problems with my English or my career plan, I would talk to Christine. It’s nice to have someone guide me though. (F8)
- It’s nice to have a talk with Christine. I appreciate her assistance with my English and other personal problems. (F2)

Most interviewees also felt the English classroom provided a more pleasant learning environment, for there was a greater level of interaction with their classmates and with the teacher. They claimed that, in the traditional class, they were not given the opportunity to speak to the teacher in English. In the poststructural feminist class, it was a different situation. They had more opportunities to speak up in the class, and the teacher would listen to their opinions attentively. Also, they enjoyed more interaction with their peers.

- This class let me have a chance to listen to female students’ voices. I enjoy sitting in this class and listening to females’ opinions. It’s nice to realize something from females’ perspectives. (M1)
We enjoy staying at this open and comfortable class. It’s very exciting to have interactions with the teacher and the classmates, especially in English. (F7)

Furthermore, both female students and male students like the feeling of being respected. Male students also expressed their approval of being able to listen to the opinions and contributions of their female counterparts, which did not occur in the traditional class. They also enjoyed sharing their feelings with female students, claiming that this class provided an environment in which they could listen to their female classmates patiently and attentively. They felt this gave them the opportunity to learn more about their female classmates and better understand their perspectives.

I think the teacher, Christine, cares a lot about students’ feelings. She respects our opinions. It’s so nice to have my voice heard and appreciated. Now I feel it is so easy to speak up in this class. (F6)

Before, it felt so awkward to speak up in class, not to mention to speak English. But with Christine’s support, I can freely express my opinions, not afraid of being mocked or criticized. (F4)

I tried very hard to take part in class discussion, but it ended up in vain. But in this unthreatening and comfortable class, I can dare to say something in class. (F3)

In addition, the participants felt that their learning anxiety had been greatly reduced, for the class provided an unthreatening learning environment.

I feel so comfortable to say something in the class. Before I was so worried and anxious about others’ opinions. (F7)

This English class is so supportive and safe, nobody is afraid of being scorned or ridiculed. (M8)

With Christine’s constant encouragement, I feel not so anxious to share my opinions with peers. (F11)

Female students felt that they were appreciated in this poststructural feminist classroom.

I felt so scared when I had to speak English. However, in this class I am not so scared when it’s my turn to speak up in class. (F3)

Before, I was so shy and passive about saying something in class,
knowing that nobody paid attention to me. Now, everything is changed. It’s nice to participate in the class interaction and to exchange my opinions with others. (F2)

The results of the interviews demonstrate that the poststructural feminist pedagogical model has a positive effect on these students’ English learning due to the following reasons. First, the poststructural feminist pedagogical model fosters students’ communication in the English class because students have more time and opportunities to express their opinions and to interact with their classmates as well as the teacher. Both the female students and male students liked the feeling of being respected. The male students expressed an appreciation at being able to share their feelings with the female students and felt it was nice to hear females’ opinions and contributions in this class. Furthermore, the poststructural feminist pedagogical model enhances the amount of dialogical interaction between teachers and students as well as among students.

To sum up, the results of quantitative and qualitative analysis demonstrate that the students instructed according to the poststructural feminist pedagogical model, no matter whether male or female, have significantly better learning achievement in English proficiency and critical thinking ability than those exposed to the traditional instruction. In addition, the poststructural feminist English class provides these students with a higher level of satisfaction than those in the traditional class.

Discussion

The purpose of this study was to develop a poststructural feminist pedagogical model and to investigate whether the model would have a positive effect on students in terms of English learning achievement, critical thinking ability, and student satisfaction. Results of this study show that students exposed to the poststructural feminist pedagogical model achieve better learning outcomes. The findings of the data analysis are described as follows:

1. The students, especially females, instructed according to the poststructural feminist pedagogical model significantly outperform those receiving the traditional banking instruction in terms of the level of their English learning achievements, in the areas of listing, vocabulary, grammar, and reading. Based on the scores on the English achievement pre- and post-test, the means of the experimental group in the listening, vocabulary, grammar, reading, and overall
section are respectively 10.45%, 8.75%, 21.58%, 8.81%, and 10.69% higher than the means of the control group. The means of the male students in the experimental group in the listening, grammar, and overall sections are respectively 9.28%, 30.15%, and 8.23% higher than the means of the control group. However, there are no significant differences in the vocabulary and reading sections. All the means of the females in the experimental group are significantly higher, respectively 10.56%, 12.58%, 23.25%, 10.78%, and 13.22% higher than the means of the females in the control group. While taking an investigation between the male and female students in the experimental group, the researcher also finds that in the grammar, reading, and overall sections, the means of the females in the experimental group are significantly higher, 10.96%, 11.39%, and 9.98% respectively, than the means of the males in the experimental group. However, there are no significant differences in the listening and vocabulary sections. It is not surprising that the poststructural feminist pedagogical model brings positive learning outcomes to students because the model is designed to achieve this through providing a classroom which is secure, in which instruction is optimized, and in which the students are empowered and therefore engaged. In this kind of classroom, students as well as teachers can work together to improve their learning outcome. The success of the poststructural feminist classroom lies in its highly interactive learning environment, in which students can actively seek help from other students whenever difficulties arise, which would partly account for the increase in the students’ learning performance. Within a secure learning environment, students’ learning anxiety with regard to foreign languages would be reduced and their involvement in the learning process would be increased, thus optimizing their learning performance. In addition, when students become involved in the learning process, they become context-dependent learners; they can try to connect the course content to their own daily experiences to facilitate their learning (Crabtree & Sapp, 2003; Shrewsbury, 1993).

2. The students, especially females, are equipped with significantly better critical thinking ability, in terms of length, focus, content, organization, and style, in the poststructural feminist classroom.

After the treatment, except the means of the style rating, the means of the experimental group in the length, focus, content, organization, and overall ratings are respectively 225.68%, 32.28%, 18.81%, 19.03%, and 50.58% higher than the
means of the control group. The means of the male students in the experimental group are respectively 227.81%, 33.15%, 15.19%, 15.05%, and 53.38% higher than the means of the control group in the length, focus, content, organization, and overall ratings. However, there is no significant difference in the style rating. All the means of the females in the experimental group are significantly higher, respectively 222.22%, 31.98%, 20.13%, 12.43%, and 61.76% higher, than the means of the females in the control group, excluding the means of the style section. Also, the means of the females in the experimental group are significantly higher, respectively 5.51%, 4.98%, and 3.15% higher, than the means of males in the experimental group in the length, organization, and overall sections. Nonetheless, there are no significant differences in the focus, content, and style sections.

The poststructural feminist educators make use of Hooks’ “engaged pedagogy” (Hooks, 1994), taking into account not only students’ emotions but also their critical thinking in the learning process. By skillfully managing their thinking structures and the intellectual criteria around them, students can sharpen their critical thinking (Paul & Elder, 2001). Students in the experimental group have a greater opportunity to express their opinions, to display their competency, and to skillfully conceptualize, apply, synthesize, and evaluate information generated by observation, experience, reflection, reasoning, and communication (Scriven & Paul, 2003). Hence, they have developed a better critical thinking ability, displaying a reasonable and reflective thinking capacity focusing on deciding what to believe or what to do (Ennis, 1984). In addition, through mutual self-reflection, these students have the capacity to understand multiple and contradictory arguments and viewpoints. They become open-minded to different viewpoints and arguments through dialogical interactions. Furthermore, they are offered the chance to develop their potential to show their creativity and originality (Walstad & Becker, 1994).

3. The students get significantly greater satisfaction in terms of instructional objective, teaching method/material, teacher’s quality, class climate/environment, and assessment, in the poststructural feminist pedagogical classroom.

The means of the experimental group in the instructional objective, instructional material/method, teacher’s qualities, class climate/environment, assessment, and overall sections are respectively 14.55%, 15.45%, 16.00%, 15.14%, 16.40 and 15.49% higher than the means of the control group. The means of the male
students in the experimental group are significantly higher, respectively 14.80%, 16.78%, 17.06%, 19.53%, and 17.40% higher, than the means of male students in the control group. However, there is no significant difference in the assessment section. The means of the females in the experimental group are respectively 14.50%, 14.35%, 14.87%, 11.75%, 15.55% and 13.95% higher than the means of the females in the control group. In the experimental group, there are no significant differences between the male students and female students on all the scales.

The poststructural feminist classroom attempts to create a comfortable environment for the learning of English, in which students are willing to participate in class discussion. Considering that knowledge is socially constructed and culture-bound (Jackson, 1997; Lather, 1992), poststructural feminist classrooms endeavor to develop multiple authorities from a variety of sources, in particular from students’ prerequisite knowledge and their life experiences, in order to create an egalitarian and democratic classroom, in which both instructors and students need to listen to and interact with others in a respectful way. While being respected, students feel secure and comfortable in daring to speak their mind and articulate their viewpoints. Knowing that their needs and opinions can be addressed, they would have gained satisfaction from the course (Schmidt, Debevee, & Comm, 1987; Terpening, Gertner, & Pitt, 1982).

To sum up, the success of the poststructural feminist pedagogical model lies in the creation of an egalitarian classroom in which dialogical interactions are promoted between students and teachers. In turn, the dialogical classroom would also lead to a more democratic, egalitarian, empowering, and reflective learning environment through “dual-voiced” or “multiple-voiced” dialogical interactions, in which both teachers and students can share power and communicate openly with each other.

Poststructural feminist pedagogy encourages students, especially female students, to find their own voice in the classroom. The research results show that, compared to the male students in the same group, the female students in the experimental group achieved a higher level of satisfaction from this form of pedagogy. Also, compared to the students in the control group, the students in the poststructural feminist class received more opportunities to express themselves and show their capacity to organize, synthesize, and express knowledge. They also developed a better critical thinking ability, which can be defined as a kind of reasonable and reflective form of thinking that focuses on forming opinions independently (Ennis, 1984).
Conclusion and pedagogical implication

The study aims to examine the effectiveness of the poststructural feminist pedagogical model and investigate whether this model can bring any potential to increase Taiwanese students’ English learning achievement, critical thinking ability and satisfaction with their class. The results show that the students, especially female students, exposed to the poststructural feminist pedagogical model gain better English learning achievement levels, superior critical thinking ability, and greater student satisfaction level. In addition, the qualitative results show that the poststructural feminist pedagogical model can foster student communication and interaction with their classmates as well as the teacher. The study aims to create a poststructural feminist class in which dialogical interactions, consciousness-raising, and reflexivity are emphasized.

The power of the poststructural feminist class lies in its elimination of subject/object, producer/consumer, giving/receiving, and teaching/learning boundaries. By blurring these boundaries, English instructors become a part of the learners, and the learners a part of the instructors. Instructors and students become partners and co-learners in the learning process. Being empowered and given the opportunity to defend their ideas, students feel comfortable because they know that their ideas and opinions will not to be criticized and ridiculed. Therefore, with the promotion of dialogue and empowerment in the class, the students in the experimental group can improve their English learning achievement, develop a better critical thinking ability, and have greater satisfaction than those receiving the traditional banking instruction.

When implementing the poststructural feminist pedagogical model, however, teachers should be aware of some implications. While proceeding with class instruction, teachers should bear in mind that they should do their best to reduce the power differentials between themselves and their students by listening to students’ voices and giving students more power over their English learning process. Also, they should endeavor to create a secure learning environment in which students with different backgrounds and opinions feel comfortable while interacting with each other. Future studies could focus on applying the poststructural feminist pedagogical model to the teaching of other educational levels. Also, it could be applied to the instruction of other courses to bring out students’ voices in their different subjects as well as to promote discussion of issues which traditionally have remained unspoken of in the classroom, such as gender, race, class, and sexual identity. It may also be successfully
applied to students with different ability levels, such as at-risk students, dropout students and repeaters.

References


Asian EFL Journal. Volume 13, Issue 4


286


English only? inda kali eh! (not likely!) – Changing the paradigm

Breda O’Hara-Davies

Duli Pengiran Muda Al-Muhtadee Billah College, Brunei Darussalam
PhD Candidate University of Technology Sydney
Breda.OHaraDavies@student.uts.edu.au

Bio Data

Breda is based at Duli Pengiran Muda Al-Muhtadee Billah College in Brunei Darussalam. With over twenty seven years of experience Breda has worked in schools and universities in Malaysia, Thailand, UK and her native Ireland and has taught across the age range. Her research interests include bilingualism, identity construction, postcolonial contexts and world Englishes.

Abstract

We live in an era of accelerated change and innovation which demands that we adapt and evolve or risk obsolescence and exclusion. This paper traces the path of English language teaching (ELT) since its inception. It examines some of the underlying theories that have guided its policies and practices at classroom level. It also questions whether ELT has kept pace with linguistic and other developments on a wider societal scale. Using an autoethnographic approach, the paper traces the parallel journey of the author in her role as an English language teacher in Brunei Darussalam. It aims at providing a reflexive account of the ways in which critical engagement with theory have impacted her attitudes, practices and ongoing evolution as a professional.

Keywords

Brunei Darussalam, English Language Teaching, English Only.

Introduction

As English Language teachers we are operatives in a much wider network; the industry of English Language Teaching (ELT). While we may naturally focus on what happens within the confines of our classroom walls from interpersonal and educational standpoints, our work also has important social, political and cultural ramifications that cannot be ignored (Phillipson 1992; Pennycook, 1994). Therefore, we need to be mindful that we are never ‘just’ teaching the language and that the very act of teaching itself is culturally and politically loaded (Pennycook, 1994).
Traditionally, however, ELT has been underscored and propelled by a belief that the spread of the English language is a natural phenomenon that is culturally and politically neutral as well as being beneficial for all. English Language teachers have unwittingly been complicit in promoting and supporting this view through a combination of its uncritical acceptance and a tendency to avoid engaging with the wider socio-political aspects of their work (Phillipson, 1992). In fact, many of the tenets by which ELT has been guided over the years have held sway due simply to an absence or at best a paucity of challenge or critical analysis.

