

The Asian EFL Journal Quarterly

March 2011

Volume 13, Issue 1



Senior Editors:

Paul Robertson and Roger Nunn



Published by the Asian EFL Journal Press

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

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

©Asian EFL Journal Press 2011

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

No unauthorized photocopying

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

Publisher: Dr. Paul Robertson
Chief Editor: Dr. Roger Nunn
Journal Production Editors: Robert Baird and Dr. Margaret Hearnden

ISSN 1738-1460

Table of Contents:

Foreword by Roger Nunn.....	6-9
1. Mohammed Zaid	10-47
- <i>Effectiveness of Organised E-mail Exchanges and Online Reading/Writing on College Students' Literacy Development and their Attitudes towards English: A Study from Saudi Arabia</i>	
2. Amin Saleh Moheidat and Abdallah Ahmad Baniabdelrahman.....	48-84
- <i>the Impact of Omani Twelfth-Grade Students' Self-Assessment on their Performance in Reading in English</i>	
3. Zubeyde Sinem Genc.....	85-113
- <i>EFL in Higher Education: Designing a Flexible Content-Based Curriculum at University-Level</i>	
4. Adel Abu Radwan.....	114-162
- <i>Effects of L2 proficiency and gender on choice of language learning strategies by university students majoring in English</i>	
5. Ayten İflazoğlu Saban and Dilek Yavuz Erkan	163-191
- <i>Writing Performance Relative to Writing Apprehension, Self-Efficacy in Writing, and Attitudes Towards Writing: A Correlational Study in the Turkish Tertiary-Level EFL Context</i>	
6. Ahmad Sharifi and Jaleh Hassaskhah.....	192-227
- <i>the Role of Portfolio Assessment and Reflection on Process Writing</i>	
7. Sun Xiao-zhao and Du Juan.....	227-269
- <i>Understanding the influence of L1 and lexical aspect in temporal acquisition: Quantitative and qualitative studies</i>	
8. Fatemeh Takallou.....	270-299
- <i>The Effect of Metacognitive Strategy Instruction on EFL Learners' Reading Comprehension Performance and Metacognitive Awareness</i>	
9. Tzung-yu Cheng.....	300-326
- <i>The Ability of