The origins of ELT can be traced back to colonial days when “the Empire became the central testing site” (Pennycook, 1998, p.131) for its development. Once trialled in the colonies the theories and practices were then imported into Britain. However, far from spreading the English language, it seems that the English “were sometimes arrogantly possessive” of it (Brutt-Griffler, 2002, p.57), especially in their colonies. Such a protectionist attitude may have stemmed in part from a desire to continue subjugating the local people by limiting their access to English due to the power it could bestow. In the Brunei context, when formal education began in 1912, it was with the opening of a Malay vernacular school (Public Relations Unit, MoE, 2003), suggesting a British preference for vernacular education for the local populace (Brutt-Griffler, 2002). English-medium education only became available in 1931 and did not become more widespread until the post-war period in the 1950s when it was needed for both the oil industry and a burgeoning civil service. This suggests that it was only economic and utilitarian motives that caused the British to relent and allow local Bruneians greater access to ‘their’ language.

From its beginning as a cultural propagandist product of Empire (Pennycook, 1994), ELT has gone on to assume a number of guises over the years. In the wake of colonialism, it reinvented itself as development aid with the British Council as its vehicle. At this time, its emphasis became education in ‘developing’ countries (arguably in itself an ethnocentric label that holds countries such as Britain up as the norm), especially former colonies and protectorates such as Brunei where it could continue to ensure that the educational élites were favourably disposed towards Britain (Pennycook, 1994) and the English language. The current discourse of ELT presents it as a global commodity that is part of a market-force response based on the assumption that the language has been chosen by a world that knows what it wants and needs (Pennycook, 1994, citing Hindmarsh, 1978). Regardless of whether it has been chosen or not, the current “reality of hegemony” (Canagarajah, 2004, p.141) of
the English language is something that all countries in which English is not the first language (L1) need to take into account. Failure to do so could result in perpetual powerlessness and disenfranchisement in a global context.

The current hegemonic status of the English language should not insulate ELT from being problematised and critically appraised. As busy professionals consumed with honing their craft, teachers generally accept practices as being pedagogically grounded and fail to interrogate whether such practices may instead be overtly ideologically inspired (Pennycook, 1998, citing Auerbach, 1993). The fact that they may even have deep roots in the cultural constructions of colonialism wherein monolingualism in English was deemed preferable to multilingualism across other languages is often not considered (Pennycook, 1998). The discourse of ELT has long been dominated by a set of guiding principles or tenets whose very longevity coupled with uncritical acceptance have lent them enduring credence.

The first tenet is that English is best taught monolingually followed by the notion that the ‘ideal’ teacher of English is a native speaker of the language – the second tenet. Thirdly and fourthly, the earlier and the more English is taught, the better the results, and finally, if other languages are used, the standards of English will be adversely affected. Robert Phillipson in his seminal 1992 book entitled “Linguistic Imperialism” re-designates each of these tenets as fallacies. He enumerates them as one: the monolingual fallacy, two: the native speaker fallacy, three: the early start fallacy, four: the maximum exposure fallacy, and five: the subtractive fallacy (Phillipson, 1992).

This paper focuses mainly on the first and fifth tenets because insistence on the first necessarily impacts the fifth and both are in turn impacted by the second tenet which concerns the linguistic background of the teacher involved.

Whether one considers these tenets to be valid or agrees that they are fallacious, one cannot but evaluate them within the reality of diverse ELT contexts in the 21st century. This calls into question the issue of the dynamism of ELT since many of these tenets were formulated in the 1960s and seem to have remained largely unchanged in the interim. Such stagnation and domination of traditionally held views cannot be regarded as reflecting students’ needs in our globalised, technological world.

However, slaying the dogma of monolingualism is far from easy. Many at both institutional and individual levels adhere steadfastly to the belief that ‘good’ language teaching means exclusive use of the target language in the classroom, in
keeping with the Direct Method or Communicative Approach which tend to characterise learners’ L1 as an “enemy that must be held at bay” (Ibarra, 2009, p.4) in the pursuit of successful acquisition of the target language.

Those who challenge the monolingual tenet on the other hand view things very differently. For them, resisting the L1 of language learners is futile since it is always going to be present in the language classroom even if it is silenced and banned from students’ lips. It cannot be banned from their minds (Ibarra, 2009). The L1 acts as a default code (Butzkamm, 2003) to which students will naturally revert especially in the early stages of learning a second or other language. For this reason “translation/transfer is a natural phenomenon and an inevitable part of second language acquisition... regardless of whether or not the teacher offers or ‘permits’ translation” (Butzkamm, 2003, p.31, citing Harbord, 1992). Taken from this vantage point, only a foolhardy language teacher would take on the combined forces of nature and inevitability.

Languages by their very nature are dynamic and innovative as they are constantly growing and changing to meet the expressive needs of those who use them. In the case of English, which is currently regarded as the world’s first “global language” (Crystal, 2003, p.1), growth and creativity are even more pronounced given its task of giving voice and text to many of the new concepts and commodities that have emerged from human ingenuity. For example, the 1991 version of the Collins Dictionary does not list the word ‘Internet’ and at that time ‘pods’ were mainly associated with vegetables such as peas and a ‘mouse’ was a small long-tailed rodent, not something you would have or want on your desk! These are just a handful of innumerable examples of how the English language has changed in recent decades.

If the English language is in a constant state of flux it might reasonably be expected that the systems devised to teach it would mirror the dynamism of their target medium and would be responsive to its mutability. However, this seems not to be the case. Instead the ELT industry appears to be mired in the past and out of synchrony with the realities of 21st century language learners.

**Autoethnography of a ‘native’ speaker English language teacher**

Having briefly outlined the existing framework in which we operate as English language teachers, I would now like to invite the reader to accompany me as I re-visit my own journey as a ‘native’ or first language (L1) English speaker teacher in Brunei Darussalam where I have lived and worked for the past thirteen years.
The Bruneian educational system is an additive bilingual one known locally as *Dwibahasa* (Public Relations Unit, MoE, 2003) featuring Standard Malay (*Bahasa Melayu*) and Standard British English. However, the de facto first language of the majority of Bruneians is Brunei Malay (Martin, 1996), a nativised form of Standard Malay. Although Standard Malay and Brunei Malay are estimated to be about 84 percent cognate (Nothofer, 1991), this still means that students are not being instructed in their mother tongue even when Standard Malay is the medium used. English is also widely used as a medium of instruction across the curriculum, making schools in Brunei more linguistically challenging than is the norm elsewhere.

English language teachers from abroad need to be alerted to the rich complexity of the language ecology of the host country in order to fully appreciate and empathise with their students’ linguistic challenges. It was some years before I came to understand the important distinction between Standard Malay and the Bruneian variety. Initially I was simply informed that for the majority of my students their first language was ‘Malay’.

Having begun my teaching career as a primary level teacher in Ireland, I had had experience of a system in which bilingual education featured. However, in the Irish context, Gaelic is an important subject in the majority of schools rather than a medium of instruction. Although I could draw some parallels, I soon realised that the Bruneian system was far more demanding in terms of linguistic skills since a lack of linguistic acuity could mean not only failure in the target language itself but serious curtailment in terms of access to knowledge in other curricular areas delivered through that language.

Despite these musings and observations, I have to confess that in the first three or four years I devoted myself to delivering the prescribed curriculum, vigilantly patrolling my students’ use of the target language (English) and cajoling them to speak *only* in English during class time. I was acting in their best interests, wasn’t I? After all, this was the received wisdom which was assiduously promoted at language teaching courses such as the Royal Society of Arts (RSA) TEFL Certificate which I had taken in the UK back in 1991 prior to embarking on my sojourn to faraway shores. This doctrine was in turn reinforced by the company for which I worked, a then British charitable organisation employed by the Ministry of Education in Brunei to supply native-speaker teachers of English to supplement the local teaching body for this subject. In those early years I am now ashamed to say that I do not recall learning any new lexical items in Malay (either Standard Malay or Brunei Malay).
Instead I relied on what I had learned from my first visit to Brunei in 1992 and from the two years I had spent in Kuala Lumpur, Malaysia in the mid-nineties. After all, I was the English teacher; I did not need to speak any Malay.

Back in those early years in Brunei I was blissfully unaware that I was in fact perpetuating the fallacies (Phillipson, 1992) by upholding the tenets (Gatenby, 1965) that governed ELT. My daily classes found me intent on finding ways to help my often silent, passive students to become more communicative in English. In retrospect I realise that my focus was often on their English language learning rather than on them as holistic beings. I was so preoccupied with ‘what’ I was teaching that I neglected the ‘who’ involved. By making my students more central to the teaching-learning process, I have since been able to connect with them in more meaningful ways because I ‘know’ them better as individuals. I can use my knowledge of their current interests and concerns to create meaningful learning experiences that might engage and inspire them. Some notable examples from recent classes include finding out that some are passionate about Art; one loves Physics but struggles with it; one is a budding entrepreneur and a master of excuses; while another is phobic about lizards. These may seem to be trivial, inconsequential data at face value, but they have enabled me to ‘reach’ students at a deeper level. Topic choices are determined by students’ interests and needs rather than being dictated solely by texts or syllabi. Our interpersonal relationships are also enhanced by my increased level of interest in them and their personal lives.

My own Irish colonial heritage has rendered me quite vehemently anti-colonial. Therefore, when I was first alerted to the fact that I could be considered to be complicit in neocolonialism (Phillipson, 1992, citing Altbach, 1982) through my active promotion of the English language I was shocked and horrified. I set out at once to exorcise whatever elements of a “colonizer within” (Pennycook, 1998, p.28; Pegrum, 2005, p.2) might exist. My initial reaction to the whole suggestion was one of indignant denial. However, following closer honest interrogation of my own attitudes and behaviours, some could indeed be construed as imperialistic, though inadvertently so. After all, I had accepted the ‘monolingual fallacy’ as ‘best practice’ and in doing so had tried to deny my students the use of their L1 in the classroom. By discouraging the use of the Malay language, I was also giving credence to the fifth fallacy which holds that the use of other languages will adversely affect and ‘interfere’ with the learning of English. For example, it is usual practice for students in my English O level repeater classes to use Brunei Malay, their
actual L1, when they are working in groups. This acts as their default code, especially in the initial stages of discussion when they are formulating and sharing ideas. Later they collaborate to present these thoughts in English in readiness for the feedback stage of the lesson. Previously I would have policed or curtailed L1 use in that initial phase and in doing so probably limited the depth of discussion through my insistence on English only. The more ‘enlightened’ me now interjects only to enquire about the issues under discussion which are then translated for my benefit. Naturally, I remain aware that the use of the L1 must always be judicious and that students’ use of the target language, in this case English, must be nurtured and optimised.

Furthermore, I questioned my very physical presence in Brunei, an ex-British protectorate, a country that falls into Kachru’s ‘Outer Circle’ (Kachru, 1986) and Canagarajah’s ‘periphery’ (Canagarajah, 1999, p. 4) which was upholding the second tenet that the ‘ideal’ teacher of English is a native speaker of that language. I was guilty on all three counts. I had also accepted that it is best if learners begin learning from a young age and that maximum exposure to the language would result in greater proficiency. All in all, I had been acting as a standard bearer for the five tenets of ELT as we know it.

My awakening came as a result of my pursuit of an MA course. I chose to base my dissertation on the ways in which ELT and English Language Learning (ELL) impact the lives of a group of contemporary Bruneian 17-19 year olds in a sixth form college with special attention given to an examination used as the qualifying standard for entry to tertiary education. I had long held misgivings about the appropriacy of this examination in the Brunei context, given its antiquity and elitist thrust.

I now discovered the said examination also functioned as a “crucial gatekeeper” (Pennycook, 1994, p.13): an active agent in condemning more than three quarters of young Bruneians to the academic scrapheap. The fact that these students have actually achieved significant proficiency levels in English, their second (or some may argue third) language is largely ignored in light of unfavourable judgements handed down by a faraway examination board. This is evidence that power continues to be wielded by proficiency in English as determined by the British standard setters, confirming Phillipson’s view that it is “one of the most durable legacies” (Phillipson, 1992, p.111) of colonial rule. Again I found myself complicit in perpetuating a form of neocolonialism in keeping with Sir Winston Churchill’s
1943 prophecy that “the empires of the future are the empires of the mind” (Pennycook, 1994, pp. 130-131).

As I continued to read I became far more critically aware of my professional role. I began to reappraise the educational system itself whilst all the while reflecting on my own actions and interactions in terms of my day-to-day work as I sought to equip my students to achieve the elusive credit grade in the qualifying examination.

I was reassured to find that the bilingual system adopted in Brunei is defined as “additive” in that a learner adds “a new language to his repertory of skills” (Cummins, 2001, p. 39, citing Lambert, 1975). In such contexts supplantation of the first language is not regarded as an issue and both languages enjoy a high level of prestige. In addition, both are seen as “majority languages” and are used as media of instruction (Baker, 1996, pp.193-194). All of these conditions seem to fit the situation in Brunei. However, danger lurks in the curtailment of the L1s (Standard Malay and Brunei Malay) to historical and ethnocultural subjects which risks these languages being regarded as less relevant to today’s world. Furthermore, having discovered the “threshold hypothesis” proposed by Cummins and Swain (1986, p.18), I began to apply it to my teaching context. This hypothesis suggests that students in a bilingual context need to reach certain levels of language competence in their L1 in order to avoid cognitive disadvantages. This did not seem to be happening with the confinement of Standard Malay to a limited number of subjects and the situation was further exacerbated by the fact that the use of Standard Malay as a medium of instruction itself constituted a second language. Cummins (2000; 2001) advocates the ongoing development of both languages and stresses the interdependence of literacy skills across languages. These theories provided a whole new lens through which to view my teaching context and alerted me to the fact that at times I might be required to teach both concept and language. It was also becoming increasingly clear that my students’ L1(s) should not be ignored.

I felt challenged by Skutnabb-Kangas’ (1995) contention that a monolingual teacher of a second language (L2) can never be a really good L2 teacher due to the lack of relativity that accompanies his/her (presumed) monoculturalism. I rejected that idea and set out to prove to myself that this was not the case. Now the tables were turned. I found myself envying my local colleagues who had the facility to communicate with their students in both varieties of Malay as well as in English. I took some consolation in the fact that I was no longer upholding the second tenet of
ELT that the ‘ideal’ language teacher was one who spoke that language natively. Far from feeling ‘ideal’ I now felt woefully inadequate. Holliday (2005) regards my thinking at this juncture as another side of the native speakerism debate. It was also at this stage that I began to see clearly for the first time how unjustified the ascendancy of monolingual ‘native’ English Language Teachers has been (O’Hara-Davies, 2006).

Salvation came soon after this when I discovered the following excerpt:

“…bilingual / bicultural teachers (this includes native-speaker teachers of English who are au fait with the language and culture of the country they are in) are in a position to enrich the process of learning by using the mother tongue as a resource, and by using the culture which the mother tongue embodies, they can facilitate the progress of their students towards the other tongue, the other culture.”

(Prodromou, 2001, ‘Conclusion’ section, para.2)

These words encapsulate who/what I wanted to be as an English teacher and discovering them was a ‘eureka’ moment for me. Not only had I found a prescription for my future professional self, I also had an answer to Skutnabb-Kangas’ challenge.

Without delay I set out to become that kind of English language teacher. Determined to exorcise whatever elements of a coloniser lurked within me or my practices, I sought at once to change my mindset. This began with my language-related thought processes. I began to think of both varieties of Malay as resources (Saxena, 2008) rather than as ‘problems’. Using both Brunei Malay and Standard Malay as resources has been invaluable in helping me ensure that lessons are more comprehensible for students. I have found this particularly in the case of concepts which cannot be explained so readily through the use of images. For example, a comprehension text featured in a past O level English Language examination had ‘temptation’ and ‘jealously’ as underlying themes. Reference to Malay equivalents (godaan and kecemburuun respectively) in a case such as this provides a quick and reliable checks-and-balances system as well as anchoring the new vocabulary items securely, ready for future use. It also casts the students in the role of language expert, thereby helping them to value their bilingualism. Students’
L1(s) can also be used to clarify confusion in the target language, especially in the context of related concepts. For example, students in Brunei tend to use ‘avoid’ and ‘prevent’ interchangeably. By referring to the Malay translation, the confusion can be explained and the distinction between them can be emphasised. Use of Malay and English in this way acknowledges the diversity of cultures within the classroom and mirrors the students’ journeys back and forth within these cultures (Kramsch, 1993). It also reinforces the notion that students’ L1 is valued and relevant regardless of the current hegemony of English.

Pursuant to my new perception of L1(s) as a resource, I began to substitute the notion of “transfer” (Svalberg, 1998, p. 341) for that of ‘interference’ in terms of L1-inspired ‘errors’ in English. Such positivity was in itself both liberating and motivational. In addition, I began to experiment with the judicious use of Malay (both Standard and Brunei) in my classes. For example, a lesson centred around idioms or aphorisms would now involve exchange of target English forms for Malay versions rather than being a one-way street as would erstwhile have been the case. This change began back in 2005 and I am happy to say I have never and will never look back. Use of and reference to Malay has now become an integral part of my methodology. I am enjoying a level of connectedness and rapport with my students that simply did not exist previously. English only? – inda kali eh! (not likely!) My paradigm has changed!

References


Students’ Reactions to School Based Oral Assessment: 
Bridging the Gap in Malaysia

Gurnam Kaur Sidhu  
Chan Yuen Fook  
Faculty of Education, Universiti Teknologi MARA, Malaysia  
yuenfookchan@yahoo.com

Sarjit Kaur Sidhu  
School of Humanities, Universiti Sains Malaysia, Penang, Malaysia

Bio Data
Gurnam Kaur Sidhu (PhD) is Professor at the Faculty of Education, University Technology MARA, 40200 Shah Alam, Selangor, MALAYSIA. Prior to this, she was attached to the National Institute of Educational Management and Leadership and at the Ministry of Education, Malaysia. With more than 30 years of experience in the field of education as a lecturer, practitioner and researcher, she is widely published including being the author of English Language Textbooks currently used in Malaysian schools. Her interests include TESL, Learner Autonomy, Educational Management and Life-Long Learning.

Chan Yuen Fook (PhD) is Professor and Research Coordinator of ACRULET (Asian Centre for Research on University Learning and Teaching) at the Faculty of Education, University Technology MARA, 40200 Shah Alam, Selangor, MALAYSIA. Prior to this he was attached to the National Institute of Educational Management and Leadership, Ministry of Education, Malaysia. Dr. Chan has more than 20 years of experience in teaching and he has presented numerous papers at both national and international conferences. His written contributions have appeared in a number of journals and proceedings. His interests are in the field of Testing and Evaluation, Educational Administration, Management, ICT and Curriculum and Instruction.

Sarjit Kaur Sidhu (PhD) is Associate Professor in Applied Linguistics and Programme Chairperson of the English Language Studies Section at the School of Humanities in Universiti Sains Malaysia. Widely published, her areas of specialisation include Teaching English as a Second Language (TESL), workplace literacies, postgraduate education, learner autonomy and policy research in higher education. Her recent co-edited books include Quality Assurance and University Rankings in Higher Education in the Asia Pacific: Challenges for Universities and Nations (USM Press and IPPTN, 2010), Globalisation and Internationalisation of Higher Education in Malaysia (USM Press and IPPTN, 2008) and a monograph on Contemporary Issues in the Global Higher Education Marketplace: Prospects and Challenges (IPPTN, 2010).
Abstract
In 2002, the Malaysian School Based Oral English Test (SBOET) was implemented and this was viewed as an initial step towards formative language assessment in the ESL classroom. Since then, it has undergone several transformative changes. To date, there is scant empirical research that has looked at the SBOET from the perspective of the test takers. This paper presents the findings of a study that examined feedback from 2,684 upper secondary ESL students from 45 schools located in 10 states in Malaysia. This descriptive study employed both qualitative and quantitative methods of analysis. The data collection process involved the use of questionnaires and focus group interviews. The findings of the study revealed that students were divided as to their opinions on the SBOET. While 55% of the students held positive opinions of the SBOET, the remaining 45% felt otherwise. Urban and rural school students were also divided in their knowledge and understanding of the SBOET. In addition, this study also explored issues regarding students’ readiness, confidence and commitment to formative oral assessment. This study has significant implications for the implementation of formative assessment in the Malaysian ESL classroom.

Keywords: English as a Second Language (ESL), School-based Oral English Test (SBOET), Upper Secondary Schools, Malaysia

Introduction
A preliminary investigation into language assessment indicates that stakeholders such as policy makers, school administrators, teachers, parents and students view assessment differently (Teasdale and Leung, 2000). For instance, policy makers perceive assessment as standards to monitor the quality of education; administrators use assessment to monitor the strengths and weakness of a programme, whilst teachers view assessment as a means to monitor a student’s progress and performance. Likewise, parents view assessment as a form of feedback on their child’s performance and as an indicator of the school’s accountability for effective learning and teaching. Students regard assessment as an indication of their ongoing performance and progress. In other words, whatever the reason for assessment, it must be stressed that high quality assessment should be put in place to facilitate high-quality learning and teaching (Airasian, 2005; Linn and Miller, 2005).

Recent trends in testing and evaluation have witnessed a global trend towards the decentralization of examinations leading a swing towards school based formative and authentic assessments (Van der Watering et al., 2008). White and Gunstone (1992) note that centralized summative assessment procedures are a limited range of tests that promote limited forms of understanding, whilst formative school based assessments are multi-dimensional with an emphasis on students’ involvement in active learning. Nitko (1995) further emphasizes that good instruction requires constant formative assessment.
that can provide constant feedback on students’ performance. This can only be obtained through alternative assessment procedures that can be carried out by teachers through school based assessment.

The need for more effective and widespread school based formative assessment has been emphasized by researchers all around the globe. At the Second International Conference of Educators in Malaysia, there was a call to reduce summative assessments as they are not a true reflection of a student’s progress and achievement (Caldwell, 2007). It was stressed that examinations tend to shrink the thinking process and this halts creativity. Hence, there is a need to prepare students to think outside the box.

Vella, Berardinelli and Burrow (1998) emphasize the fact that assessment procedures used must be able to determine if learners have developed knowledge skills and attitudes as a result of the teaching and learning process. Consequently, students must be given ample time and learning opportunities to display their learning abilities such as the development of critical and creative thinking skills and their ability to solve problems and make decisions. Such features are encapsulated by formative assessment and have been incorporated into the school based oral English test (SBOET hereafter) currently implemented in Malaysian public schools. In summary, this study on students’ reactions to SBOET is located within the larger theoretical framework of a global trend of decentralization in assessment.

Literature Review

In recent years, assessment policies and practices in most disciplines including language teaching programmes worldwide have been changing in a number of ways, both at system and classroom level with many teachers being concerned “with the relationship between assessment and learning” (Brindley, 2001, p. 127). In line with such a shift, teachers view assessment as an activity which is integrated into the curriculum with the aim of improving learning, rather than a ‘one-off summative event’ (Brindley, 2001).

With English fast developing as a global language of choice for communication, business, commerce and academia (Crystal 1997), there is a need to ensure that students develop good speaking skills so that they can operate in today’s globalised economies. As such, in second language learning contexts, the ability to develop linguistic proficiency with respect to oral language skills is certainly a key emphasized learning objective in Malaysian ESL classrooms. The objective of teaching spoken language is
the development of the ability to interact successfully in that language and this involves comprehension as well as production.

Speaking tasks involving “peer-to-peer interaction as well as group based tasks are increasingly being incorporated into language proficiency assessments, in both large-scale international testing contexts, and in small-scale course-related contexts (Ducasse and Brown, 2009, p. 423.)” With such trends becoming predominant, the basic problem of testing spoken ability is the same as for testing writing in that teachers want to set tasks that form a representative sample of the population of tasks that they expect their students to be able to perform. Weir (1993) proposes a three-part framework comprising the operations that are involved in spoken interaction (activities or skills), the conditions under which the tasks are performed and the quality of the output (the expected level of performance in terms of various relevant criteria). In most ESL classroom contexts, when teachers need to test whether learners can speak effectively, they need to make students take part in direct spoken language activities in order to assess whether their learners can perform relevant tasks and adapt their speech to the circumstances, making decisions under time pressure, implementing them fluently and making any necessary adjustments as unexpected problems arise. In such a situation, classroom based formative assessment provides the relevant platform needed for more authentic testing in the speaking classroom.

School based oral English tests are becoming the preferred method of assessment in many foreign and second language contexts since the inception of oral proficiency testing. It is not surprising to see such change in focus given the fact that many developing countries have introduced literacy benchmarks and assessment regimes in their school systems which cater predominantly to ESL or EFL students. However, as a result of progressive changes in the view of what ‘speaking’ in tests should consist of, since the late 1980s, “pair or group tasks involving peer-to-peer interaction have increasingly been used” (Ducasse and Brown, 2009, p. 424).

This shift in focus from summative to more formative and authentic assessment in the language classroom is linked to the claims of positive washback (impact that tests have on teaching and learning) on the classroom as well as the belief among policy makers and teachers that such tasks are thought to be more representative of best practice in classroom activities. Indeed, the major issue of washback in the field of assessment has been a concern with the systemic validity of tests or the effect a test has on classroom practice (Bachman and Palmer, 1996). For teachers, the pragmatic perspective provided by peer-to-peer assessment is certainly more time and cost
effective, seeing that candidates can be tested together and raters assess two or more candidates simultaneously (Ducasse and Brown, 2009). Such motivation in testing spoken language of candidates has also promoted the acknowledged value of authentic tasks compared to interview tests that result in mainly ‘test discourse’ and did not represent normal conversation advocated by the communicative approach to language teaching. Testing speaking in a second language invariably involves assessing information transfer whereby students can engage in negotiating meaning in pair or group-based tasks. Such ‘interactional competence’ allow raters to assess how speakers structure and sequence their speech and how they apply turn-taking rules. It can also be defined in terms of how speakers co-construct the spoken performance (Jacoby and Ochs, 1995).

In assessing oral language skills, many teachers incorporate aspects of formative learning in their ESL classrooms, but it is less common to find it being practiced systematically. By implementing formative learning, teachers can have a better feel of how to identify and react to their students’ needs. In classrooms featuring formative assessments, teachers make frequent interactive assessments of students’ understanding and this enables them to adjust their teaching to suit individual students. If formative learning is used as a guiding framework in learning contexts, teachers can change the way they interact with their students, how they set up learning goals, guide students towards learning and how they define student success. The benefits of formative assessment have been shown to be highly effective in raising the level of student attainment, increasing equity of student outcomes and improving students’ abilities to learn.

Recognizing the many benefits of formative school based assessment, many educational institutions around the world, including Malaysia, have witnessed a paradigm shift towards decentralizing assessment. Such an opinion was first articulated by the former Malaysian Minister of Education, Tan Sri Musa Mohamed (TheStraitsTimes INTERACTIVE, 2003) when he said “we need a fresh and new philosophy in our approach to exams . . . we want to make the education system less exam-oriented and (we) are looking at increasing school based assessment as it would be a better gauge of students' abilities.”

Later in 2007, it was reported that a proposal was underway to decentralise assessment in Malaysian public schools (Simrit Kaur, 2007). According to Simrit Kaur (2007), the Malaysian Examination Syndicate (MES) proposed the following forms of assessments for Malaysian public schools: school assessment, central assessment,
central examination, psychometric tests and physical activity assessment. It was highlighted that school based assessment would be planned, developed, examined and reported by teachers in schools. At the same time, the Director of MES hinted that there was a plan to abolish two major examinations, i.e. the Primary Year Six UPSR (Primary School Assessment) centralised examination by 2012 and the Secondary Three PMR (Lower Secondary School Assessment) centralised examination by 2015. He stressed that his ministry would provide the necessary ‘tools and training’ to empower teachers to conduct school-based assessment effectively in their schools (Simrit Kaur, 2007).

**Background to the Study**
In Asia, the Philippines, Hong Kong, and Singapore have been officially immersed in English since World War II. The pursuit of higher academic status coupled with demands of keeping pace with the dynamic pace of globalization has expanded English education in Taiwan. Many universities now propose the use of English to teach in specific disciplines. Malaysia and Indonesia are now mandating English competence in government and business (Cheng, 2010, p. 25). Keeping in line with global trends of testing and evaluation, both policymakers and educators in Malaysia have taken steps to introduce school based assessment (SBA) as a catalyst for educational reform. Therefore the School Based Oral Assessment (SBOA), which was introduced in all secondary public schools in 2002, was viewed as a form of leverage for instructional improvement to assist teachers to find out what students are learning in the classroom and how well they are learning it.

The School Based Oral English Assessment (SBOEA) or *Ujian Lisan Berasaskan Sekolah* (ULBS) is a form of school based assessment that is formative and on-going and is implemented based on the guidelines provided by the MES in the ULBS Manual. It is based on the rationale that it is an authentic assessment in comparison to the conventional Oral English Test (OET) that was administered at the end of each academic year by a centralized body requiring students to take a 10-minute oral test. It was also highlighted that the one score meted out to students from centralised summative examinations was not an indicator or true reflection of a student’s actual communicative competency.

Therefore, the SBOEA was introduced with an aim to help develop students’ English language ability in accordance with the learning objectives outlined in the

According to the ULBS Manual, the SBOEA is a more valid form of assessment as it is in line with and based on the latest English Language syllabus requirements laid out by the Curriculum Development Centre (Kementerian Pendidikan Malaysia, 2003a and 2003b). It is also said to be more reliable as the chief evaluator is the classroom teacher; an examiner who is well-acquainted with the students, their needs and their learning styles and hence can truly assess a student’s actual language competency. The manual also emphasizes that the SBOEA is more authentic and versatile as the communicative tasks are integrated into everyday teaching and learning. This not only allows teachers to assess students’ oral communication skills in a variety of situations but also allows teachers to assess a range of speaking skills.

Since the implementation of the SBOEA in 2002, it has undergone several changes and the opinions of various parties, especially teachers and testing experts, have been taken into consideration to further fine tune and address issues and concerns surrounding the implementation of the SBOEA. Nevertheless, after more than six years (this nationwide study was conducted in 2008) into the implementation, there has been scant empirical research carried out to investigate the learners’ perspectives of the implementation of the SBOEA. Therefore, this study aims to explore students’ perspectives of this formation oral assessment – specifically the upper secondary Malaysian students sitting for the School Based Oral English Test.

Under the SBOET, the upper secondary students are required to sit for three oral assessments. The assessments are conducted by their classroom TESL teacher during their regular English language periods. The assessments are conducted during a stipulated time frame. The first assessment in Form Four is carried out during the months of April to June, while the second assessment is from July to September. In Form Four, students are assessed on Model 1 (Individual Presentation) and Model 2 (Individual presentation with prompts). In Form 5, students sit for one oral assessment and are given the option to choose between Model 3 (Pair Work) or Model 4 (Group work). Students are also given the option to choose their own activities from a given list of topics based on the curriculum specifications for Malaysian Secondary Schools. In the SBOET, teachers are encouraged to assist students in their preparation in aspects such as choosing the activity, the topic and deciding and preparing the task. During the assessment, candidates are allowed to refer to the text. However, reading directly from
the texts is not allowed. Candidates are also encouraged to be creative in carrying out the task.

Under the SBOET, students will be informed of their marks after each assessment and if they are not satisfied with their grades they can repeat their presentation. Once they are satisfied with their performance, they acknowledge their grade by signing the Oral English Test Individual Profile Score Sheets (Borang Profil Individu). After three assessments, the assessor (teacher) then selects the highest score among the three assessments and completes the Master Score Sheet (Borang Markah Induk). The total marks from the three assessments are added and the student’s oral assessment is graded based on the criteria given.

To ensure that each new paradigm, policy or programme is beneficial to students, teachers and all stakeholders, systematic and effective evaluation must be carried out. Feedback received can be used to provide valuable information on the strengths and limitations of its implementation so that the necessary steps and measures can be taken to enhance the implementation. Thus, this study explored the perspectives of the students with regards to the SBOET. More importantly, it would impress upon the relevant authorities to take the necessary steps and measures to further enhance its implementation as school based assessment is here to stay and should not be seen as another swing of the pendulum.

**Purpose of the Study**

The main aim of this study was to investigate students’ perspectives on the formative assessment of the School Based Oral English Test (SBOET). Hence, the study sought to answer the following research questions:

1. What are the students’ perceptions of the formative School Based Oral English Test with regards to the following:
   a. opinions of the SBOET
   b. dissemination of information on SBOET
   c. implementation of SBOET,
   d. feedback on SBOET
   e. SBOET models of assessment
   f. Scoring criteria of SBOET

2. To what extent are students prepared for the School Based Oral English Test?
3. What challenges do students face as test takers of the School Based Oral English Test?

Methodology
The study employed a descriptive design which allowed a qualitative and quantitative description of the relevant features of the data to be collected based on different variables outlined such as gender, year of study, main language spoken at home, academic stream, type of school, tuition, and the PMR (Lower Secondary) grade for English language.

Population Sample
The population sample for this study involved a total of 45 schools from the 10 states which were divided into the following five zones: North Zone (Kedah and Perak), East Zone (Terengganu and Kelantan), Central Zone (Selangor and Federal Territory of Kuala Lumpur), South Zone (Johor and Melaka), and East Malaysia (Sarawak and Federal Territory of Labuan). The study only involved Secondary Four and Secondary Five students who were involved in the SBOET.

Instrumentation
Data for the study were collected via a Student Questionnaire and focus group interviews. Both these instruments were used to examine students’ opinions, perceptions of and readiness for the SBOET. Focus group interviews involved a total of 50 students (10 from each zone). The ten students in each focus group interview comprised five from Secondary Four and five from Secondary Five ESL classrooms.

The Student Questionnaire consisted of four parts. Part I sought the respondents’ demographic data while Parts 2 and 3 examined students’ perceptions of and readiness for the SBOET. In Part 2 of the questionnaire, students responded to items using a dichotomous scale of ‘Yes’ and ‘No’, but for Part 3 students responded using a 4-point Likert scale. A response of 1 indicated that the students ‘Strongly Disagreed’ to the statement whilst a response of 4 showed that the respondent ‘Strongly Agreed’ to the statement. Finally, Part 4 consisted of six open-ended questions that looked into the students’ perceptions of the SBOET. The internal consistency of the items in the Student Questionnaire was validated with the Cronbach Alpha Reliability Test. The result of the reliability of the students’ questionnaire was at the acceptable level (r = 0.75).
Data Analysis
Simplistic statistics such as percentage, frequency, means, standard deviation, significant differences, and comparing several means were used to analyse quantitative data in this study. In terms of the qualitative data gathered from focus group interviews, the data were coded, categorized and then elaborated based on the feedback given. The data obtained from both questionnaires and interview sessions were constantly compared and contrasted before any conclusions were drawn.

Results and Discussion
A total of 2684 students from a total of 45 schools responded to the questionnaire. The demographic data revealed that the majority of the respondents were females (60.3%) while the remaining were males (39.7%). The results further indicated that 63.7% of the students were in Secondary Five (5) whereas 36.3% were Secondary Four (4) students. With regards to ethnicity, the majority of the respondents were Malays (71.2%), followed by Chinese (17.7%), Indians (6.4%), Iban/Kadazan (1.3%) while others constituted 3.4%. In this study, 82.7% of the respondents were from day-time non-residential schools, 11.7% were from residential schools while the remaining 5.6% were from state run schools. Slightly more than half (51.2%) of the students were from the science stream, a third (35.8%) were from the arts streams while the remaining 9.6% and 3.4% were from the vocational/technical and other academic streams respectively.

Students’ Perceptions of the School Based Oral English Test (SBOET)
The questionnaire sought students’ perspectives of the SBOET from perspectives such as their opinions of the implementation SBOET, the dissemination of information and their perception of the scoring and feedback of SBOET.

Opinions on SBOET
Students were rather divided as to their opinions on the SBOET. Approximately 55.0% of the students held positive opinions of the SBOET. Qualitative data obtained from interview sessions further revealed that students who viewed it positively see the SBOET as a chance to improve their oral skills as it allowed them to repeat their presentations. They also indicated that they could collaborate and they learnt much from their peer presentations. In addition, students reported that the SBOET helped them in building their confidence as their teachers kept reminding them that they could
obtain ‘extra’ or ‘bonus’ marks which might help them obtain a better grade for the English Language SPM paper (equivalent to the UK ‘O’ Levels). The majority of the respondents viewed the SBOET as a ‘necessary’ item as they claimed that it is an important indicator of their future and this group felt it has been ‘fun’ taking the SBOET. In addition, by observing the presentations of their peers they could get better ideas and a chance to emulate and model their presentations after good examples were highlighted by their teachers.

The remaining 45.0% of the students who viewed the SBOET negatively gave several reasons behind their negative responses. During the focus group interviews, a few students highlighted that classroom presentations under the SBOET were often marred by teasing peers which at times demoralized them. According to a student from the state of Kelantan, he did not like the SBOET because of his strong native “Kelantan” accent. He said, “my friends laugh at me when I speak English using my Kelantan accent.” Another student from Selangor highlighted she felt uncomfortable speaking English because “my (her) friends are all very good and I am not good and they make funny face at me when I speak the English.”

Qualitative data from interview sessions also revealed that students’ presentations were often interrupted by noisy, disruptive and inattentive students. Others claimed that the SBOET was a futile effort as they could not hear the presentations of their peers as the classroom was often noisy or there was a lot of noise from neighbouring classes. This often left student presentations barely audible. Consequently, students were disinterested in the oral presentations of their peers. Some respondents felt that the SBOET was a ‘burden’ as it requires a lot of preparation time and was a ‘waste of time’ because it takes them away from their study time in their classrooms. In addition, they also have to juggle the preparation of the oral test and their daily assigned homework. Finally, there were a few students (15.0%) who felt that the SBOEA is not an indicator of their overall language proficiency as it is not testing their abilities in speaking English.

A majority (90.3%) of the students also agreed that the oral assessment is a compulsory component of their English language assessment. Another 81.6% claimed that they took the assessment seriously. Further analysis of the findings revealed that there is a significant difference (F=11.276, p=< 0.05) among students of the Average English Proficiency as compared to the High and Limited English Proficiency students in terms of the compulsory nature of the test.
The findings displayed in Table 1 reveal that most of the students also agreed (mean= 3.11) that school based assessment had helped improve their speaking skills. They also added that the classroom environment in which the SBOET was conducted was suitable and conducive (mean 2.67). Nevertheless, they felt that they would prefer only one assessment in each year compared to the current two assessments in Form Four and one in Form Five. The students also felt that they would prefer to be graded by another English teacher (mean =2.04) rather than their own TESL teacher.

Table 1: Students’ Perceptions of the SBOET (n= 2684)

<table>
<thead>
<tr>
<th>Students’ perceptions of the SBOET</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>The school based oral assessment has helped me improve my speaking skills</td>
<td>3.11</td>
<td>.72</td>
</tr>
<tr>
<td>There should be only one oral assessment for Form Four and one for Form Five</td>
<td>2.81</td>
<td>1.01</td>
</tr>
<tr>
<td>My classroom environment is suitable for my oral assessment</td>
<td>2.67</td>
<td>.89</td>
</tr>
<tr>
<td>Oral assessment should be graded by another English language teacher</td>
<td>2.04</td>
<td>.95</td>
</tr>
</tbody>
</table>

*Scale: 1=Strongly Disagree  2=Disagree  3=Agree     4=Strongly Agree*

**Dissemination of information on the SBOET**

With regards to the dissemination of information, students were asked if their teachers briefed them on the oral English assessment. The findings presented in Table 2 show that a large majority (90.5%) of the respondents were briefed by the teachers but it is important to note that close to 10.0% of the respondents were not informed. In addition, 93.9% reported that their teachers did inform them of the oral English activity. Interview sessions further highlighted that students were informed verbally by their respective TESL teacher of the oral assessment. Nevertheless, a majority of students indicated that their schools did not issue a written circular to their parents.

Table 2: Dissemination of Information on the SBOET (n= 2684)

<table>
<thead>
<tr>
<th>Dissemination of information on the SBOET</th>
<th>Frequency (n=2684)</th>
<th>Yes Percent (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>My teacher briefs me on the Oral English Assessment</td>
<td>2430</td>
<td>90.5</td>
</tr>
<tr>
<td>My school informs my parents of the Oral English Assessment</td>
<td>1292</td>
<td>48.1</td>
</tr>
<tr>
<td>The school and class Oral English Assessment schedule is put up on the board</td>
<td>1049</td>
<td>39.1</td>
</tr>
<tr>
<td>My teacher informs me of the oral English activity</td>
<td>2521</td>
<td>93.9</td>
</tr>
</tbody>
</table>

*Scale: 1=Yes, 0=No*
Effective implementation of the SBOET also requires school authorities to inform both students and parents regarding the SBOET. By doing so, parents can help monitor and assist their children to do well in their oral test. In this study, close to two thirds (60.9%) of the students highlighted that both the classroom teacher and the school authorities did not put the SBOET schedule on the notice boards. Furthermore, only 48.1% of the students’ parents were informed of the SBOET. This indicates that there is poor dissemination of information to both students and their parents. Additionally, it also indicates the ineffective implementation of the SBOET.

Table 3: ANOVA of Informing Parents of the SBOET (n= 2684)

<table>
<thead>
<tr>
<th>Item</th>
<th>Category</th>
<th>Item</th>
<th>Yes</th>
<th>N</th>
<th>M</th>
<th>SD</th>
<th>F</th>
<th>Sig</th>
</tr>
</thead>
<tbody>
<tr>
<td>My school informs my parents of the oral English assessment</td>
<td>Proficiency</td>
<td>HEP</td>
<td>42.1</td>
<td>880</td>
<td>1.58</td>
<td>0.49</td>
<td>14.11</td>
<td>.00</td>
</tr>
<tr>
<td></td>
<td></td>
<td>AEP</td>
<td>50.1</td>
<td>1085</td>
<td>1.50</td>
<td>0.50</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>LEP</td>
<td>55.4</td>
<td>730</td>
<td>1.45</td>
<td>0.50</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Location</td>
<td>Urban</td>
<td>41.3</td>
<td>1086</td>
<td>1.59</td>
<td>0.49</td>
<td>23.46</td>
<td>.00</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Suburban</td>
<td>53.2</td>
<td>823</td>
<td>1.47</td>
<td>0.50</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Rural</td>
<td>56.2</td>
<td>721</td>
<td>1.44</td>
<td>0.50</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Scale: 1=Yes, 2=No

Further analysis using the one-way ANOVA results (Table 3) indicated that parents of the students in the rural schools (mean=1.44) are better informed about the SBOET in comparison to the parents of students in the urban (mean=1.59) and suburban (mean=1.47) schools (F=23.462, p=.000). The data also revealed that parents of students possessing limited English proficiency (LEP) are better informed about the SBOET compared to parents having High English proficiency (HEP) and Average English Proficiency (AEP) students. Both these differences were significantly different (F = 14.113, p = < 0.05).

The findings (Table 4) in this study indicated that there was a significant difference in the dissemination of information by teachers to students. Students in rural localities were better informed of their oral presentation compared to their counterparts in suburban and urban areas. The findings also revealed that HEP and AEP students were better informed of their oral test than LEP students.
Table 4: ANOVA of Informing Students of the SBOET (n=2684)

<table>
<thead>
<tr>
<th>Item</th>
<th>Category</th>
<th>Item</th>
<th>Yes (%)</th>
<th>N</th>
<th>M</th>
<th>SD</th>
<th>F</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>My teacher informs me of the Oral English activity</td>
<td>Proficiency</td>
<td>HEP</td>
<td>97.5</td>
<td>807</td>
<td>1.02</td>
<td>0.16</td>
<td>18.31</td>
<td>.00</td>
</tr>
<tr>
<td></td>
<td></td>
<td>AEP</td>
<td>96.9</td>
<td>1089</td>
<td>1.03</td>
<td>0.17</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>LEP</td>
<td>91.9</td>
<td>739</td>
<td>1.08</td>
<td>0.27</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Location</td>
<td>Urban</td>
<td>95.9</td>
<td>1092</td>
<td>1.04</td>
<td>0.20</td>
<td>8.53</td>
<td>.00</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Suburban</td>
<td>93.5</td>
<td>820</td>
<td>1.06</td>
<td>0.25</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Rural</td>
<td>97.8</td>
<td>730</td>
<td>1.02</td>
<td>0.15</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Scale: 1=Yes, 2=No

Implementation of the SBOET

With regards to the implementation of best practices, respondents were asked if they were given the freedom to choose from the list of activities. Findings displayed in Table 5 show that a large majority (80.0%) of the respondents agreed that they were able to choose any of the activities from the given list. This is encouraging as it gives students the freedom to talk on a topic of their choice. In addition, a majority (83.8%) of the students also stressed that they were allowed to choose their own partner for group/pair work. This was also articulated by students during the interview sessions. According to a student from Kedah, she liked the SBOET as it “gives me (her) the chance to choose to talk about something I like talking about.” This is line with the notion that students can also be empowered and involved in the democratization process of teaching and learning as echoed by the Ministry of Education.

The school based assessment is formative in nature and teachers and students are encouraged to work together to enhance student learning. Therefore in this study, students were asked if they discussed their topic with their teachers before their oral presentation. The findings revealed that only 59.0% responded favourably. Furthermore slightly more than half (52.5%) of the respondents agreed that their teachers guided and assisted them in preparing their oral presentation scripts.

Table 5: Best Practices in the SBOET (n=2684)

<table>
<thead>
<tr>
<th>Best practices in the SBOET</th>
<th>Frequency (n=2684)</th>
<th>Percent (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>I can choose any activity from a given list.</td>
<td>2148</td>
<td>80.0</td>
</tr>
<tr>
<td>I am allowed to choose my own partner for group/pair work.</td>
<td>2250</td>
<td>83.8</td>
</tr>
<tr>
<td>I discuss my topic with my teacher before oral presentation</td>
<td>1605</td>
<td>59.8</td>
</tr>
</tbody>
</table>
My teacher corrects and edits my script before oral presentation 1410 52.5
My teacher informs me of my marks after the oral presentation 1902 70.9
After my presentation my teacher tells me my weaknesses 1846 68.8
I am happy with the marks that my teacher gives me 1489 55.5

Further analysis using ANOVA exhibited that there was a significant difference between proficiency (F=14.69, p=.00) and localities (F=114.23, p=.00) with regards to teacher guidance before an oral presentation (Table 6). The results indicated that LEP students obtained more guidance from their teachers on their oral presentation script than AEP and HEP students. Comparatively, teachers from the rural areas also gave more assistance and guidance to their students on their oral presentation scripts than teachers in urban and suburban areas.

Table 6: ANOVA of Teacher Assistance and Guidance before SBOET (n= 2684)

<table>
<thead>
<tr>
<th>Item</th>
<th>Category</th>
<th>Item</th>
<th>Yes</th>
<th>N</th>
<th>M</th>
<th>SD</th>
<th>F</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>My teacher assists and guides me before oral presentation</td>
<td>Proficiency</td>
<td>HEP</td>
<td>49.2</td>
<td>807</td>
<td>1.51</td>
<td>0.50</td>
<td>14.69</td>
<td>.00</td>
</tr>
<tr>
<td></td>
<td></td>
<td>AEP</td>
<td>51.3</td>
<td>1083</td>
<td>1.49</td>
<td>0.50</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>LEP</td>
<td>61.9</td>
<td>738</td>
<td>1.38</td>
<td>0.49</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Location</td>
<td>Urban</td>
<td>39.2</td>
<td>1091</td>
<td>1.61</td>
<td>0.49</td>
<td>114.23</td>
<td>.00</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Suburban</td>
<td>55.1</td>
<td>820</td>
<td>1.45</td>
<td>0.50</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Rural</td>
<td>73.9</td>
<td>717</td>
<td>1.26</td>
<td>0.44</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Feedback on Students’ Presentations

In the study, respondents were also asked about teacher feedback on their presentations. Only 68.8% of the respondents agreed that their teachers provided feedback and explained their weaknesses. Even though the majority of the respondents responded favourably that they always received feedback, as a best practice, teachers must improve on this aspect as students needed immediate feedback on their weaknesses in the SBOET so as to further improve their presentations in future assessments. Further analysis however indicated (Table 7) that there was a significant difference between proficiency (F=8.97, p=.00) and localities (F=22.67, p=.00) with regards to teacher feedback on the SBOET performance (Table 8). ANOVA analysis showed that HEP
students received better feedback from their teachers on their oral presentation than AEP and LEP students. Besides that, teachers from the rural areas were also more concerned about their students’ weaknesses and provided better feedback on their students’ oral presentation than teachers in urban and suburban areas.

Table 7: ANOVA of Teacher Feedback on SBOET Performance (n= 2684)

<table>
<thead>
<tr>
<th>Item</th>
<th>Category</th>
<th>Yes</th>
<th>M</th>
<th>SD</th>
<th>F</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>After my presentation my teacher tells me my weaknesses</td>
<td>Proficiency</td>
<td>HEP</td>
<td>76.2</td>
<td>1.24</td>
<td>0.43</td>
<td>8.97</td>
</tr>
<tr>
<td></td>
<td>AEP</td>
<td>70.1</td>
<td>1.30</td>
<td>0.46</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>LEP</td>
<td>66.6</td>
<td>1.33</td>
<td>0.47</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Location</td>
<td>Urban</td>
<td>65.7</td>
<td>1.34</td>
<td>0.48</td>
<td>22.67</td>
</tr>
<tr>
<td></td>
<td>Suburban</td>
<td>70.1</td>
<td>1.30</td>
<td>0.46</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Rural</td>
<td>80.2</td>
<td>1.20</td>
<td>0.40</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Again, even though the respondents agreed that they received feedback on their weaknesses, 55.5% of students displayed unhappiness with the marks given by their teachers. The respondents were again split in their responses. This is a worrying trend as it indicates students’ dissatisfaction on the marks given by their teachers. Thus, it also raises several concerns about the credibility of the teachers in carrying out the SBOET, in particular the aspects of reliability and the validity of scoring. This is probably one of main reasons why respondents indicated that they would prefer to be graded by another English teacher (mean =2.04) rather than their own TESL teacher.

Students’ views on the SBOET Models

The open ended questions in the Students’ Questionnaire sought students’ views on the SBOET Models. The findings showed that the majority of the students preferred Model 4 (Group work). Interview sessions further revealed that students felt that they were more confident when presenting in groups. Students also highlighted that group work activity enhanced their abilities to discuss with each other before their presentations, thus providing the opportunities to exchange ideas and gain new information. In addition, the respondents also stressed that through group work, they were able to cooperate and work as a team. The strong preference for group work enhanced their individual talent, motivation and time management in taking the test. The respondents agreed that group work is a much easier task than the other two models as they felt less nervous and had more fun and enjoyment with their friends, hence, assisting them in
overcoming their shyness. As a whole, the respondents also agreed that the group work model helped them to improve their overall speaking, presentation and language skills.

The respondents listed the individual presentation as a second preference. Focus group interviews highlighted that through individual presentations a candidate can easily prepare at his or her own pace, is free to choose what they wanted to do with their presentations, and had the freedom to choose their own topic. Students also articulated that an individual presentation could boost their self-confidence through improving their speaking skills and developing their overall language proficiency. The respondents believed that through an individual presentation, a candidate can gain the attention of the assessor.

The least preferred model was Model 2 (Individual work with prompts). The respondents reported that Model 2 limits their abilities and opportunities to contribute both ideas and the topic of their presentations. Nevertheless some students also voiced discontent over Model 1 as it encouraged them to memorize and did not allow them the flexibility of free conversation.

The students were also asked if listening to the presentation of their peers helped them in any way. A large majority (76.0%) said that they had learned from their classmates’ oral presentations by imitating and modelling after the good examples highlighted by the teacher. They also believed that they could gain new ideas and new experiences from the different examples of oral presentations delivered by their classmates. In addition, they believed that their classmates could give opinions, ideas and more information as they interacted with one another. Furthermore, they agreed that by observing the oral presentations of their peers, they became more motivated and inspired during their own oral presentations.

Apart from that, the respondents felt that they would become more relaxed and confident as they are able to discuss with their classmates who have presented. In addition, they would be able to identify their level of English and specific weaknesses, namely in pronunciation and grammar to improve their speaking skills.

Despite the positive aspects highlighted, there were some indications that they would not have learned from their classmates’ oral presentation. This was clearly evident when they reported that some of their classmates spoke softly, presented uninteresting topics and encountered difficulties in being understood by others.

The findings also revealed that many of the respondents could not devote full attention during the presentations as the classroom was noisy and some of their classmates were busy with their own work. In addition, the respondents reported that at
times, the oral test was carried out on a one to one basis with the teacher. When this happened, they were not given the opportunity to observe and learn from their classmates’ presentations. A majority commented that they should not emulate the weaker students that presented their memorized scripts during the presentation and those who were not able to answer questions spontaneously. Finally, the respondents confirmed that they did not learn anything from their classmates as they felt that the oral presentation was not important and the time allocated for the presentation was only for ten minutes.

**Students’ Readiness for the SBOET**

The questionnaire in this study also sought to gain information on students’ level of readiness for the SBOET. It included items that investigated their level of confidence and preparation for the SBOET.

<table>
<thead>
<tr>
<th>Students’ perceptions of oral presentations</th>
<th>M</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>I am given enough time to deliver my oral presentations.</td>
<td>3.01</td>
<td>.72</td>
</tr>
<tr>
<td>I am given enough time to prepare for my oral presentation</td>
<td>2.98</td>
<td>.78</td>
</tr>
<tr>
<td>I should memorize my script for my oral presentations</td>
<td>2.94</td>
<td>.84</td>
</tr>
<tr>
<td>It is good to read while delivering my oral presentation</td>
<td>2.44</td>
<td>.95</td>
</tr>
<tr>
<td>I should be given a Student’s Kit to prepare better for my Oral Assessment</td>
<td>2.97</td>
<td>.83</td>
</tr>
<tr>
<td>I am confident during my oral presentations</td>
<td>2.56</td>
<td>.73</td>
</tr>
</tbody>
</table>

*Scale:* 1=Strongly Disagree 2=Disagree 3=Agree 4=Strongly Agree

The findings exhibited in Table 8 show that the respondents portrayed a moderate level of agreement to most of the items. Students indicated that they felt they are given sufficient time to prepare (mean=2.98) and deliver their oral presentations (mean=3.01). The results also revealed that students were also aware of the fact that they should not read when presenting their script (mean=2.44). In addition, students indicated that they should be provided with a Student’s Kit (mean 2.97) for the SBOET.
Table 9: ANOVA of the Level of Confidence amongst Students during Oral Presentation

<table>
<thead>
<tr>
<th>Item</th>
<th>Crosstab</th>
<th>Item</th>
<th>N</th>
<th>M</th>
<th>SD</th>
<th>F</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>I am confident during my oral presentations</td>
<td>Proficiency</td>
<td>HEP</td>
<td>821</td>
<td>2.79</td>
<td>.74</td>
<td>11.0</td>
<td>.000</td>
</tr>
<tr>
<td></td>
<td></td>
<td>AEP</td>
<td>1108</td>
<td>2.81</td>
<td>.65</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>LEP</td>
<td>743</td>
<td>2.66</td>
<td>.82</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Location</td>
<td></td>
<td>Urban</td>
<td>1128</td>
<td>2.78</td>
<td>.66</td>
<td>3.70</td>
<td>.25</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Suburban</td>
<td>824</td>
<td>2.79</td>
<td>.75</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Rural</td>
<td>720</td>
<td>2.70</td>
<td>.81</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The ANOVA analysis indicated that there was no significant difference among students of different proficiency levels (F=.81, p=.44) pertaining to the need of a Student’s Kit in preparation for the SBOET, but there was a significant difference in terms of locality (F=.01, p=.00) between urban, suburban and rural students. Students coming from rural (mean=3.0) and urban (mean=3.00) localities expressed a higher need for a Student’s Kit when compared to their peers in the suburban (mean=2.89) schools.

However, students did not exhibit a high level of confidence for the delivery of their presentations (mean=2.56). Further analysis of the findings (Table 9) revealed that slightly more than 73.0% of the HEP and AEP students expressed a higher confidence in comparison to only 64.9% of the LEP students. The ANOVA analysis also revealed that there is a significant difference in the level of confidence between HEP, AEP and LEP students (F=11.0, p=.00). Results indicated that AEP and HEP students demonstrated a higher level of confidence than the LEP students. However, students are equally confident during the oral presentations despite their location (F=3.70, df=2, p=.25).

Table 10: Students’ Readiness for the SBOET (n = 2684)

<table>
<thead>
<tr>
<th>Students’ readiness for the SBOET</th>
<th>Frequency (n=2684)</th>
<th>Percent (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>I am comfortable doing my oral presentation in front of my classmates</td>
<td>1351</td>
<td>50.3</td>
</tr>
<tr>
<td>I prefer doing my oral presentation privately with my teacher</td>
<td>1746</td>
<td>65.1</td>
</tr>
<tr>
<td>I discuss my topic with my teacher before oral presentation</td>
<td>1605</td>
<td>59.8</td>
</tr>
<tr>
<td>For my oral presentations I get help from:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>· my family members</td>
<td>745</td>
<td>27.8</td>
</tr>
<tr>
<td>· my friends</td>
<td>2137</td>
<td>79.6</td>
</tr>
<tr>
<td>· tutor / tuition teacher</td>
<td>569</td>
<td>21.2</td>
</tr>
</tbody>
</table>
For my oral presentations I gather materials from:

- newspapers and magazines: 1424 (53.1)
- Internet: 762 (28.4)
- the library: 1108 (41.3)
- Others (please state): 829 (30.9)

I understand the different aspects on which my oral presentation is graded.
I am aware of the scoring criteria for my oral presentation.

The lack of confidence among the students was further established when almost 49.7% of them admitted that they were not comfortable doing their oral presentation in front of their classmates (Table 10). In addition, more than half of the students (65.1%) indicated that they preferred to do their oral presentation privately with their teachers. This, however, defeats the purpose of the SBOET, where the speaking skill is seen as a productive skill that requires an audience. Hence teachers must ensure that test takers for the SBOET do their presentations in front of their classmates as stated in the SBOET Manual.

When asked if they had discussed their topic with their teachers before their oral presentation, almost 60.0% of the respondents said they had. However, this is still a low indication of teacher support as only 59.8% of the respondents sought help from their teachers. Nevertheless, students agreed that their teachers guided and assisted them in preparing their scripts before the SBOET, as the majority of the respondents (79.6%) received help from their friends, followed by the students’ family members (27.8%) and only 21.2% of the students’ received help from their tutors and the tuition teachers. It is clear that their peers played a very crucial role in supporting and assisting them for their presentations in the SBOET.

Respondents were also asked how they prepared for the SBOET. A majority of the respondents (53.1%) reported that they gathered materials from newspapers and magazines, while 41.3% of them used library materials and the remaining 28.4% of the respondents surfed the internet for information. This clearly indicates that most of the respondents are still comfortable with the references and materials from the traditional sources of information instead of utilizing the internet for more current materials. This is very much similar to another study conducted by Chen and Squires (2010) who investigated vocational students’ perceptions on standardized English proficiency tests.
in Vietnam. Their study also reported that close to 63% of the students referred to traditional sources such as books with only 12% using resources provided by the self-access centre at the university, when preparing for their.

If students were ready for the SBOET, they should have a good understanding of it. Hence, students were asked if they understood how the oral presentation was graded. Findings indicated that 69.7% indicated an awareness of the different aspects to which their oral presentation is graded and a similar percentage (69.3%) revealed an awareness of the scoring criteria for their oral presentation. Further analysis indicated that there is a significant difference on the perceptions based on both proficiency level (F= 5.79, p=.00) and their location (F=28.77, p=.00). The results indicate that HEP and AEP students show a higher level of understanding of the grading system in comparison to the LEP students. In terms of location, both urban and suburban students indicated a higher level of understanding compared to rural students.

Table 11: ANOVA of Awareness of the SBOET Scoring Criteria

<table>
<thead>
<tr>
<th>Item</th>
<th>Crosstabs</th>
<th>Item</th>
<th>Yes</th>
<th>No</th>
<th>N</th>
<th>M</th>
<th>SD</th>
<th>F</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>I am aware of the scoring criteria for my oral presentation</td>
<td>Proficiency</td>
<td>HEP</td>
<td>72.3</td>
<td>27.8</td>
<td>800</td>
<td>1.28</td>
<td>0.45</td>
<td>11.6</td>
<td>.00</td>
</tr>
<tr>
<td></td>
<td></td>
<td>AEP</td>
<td>76.1</td>
<td>23.9</td>
<td>1069</td>
<td>1.24</td>
<td>0.43</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>LEP</td>
<td>65.7</td>
<td>34.3</td>
<td>712</td>
<td>1.34</td>
<td>0.48</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Location</td>
<td>Urban</td>
<td>77.6</td>
<td>22.4</td>
<td>1072</td>
<td>1.22</td>
<td>0.42</td>
<td>37.2</td>
<td>.00</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Suburban</td>
<td>75.4</td>
<td>24.6</td>
<td>800</td>
<td>1.25</td>
<td>0.43</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Rural</td>
<td>59.9</td>
<td>40.1</td>
<td>709</td>
<td>1.40</td>
<td>0.49</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Similarly, there was a significant difference in terms of awareness of the SBOET scoring based on both proficiency (F=11.62, p=.00) and location (F=37.25, p=.00). Findings displayed in Table 11 indicate that HEP and LEP students are more aware of the scoring criteria than AEP students. In terms of location, urban and suburban students indicated a higher level of awareness of the scoring criteria compared to students from the rural schools.

Challenges faced by students as test takers of the SBOET

This study also investigated the challenges students faced during their SBOET. The findings indicated that students’ responses ranged from test anxieties to logistical issues. As test takers, the respondents agreed that they constantly faced problems with self-confidence, nervousness and a substantial amount of pressure.
One of the main problems was their English language competency. A large majority (87.0%) cited that their limited language proficiency often left them with a loss for words to express themselves accurately. Interview sessions further indicated this limitation also resulted in them being unable to express their thoughts clearly and coherently so that they could be easily understood by the assessors. The respondents also cited pronunciation as a problem and some indicated that they often mispronounced words and sounds. The lack of general and technical vocabulary was also an identified problem by the respondents. Furthermore, they agreed that their knowledge of the structure of the language, i.e. grammar, was equally weak. For instance a student from Terengganu, had this to say:

“ I think my problem and also my friends’ problems are the same. We all have poor pronunciation and we find it difficult to talk in English and many times when we talk in the oral test we mix the Malay and the English words together . . . many of my friends also have poor grammar.”

A sufficient percentage (72.0%) also cited the task of having to prepare a written script prior to their presentations as a great challenge. The first problem in the script preparation was regarding the choice of topics given by their teachers as the respondents felt that most of the topics assigned were way beyond their linguistic abilities. This was further aggravated by the time allocation for the script preparation. A large majority of the students felt that they needed more time to research materials and get more information for the assigned topic.

Many students (89.0%) indicated that presenting before a live audience (i.e. their peers) was a great challenge. They highlighted that they felt very uncomfortable presenting in front of their teachers and their friends. On the other hand, some respondents felt that their presentations were not taken seriously as their classmates were not paying attention and as a result, confirmed their feelings of discomfort during their presentations. Others felt that since teachers did not place much emphasis on classroom control and discipline, the whole exercise of the SBOET seemed like a futile activity. This was articulated succinctly by a Secondary Five student from Kuala Lumpur. She said:

“I do not like this new oral test. I do not like to present my talk in front of the class. Many students laugh and I cannot speak because I too want to laugh. Then if someone is presenting, my friends sometimes do not pay attention. The next class is sometimes very noisy and no one can hear the students speaking in front. . . so we all do our own work”
Finally, most of the respondents cited the logistical issue as a very serious problem. As mentioned earlier, a few students felt that their classrooms were not very conducive environments for the SBOET as they were rather noisy and their presentations could not be heard. To make matters worse, some students were ridiculed or distracted during their oral presentations by their audience/classmates. Due to the limited space in the classroom, as the respondent is giving his or her presentation, the teacher as the assessor is bound to get distracted easily as he or she needs to maintain order in the classroom. Therefore, some respondents felt the teachers did not grade them accurately as they were more involved in maintaining class control than listening to their presentations. Hence, the classroom environment was perceived by the respondents as “unsuitable” for an oral assessment.

**Suggestions on improving the implementation of the SBOET**

The students were also asked to suggest ways to improve oral presentations in their classrooms. Most of their suggestions focused on practical routine issues, strategies employed by the students for self improvement and improving logistical issues by having a suitable environment and a good facility for the implementation of the English oral assessment in the schools.

A majority (68.0%) of the students said that they would like to see a more proactive role of their teachers in encouraging students to speak in English. For this to happen, they suggested that more time should be allocated for practice and more mock presentations prior to taking the SBOET. They would also like to be given a more formal briefing about the oral test. Through more mock presentations, the respondents believed that their teachers could provide more feedback to improve their weaknesses. In addition, the respondents also believed that teachers could reward the students with the best presentations as a form of incentive as they served as a good role model to others. Some students (25.0%) suggested that the oral test be done in ‘private’ where it is only between the teacher and the student. This suggestion might be in conflict with the procedures of the SBOET where the role of the audience is important.

The respondents also made several suggestions for enhancing the role of peers as they believed that their classmates could also have an impact on their performance in the oral test. The respondents suggested that their classmates should be a more serious audience during their oral presentations.

Finally, the respondents suggested that the oral test be conducted in a ‘more comfortable’ venue instead of the classroom. They believed that a test venue with
complete amenities would reduce the amount of test anxieties. If administrative, technical and logistical issues are handled professionally, many of the respondents felt they could perform better in the SBOET. The respondents also suggested that there should be more than one rater to assess their oral presentations so that issues pertaining to bias, inter-rater reliability and validity of the oral test could be minimized.

**Conclusion**

This study indicated that a majority of the students (81.6%) took the SBOET seriously and they (90.3%) know that it is a compulsory assessment. This is heartening as the formative school based oral test should be viewed as a platform for continuous improvement in their speaking skills. Nevertheless, students were rather divided as to their opinions and effectiveness of the SBOET as only 55.0% of them viewed it positively, whilst the remaining 45.0% held negative perceptions towards it. A divided view towards the SBOET is also found in a study conducted by Chen and Squires (2010), where their respondents were also divided in their opinion regarding the effectiveness of standardized oral tests. Such findings imply that teachers need to bridge the gap between these two groups and ensure sufficient information is given to help students understand the many benefits of formative assessments like the SBOET.

A majority of the students (90.5%) also revealed that they were briefed by their teachers and another 93.9% claimed that their teachers adhered to given guidelines as they were given a list of tasks to choose from. Nevertheless, dissemination of information to parents was however lacking. Close to 51.9% indicated that the school did not notify their parents and another two thirds (60.9%) of the students highlighted that school authorities did not put the SBOET schedule on the notice boards. Further analysis using the one-way ANOVA results indicated that parents of the students in the rural schools and students with limited language proficiency were better informed compared to the other groups. This again reflects the need for schools, especially TESL teachers and school administrators, to address the issue of effective dissemination of information on school based assessments.

The implementation of continuous formative school based assessments stress the importance of immediate feedback, but only 68.0% of the students indicated that they received feedback from their teachers after each assessment. Besides that, slightly more than half of the students (55.5%) were unhappy with the marks given by the teachers and some students questioned the validity and reliability of the scoring system. ANOVA analysis showed that HEP students received better feedback from their
teachers on their oral presentations compared to the AEP and LEP students. In fact
Long (2006) highlighted that school based assessments often come with a host of
challenges, one of which that needs to be addressed by both teachers and examination
bodies is the establishment of a system of teacher moderated school-based assessment
that can be trusted, and is valid and still reliable. These findings indicate the
implementation of the SBOET has not achieved a high compliance rate and this calls
for more regular monitoring by the relevant school authorities.

Furthermore, it is interesting to note that most of the students were reasonably
confident when taking the oral English test and results revealed that students’ level of
certainty corresponded with their level of proficiency. Those who lacked confidence
attributed it to the fact they were not comfortable presenting in front of their classmates
and 65.1% preferred to present privately with their teachers. Furthermore, a majority of
the LEP students memorized their script. To bridge this gap, it is pertinent that teachers
take into consideration students’ abilities and provide the necessary guidance and help
if the full benefits of formative assessment are to be seen.

Findings also revealed that more than half of the students (55.0%) pointed out
that their classrooms were not conducive venues for their presentations. This implies
that most schools are not equipped with the proper infrastructure and such issues of
logistics need to be taken into consideration for the successful implementation of the
SBOET. For school-based assessment to be successfully carried out, external validity
has to be taken into consideration. Schools must ensure that the classroom atmosphere
is conducive. External validity can be further enhanced with smaller enrolment per
class. The classroom must be equipped technologically (LCD, computers and screen) to
support the candidates’ presentations. Such support system should be in place for
school-based assessment to be successfully implemented.

To further enhance the implementation of formative assessment such as the
SBOET, both students and parents must be well informed of all assessments. The idea
of having a Student Kit which details the requirements of the SBOET is perhaps a good
suggestion. The Student SBOET Kit should include a manual that outlines the aims and
objectives of the SBOET together with the models to be assessed and the scoring
procedures. If possible it should also include a CD courseware that can provide samples
of each of the four SBOET Models so that students have some knowledge of what is
expected of them. This kit could also serve as a source of information for students’
parents/guardians. Consequently the relevant school authorities should also look into
ways in which dissemination and implementation of school based assessment can be
further improved and enhanced. It is perhaps timely that schools and the Ministry of Education take advantage of modern technology enhanced systems, such as emails, and set up e-portals for all stakeholders involved in school based assessments. Such technology enhanced mechanisms would without doubt help bridge the gap between all parties, especially teachers, parents and students.

The findings of this study reveal that Malaysian public schools are in the initial stages of the implementation of school based assessments and are therefore not without their share of challenges. This, however, should not be taken as a step towards dismissing formative on-going assessments, because researchers such as Cromey and Hanson (2000) stress that schools implementing school-based assessment often produce positive results for the schools, teachers and students. Adding to this discourse, Mitchell (1992) and Borich and Tombari (2004) assert that school based assessments have a significant impact on teaching style and student achievement and therefore teachers will need a number of different ways for assessing both the product and process of student learning. They also reiterate that the initial stages of implementation would demand considerable investment in terms of resources to enhance the teaching and learning process. Therefore, schools should be provided the necessary support, as school-based assessments are a step in the right direction because formative and continuous assessments reflect a more humanizing approach to testing and evaluation in schools.

Finally, the implementation of the SBOET, as perceived by the students, have shed light on relevant areas of concern which need to be addressed in order to bridge the gap between teachers and students. Therefore, teachers and schools should always strive to obtain “valid and reliable information before making important decisions that can influence student learning” (Airasian, 2005, p. 21). It is also hoped that the authorities concerned would look into these issues and put into place effective mechanisms in order to bridge the gap so that the true benefits of school-based formative assessments can be reaped by all parties concerned.

References


Launched in 2006, A Glossary of corpus linguistics is part of a series of glossaries by Edinburgh University Press, similar to the ones in Sociolinguistics, Applied Linguistics or Cognitive Linguistics. Authored by specialists from the Department of Linguistics and English Language at Lancaster University, the publication is welcome to a fast evolving research area.

The book opens with a warning about website addresses. Due to the speed at which they change, the authors have decided to include only those which are likely to remain for a longer period of time. Still, readers do not find the online address of the British National Corpus, for instance. Also in the ‘introductory notes’, a list of over 150 acronyms is offered even though, for the sake of consistency, the authors have decided to list the terms in full in the glossary.

The entries in the glossary are organized alphabetically, which helps finding specific terms such as ‘header’ or ‘raw corpus’. This way, the book may work like a dictionary. Although no theme clustering is offered by the authors, the publication might be summarized into six main categories.

First, there are numerous references to distinct concepts in corpus linguistics such as ‘alignment’, ‘hapax legomena’ and ‘representativeness’, to cite a few. There are also related notions in the fields of computational linguistics and statistics. Other entries which do not seem to be straightforwardly related to corpus linguistics at first also find a place in the glossary, for instance, ‘dictionary’, ‘ethics’, ‘form-focused teaching’ and ‘Java’.

A second category refers to existing corpora. There are entries on the Brown Corpus, the British National Corpus, the American National Corpus and the International Corpus of English. The publication additionally encompasses more
recent compilations such as the *Corpus del Español,* ‘a 100-million-word corpus of Spanish created by Mark Davies’ (p. 49) in 2001/2002. Some less popular corpora are also referenced in the glossary – for instance, *Cronfa Electroneg o Gymraeg* (representing modern Welsh) and the Guangzhou Petroleum English Corpus (developed under the auspices of the Chinese Petroleum University and the Jiao Tong University Corpus for English in Science and Technology).

Thirdly, several computer programs are mentioned in the volume. Not only are concordancers defined, but such programs – for example, ConcApp by C. Greaves, Concordance by R. J. C. Watt, and Concordancer / *Le Concordanceur* by D. W. Rand – are also described. Parsers and taggers find a home in the volume together with related concepts (‘parsing’, ‘part-of-speech tagging’, ‘skeleton parsing’, and ‘tag transition probabilities’), programs (‘Constraint Grammar Parser of English’, ‘Link Grammar Parser’, ‘Minipar’, ‘TAGGIT’ and ‘Trigrams’n’Tags’) and parsed and/or tagged corpora (‘CHRISTINE Corpus’ and ‘Gothenburg Corpus’).

Another category has to do with statistical tests, which can also be found in the glossary. Brief explanations are provided for the difference between parametric and non-parametric tests as well as their most commonly used types. Chi-square, log-likelihood and Fisher’s Exact Test, among others, are also explained.

A fifth group of entries encompasses those related to well-known journals and associations in corpus linguistics. The former includes references to the *International Journal of Corpus Linguistics* and the *Journal of Quantitative Linguistics.* The latter comprises, for instance, the Association for Computational Linguistics and its European chapter, the Association for Literary and Linguistic Computing, the European Association for Lexicography and the European Language Resources Association. Emphasis is placed on the European associations, perhaps due to the fact that the authors are based in the UK.

Finally, there are references to projects, archives, databases and universities. One example is Project Gutenberg, ‘a massive internet archive of over 16,000 copyright-free books stored as machine-readable text’ (p. 135). The Alex Catalogue of Electronic Texts resembles Project Gutenberg for it is ‘an archive of on-line, freely available texts that are copyright free’ (pp. 8-9). As far as historical databases are concerned, there is a reference to the Chadwyck-Healey Databases. The only location mentioned in the glossary is the University Centre for Computer Corpus Research on Language (UCREL), situated at the University of Lancaster, where the authors work.
In relation to further reading, some of the entries refer to publications in which readers may find out more information on a specific topic. As the bibliography is concerned, almost 200 works are listed. The volume refers to 4 publications from 2001, 9 from 2002, 3 from 2003, 4 from 2004 and only 1 from 2005, the year before the glossary was published.

The references in this volume are in most cases adequately explained and exemplified. In a user-friendly way, it provides a number of cross-references either in the middle or at the end of an entry. In the first case, they are in bold type. This system makes it easy for readers to find out what they are looking for if they need further information.

It is true that writing a book necessarily implies a selection. In the glossary, however, the absence of some terms should be reconsidered. For instance, despite the reference to Corpus del Español, the volume does not include Davies’s interface to search the British National Corpus (http://corpus.byu.edu/bnc) nor does it dedicate an entry to Davies and Ferreira’s O Corpus do Português (http://www.corpusdoportugues.org), totaling over 45 million words of Brazilian and European Portuguese from 1300s to 1900s. By the same token, readers may feel a lack of balance in the choice of institutions listed – mainly European associations and the university where the authors work. In terms of presentation, it would have been advisable if the entries had also been grouped thematically. As this is not the case, readers may find themselves flipping through the publication to learn about the different aspects of an issue they are interested in.

All these minor shortcomings, however, may be solved in a second edition. A Glossary of Corpus Linguistics does in fact provide very useful introductory information on corpus linguistics. The publication adds to the field by playing a twofold role: it stands as a source of concepts that corpus linguists should use worldwide and, at the same time, may work as an introduction to novices in the area.
Adult Language Learners

Reviewed by Stephanie A. Wilton
Kobe College High School
Nishinomiya, Japan

Adult Language Learners: Context and innovation addresses specific issues facing teachers who work with adult language learners in a wide variety of contexts such as university, continuing education, vocational training, and language teacher certification programs.

After a preface by the series editors and an introductory chapter by the volume editors, the book is divided into three sections: Teacher Development, Extending Learner Autonomy, and Innovations within a Course. Each section is composed of five to six descriptive articles, which address a range of contexts and course types that relate to the particular theme.

The first section, Teacher Development, is composed of six articles. It starts with Brandt’s chapter “Thinking Locally, Training Globally: Language Teacher Certification Reappraised,” in which the author suggests that program administrators take into account the local teaching context and include space for more reflection and feedback in their programs. In the next article, “Mind the Gap: Second Language Acquisition Theory into Practice,” McCormack describes how actual case studies and classroom practice have been integrated into a required second language acquisition course in an MA TESOL program. Following this, chapters four and five concentrate on technology used in teacher training programs. Kim discusses integrating technology training into a university TESOL course in “Podcasting and Online Journals as ESOL Resources.” And, in chapter five, “E-portfolios for Lifelong Teacher Development,” Baker, Crawford, and Jones demonstrate how online portfolios were used to support a distance learning post-graduate certificate program. After examining teacher training programs in the first four articles, the final two chapters of this section shift focus slightly to address language teacher practices.

In
“Teacher Enthusiasm in Action,” Ding presents specific ways for teachers to evaluate how they project themselves in the classroom, and Oanh compares how teachers can use memorization effectively in communication-based classes in “Memorization in Language Teaching: Vietnam and the United States."

The second section, *Extending Learner Autonomy*, also contains six articles. In the first, Murray describes a university course in which students choose their materials and set their own learning goals in the first chapter of this section, “A Self-Directed Learning Course.” Next, Lamping discusses how mentoring relationships help students build literacy in “Stepping into a Participatory Adult ESL Curriculum.” Continuing in the vein of encouraging student interaction in the classroom, Alexander illustrates how group work can encourage learners to take responsibility of their learning in “Learning Teams in Edinburgh.” Following these two chapters on learner interaction in the classroom, the emphasis shifts to building learner autonomy through classroom activities. In “An Interactive Approach to Book Reports,” Andrade demonstrates “one way of transforming the common book report into an interactive, integrated skills project” (p. 89), and Dias describes how students build computer and research skills while delving into controversial issues relevant to their community in “A Web of Controversy: Critical Thinking Online.” Finally, “Researching Pains: Iranian Students Exploring Medical English,” by Ghahremani-Ghajar, Mirhosseini, and Fattahi, illustrates how students personalize their learning by conducting research in English on a close relation’s physical or mental pain.

Each of the five articles in the final section, *Innovations within a Course*, portrays an array of inventive ideas for shaping curricula. Strong’s article depicts a course in which students explore foreign culture within their own city in “Field Trips with Japanese Student Ethnographers.” The next chapter, “Role-Playing with Fire: Hot Topics and Heated Discussions," by Stillwell, discusses methods for creating “an appropriate environment for the safe use of role-play” (p. 127). This is followed by “A Case for Discussion," in which Smith shows how case studies can be integrated into an English for academic purposes course. Next, “Motivating Thai University Students with Radio Drama,” by Kubanyiova, describes how a student-directed project helped increase motivation in an integrated skills course. And in the final article of the book, co-authors Augusto-Navarro, de Abreu-e-Lima, and de Oliveira demonstrate a process of designing and reforming an English for specific purposes.
course to meet the needs of the learners and their employer in “Ongoing Needs Analysis: English for Aviation in Brazil.”

This volume indeed contains a wide variety of articles addressing a multitude of contexts and course types. With limited technical vocabulary, this book is an easy read for teachers with any amount of exposure to the field. Overall, this volume manages to present highly accessible and practical ideas for the classroom, grounded in applied theory from the ESOL field.
Language and Education


Reviewed by Servet Celik
Karadeniz Technical University,
Turkey

Language and education, the 9th volume in the Collected Works of M.A.K. Halliday, is compiled from nineteen previously published papers and intended for language professionals. The book is organized into five parts and preceded by a previously unpublished speech by Halliday that serves as a preface to the text.

Following the editor’s brief summary, Part I delves into mother tongue education. The first chapter, “Linguistics and the Teaching of English,” underlines the role of linguistics in teaching English. Moving into the second chapter, “A ‘Linguistic Approach’ to the Teaching of the Mother Tongue,” Halliday discusses the links between language, individual, and society. In the third chapter, “Some Thoughts on Language in the Middle School Years,” Halliday argues that language is a resource “that a child constructs […] in interaction with those around him, rather than a set of structures or rules that he has to acquire” (p. 61). The next chapter, “Differences between Spoken and Written Language,” discusses literacy education at the secondary level, highlighting the differences between speech and writing. In the fifth chapter, “Language and Socialization: Home and School,” Halliday stresses the significance of language in the transmission of culture. Chapter 6, “Literacy and Linguistics: A Functional Perspective,” scrutinizes literacy from a linguistic standpoint that considers literacy as something which utilizes the “conceptual framework of linguistics […] as a way of understanding it” (p. 97).

Part II explores second language learning. Beginning with the seventh chapter, “General Linguistics and Its Application to Language Teaching,” Halliday discusses the definition of language and formal and contextual comparisons of languages. In the next chapter, “Is Learning a Second Language like Learning a First Language All Over Again?” Halliday argues that, despite the dissimilarities between natural and institutional learning, L1 and L2 learners demonstrate a comparable determination to thrive. However, the ninth chapter, “Learning Asian Languages,” illustrates that in spite of the drive to succeed noted in Chapter Eight, a child’s lack of multilingual experience might interfere with learning a new language.
Part III begins with the tenth chapter, “National Language and Language Planning in a Multilingual Society,” exploring the issue of language planning relative to the circumstances in a given culture. Halliday cites examples from around the world, indicating that it is critical to observe what is happening on a global level, rather than focusing exclusively on one’s own nation. Moving to the eleventh chapter, “Some Reflections on Language Education in Multilingual Societies,” he illuminates that language education in multilingual societies requires understanding language education as more than just a teaching of languages. He affirms that all aspects of learning must be recognized, and theory and practice should be closely linked. In the twelfth chapter, “Where Languages Meet: The Significance of the Hong Kong Experience,” Halliday explores the collision of languages in a community where everyone is expected to know two languages. He acknowledges concerns about over-exposure to English, but speculates that familiarity with the language might effectively preclude linguistic imperialism.

Part IV launches with “The Notion of ‘Context’ in Language Education,” which tackles context in relation to language education. Halliday’s analysis considers the interplay between language as system and language as text, and situational and cultural contexts. Then, in the fourteenth chapter, “Language across the Culture,” Halliday suggests that culture and curriculum are diffused through language, and the notion of language across the curriculum serves as a diversifying and unifying concept that helps maintain “the flow of meaning across the culture” (p. 305). Chapter XV, “Contexts of English,” highlights the post-colonial transformation of English. Halliday argues that English has become “resemanticized” (p. 323) in diverse cultures, ultimately leading to a mixture of Englishes.

Part V is devoted to educational linguistics. The sixteenth chapter, “A Response to Some Questions on the Language Issue,” marks Halliday’s answers to several questions which underscore how linguists can help achieve a just society. With the seventeenth chapter, “Some Basic Concepts of Educational Linguistics,” he outlines the multifaceted construct of language in education. Together with the eighteenth chapter, “On the Concept of ‘Educational Linguistics,’” this section accentuates the idea that educational linguistics targets a language-based model of teaching and learning. In the last chapter, “A Language Development Approach to Education,” Halliday investigates the discontinuity between home and school in terms of language development and encourages linking children’s early learning to formal schooling.

Thanks to wisely selected papers from four decades that emphasize research into topics such as first and second language learning, this book provides brilliant insight into contemporary issues in language education. Although the text is not always easy to read, Halliday’s personable style, drawing on his own experience and his inclusive examples, make this work another Halliday classic for language teachers and linguists.
Reading in a Second Language: Moving from Theory to Practice (Cambridge Applied Linguistics Series)

Jim Bame
Utah State University
Logan, Utah
U.S.A.

Reading in a second Language: moving from theory to practice by William Grabe is an impressive examination of what academic reading is and how to learn to read fluently. In order to accomplish this, the author selects core first language research (L1) reading themes, explains them, and examines the themes as they are applied to English as a second language (L2) research.

The book contains major sections with eighteen chapters, each of which ends with a useful section entitled Implications for Instruction that offers suggestions of how teachers, materials developers, and course designers could apply each chapter’s information to classrooms.

Part 1 includes 5 chapters explaining the foundations of academic reading. Chapter 1 outlines the purposes for academic reading and defines it. The next chapter outlines how reading works at lower processing levels of word recognition and semantic proposition encoding. In chapter 3, how reading comprehension emerges through top-down models of text and situation comprehension is described, along with reader resources and reading skills, such as strategies, goals, inferences, background knowledge, and comprehension monitoring. The next chapter depicts a number of complex cognitive concepts and issues in reading, such as implicit and explicit learning, automaticity, attention, inference, and background knowledge. The final chapter in this section outlines types of cognitive models and key concepts of well-recognized models.

Part 2 contains 4 chapters depicting learner differences. Chapter 6 overviews issues in learning to read in different L1s, possible L1 word recognition issues and their transfer to L2 reading. In the next chapter, the author surveys differences between L1 and L2 reading, specifically linguistic and processing, educational and developmental, and socio-cultural and
Institutional differences. Chapter 8 outlines several social-context sources of reading ability variation, such as social factors in L1 literacy, language-minority learners in public schools, and ESL and EFL students. The final chapter in this part describes definitions for motivation in reading, theories of motivation, and motivation for reading in L1 and L2 contexts.

Part 3, “Developing L2 Reading Comprehension Abilities,” consists of 4 chapters. This part, along with chapters 1-3, 9 and 14-16, would be of most interest to reading teachers. This is because these sections of the book provide very helpful ideas about applying current scholarly investigations about reading and readers to a research-guided teaching approach, designing units based on the approach, and implementing the units with day-to-day activities. Chapter 10 outlines L2 students developing main-idea comprehension, and the next chapter examines how to develop strategic readers. Chapter 12 describes the components of how texts signal discourse, text genres, narrative and expository texts, and patterns of discourse in text. Vocabulary and Reading Comprehension, chapter 13, outlines why vocabulary is the most critical resource for reading.

Part 4 expands on skills and instruction which help a reader become fluent. Chapter 14 defines and outlines L1 research in developing fluency in multiple settings, fluency’s relationship with comprehension, and word recognition fluency. L2 perspectives on word- and passage-reading fluency are then outlined. The next chapter examines research on extensive reading’s many benefits in L1 and L2 contexts. Research concerning curriculum development and instructional strategies are summarized in chapter 16. The last chapter in this section examines various aspects of assessment.

The book’s final chapter glances at important issues not examined thoroughly in the book, such as reading and writing relationships, teacher training, reading and technology, reading and new media, and reading disabilities.

Grabe’s book is a comprehensive discussion of an impressive quantity of research. The book synthesizes the significant highlights and organizes it in a lucid, readable style. From the beginning of the text, Grabe clearly outlines the book’s goals and limitations and accomplishes the former masterfully, all the while being cautious in his claims, allowing the weight of accumulated research findings to substantiate his conclusions. If there is a weakness in this book, it is that it may be a challenging read for a general audience. However, this book should be read by all those involved in most any facet of influencing readers learning to read, including policy makers and administrators, as well as teacher trainers, graduate students, researchers, and teachers.
### Associate Production Editors

<table>
<thead>
<tr>
<th>Name</th>
<th>Institution 1</th>
<th>Institution 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alison Smith</td>
<td>United Arab Emirates University</td>
<td>UAE</td>
</tr>
<tr>
<td>Robert Baird</td>
<td>Southampton University</td>
<td>UK</td>
</tr>
<tr>
<td>Dr. Toshinobu Nagamine</td>
<td>Prefectural University of Kumamoto</td>
<td>Japan</td>
</tr>
<tr>
<td>Dr. Wen-chi Vivian Wu</td>
<td>Providence University</td>
<td>Taiwan</td>
</tr>
<tr>
<td>Dr. Aradhna Malik</td>
<td>Vinod Gupta School of Management</td>
<td>India</td>
</tr>
<tr>
<td>Dr. Margaret Hearnden</td>
<td>Department of Educational Studies</td>
<td>University of York</td>
</tr>
</tbody>
</table>

### Editors

<table>
<thead>
<tr>
<th>Name</th>
<th>Institution 1</th>
<th>Institution 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Naoki Fujimoto-Adamson</td>
<td>University of Niigata Prefecture</td>
<td>Japan</td>
</tr>
<tr>
<td>Mr. John Baker</td>
<td>Chung Hwa College of Medical Technology</td>
<td>Taiwan, China</td>
</tr>
<tr>
<td>Dr. Kota Ohata</td>
<td>International Christian University, Tokyo</td>
<td>Japan</td>
</tr>
<tr>
<td>Dr. Malcolm Benson</td>
<td>Hiroshima Shudo University, Japan</td>
<td></td>
</tr>
<tr>
<td>Dr. John Thurman</td>
<td>Hokkaido University, Japan</td>
<td></td>
</tr>
<tr>
<td>Dr. Radha M.K. Nambiar</td>
<td>School of Language Studies &amp; Linguistics, Universiti Kebangsaan Malaysia</td>
<td>Malaysia</td>
</tr>
<tr>
<td>Annie Hong Qin Zhao</td>
<td>Education Department, University of Bath</td>
<td>UK</td>
</tr>
<tr>
<td>Dr. Toshiyuki Takagaki,</td>
<td>Faculty of Art &amp; Culture, Onomichi University,</td>
<td>Japan</td>
</tr>
<tr>
<td>Dr. Jane Mok Fung Yee</td>
<td>Department of English, Hong Kong Polytechnic</td>
<td>Hong Kong</td>
</tr>
<tr>
<td>Dr. Nooreiny Maarof</td>
<td>Dept of Methodology &amp; Educational Practice</td>
<td></td>
</tr>
<tr>
<td>Annie Hong Qin Zhao</td>
<td>Education Department, University of Bath</td>
<td>UK</td>
</tr>
<tr>
<td>Dr. Aysha Viswamohan</td>
<td>Dept of Humanities &amp; Social Sciences</td>
<td>Indian Institute of Technology</td>
</tr>
<tr>
<td>Dr. Ahmed Shakir El-Kilabi</td>
<td>Department of English, Nizwa College of Education,</td>
<td>Oman</td>
</tr>
<tr>
<td>Dr. Benedict Lin</td>
<td>Centre for Research in Applied Linguistics</td>
<td></td>
</tr>
<tr>
<td>Dr. Mohamad Ali Salmanti-Nodosuhaan</td>
<td>University of Zanjan</td>
<td>Iran</td>
</tr>
<tr>
<td>Mr. David Brown</td>
<td>The University of the Thai Chamber of Commerce</td>
<td>Thailand</td>
</tr>
<tr>
<td>Dr. Arda Arikant</td>
<td>Intensive English Language Institute</td>
<td>USA</td>
</tr>
<tr>
<td>Dr. Tan Bee Tin</td>
<td>Department of Applied Language Studies and Linguistics</td>
<td>University of Auckland</td>
</tr>
</tbody>
</table>

---

Dr. Mohammad Ali Salmanti-Nodosuhaan | University of Zanjan | Iran

---

Dr. Mr. Tan Bee Tin | Department of Applied Language Studies and Linguistics | University of Auckland | New Zealand
Dr. Keiko Sakui
Kobe Shoun Women University
Japan

Dr. Afefa Banu
Department Of English
King Khalid Women University
Abha,

Dr. Zahariah Pilus
International Islamic University,
Malaysia

Grace Hui Chin Lin
National Sun Yat-sen University,
Kaohsiung

Dr. Jun Zhao,
Marshall University, WV, USA

Dr. Cem Balciakılı
Gazi University,
Turkey

Gospel Ikpeme
Pearson Education UK & Longman
Nigeria

Dr. Nahla Shalhoub Bacha Lebanese
American University, Lebanon

Dr. Jianying Du
Huazhong University of Science and Technology
China

Margot McCamley
King Saud University
Saudi Arabia

Kenneth Ong Keng Wee
Nanyang Technological University
Singapore

Barbara Skinner
University of Ulster
Ireland

Dr. Elke Stracke
Faculty of Arts and Design
University of Canberra
Australia

Dr Stephanie Houghton
University of Kitakyushu
Fukuoka, Japan

Dr. Lawrence Jun Zhang
Nanyang Technological University
Singapore

Dr. Alexander Gilmore
Kyoto University
Japan

Elaine Schmid
University of Birmingham
UK

Nicholas Yates
Zayed University, Abu Dhabi,
UAE

Dr. Ji, Meng
National Institute of Chinese and Japanese Studies
UK

Dr. Angelia Lu
Nanyang Technological University,
Singapore

Dr. Andy (Anamai) Damnet Kasetsart
University, Kamphaengsaen,
Nakhonpathom, Thailand

Dr. Zahariah Pilus
International Islamic University
Malaysia

Dr. Hua hui Zhao
PhD Bristol, UK,
China

Stan Pederson
Faculty of Education,
Kumamoto University,
Japan

Dr. Neil Cowie
Foreign Language Education Centre,
Okayama University
Japan

Susana Gómez Martínez
University of Valladolid
Spain

Rebecca Todd Garvin
Indiana University of Pennsylvania
USA

Dr. Ahmad Al-Hassan
University of Petra,
Jordan

Mohammad Mehdi Soleimani
Karaj University,
Iran

Joseph Falout
Nihon University,
Japan

Dr. Grace Hui Chin Lin
National Texas A & M University
USA

Aisling O’Boyle
Queen’s University Belfast
Ireland

Dr. Sebastian Rasinger
Anglia Ruskin University,
Cambridge
UK

Ali Fuad Selvi
University of Maryland
USA

Distinguished & Honored Advisors

Professor Rod Ellis
University of Auckland
New Zealand

Professor Dr. Z.N. Patil
The English and Foreign Languages University
Hyderabad, India

Dr Reima Sado Al-Jarf.
College of Languages and Translation,
King Saud University,
Riyadh, Saudi Arabia

Prof. Vijay Bhatia
City University Hong Kong
Hong Kong

Dr. Phyllis Ghim-Lian Chew
Nanyang Technological University
Singapore

Professor David Nunan
Director & Chair of TESOL
Anaheim University, USA

Prof Emeritus Sandra L. McKay
San Francisco State Uni
Project Manager
Hawaii Pacific University, USA

Prof. Paul Nation
Victoria University
New Zealand
<table>
<thead>
<tr>
<th><strong>Senior Advisors</strong></th>
</tr>
</thead>
</table>
| **Professor Claire Kramsch**  
University of California  
Berkeley, U.S.A. |
| **Dr. Amy Tsui**  
Faculty of Education  
University of Hong Kong |
| **Dr. James P. Lantolf**  
Centre for Language Acquisition  
Pennsylvania State University  
U.S.A. |
| **Dr. Francis Mangubhai**  
University of Southern Queensland  
Australia |
| **Professor Robert Phillipson**  
Faculty of Languages, Communication and Cultural Studies  
Copenhagen Business School  
Denmark |
| **Dr. Alan Tonkyn**  
The University of Reading  
Applied Linguistics Dept  
UK |
| **Professor Terry Shortall**  
Birmingham University  
Birmingham  
UK |
| **Marc Helgesen**  
Miyagi Gakuin Women's University  
Sendai, Japan |
| **Dr. Robert David Carless**  
Hong Kong Institute of Education  
Hong Kong |
| **Dr. Jeong-Bae Son**  
University of Southern Queensland  
Australia |
| **Dr. Robert J. Dickey**  
Gyeongju University  
Korea |
| **Dr. Yasuko Kanno**  
Asst’ Professor English  
University of Washington  
U.S.A. |
| **Dr. Luke Prodromou**  
Leeds University  
UK |
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 (5th 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 (5th 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

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](http://www.apastyle.org/aboutstyle.html)

APA Citation Style: [http://www.liu.edu/cwis/CWP/library/workshop/citapa.htm](http://www.liu.edu/cwis/CWP/library/workshop/citapa.htm)

APA Style Workshop:

[http://owl.english.purdue.edu/workshops/hypertext/apa/index.html](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.

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.

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.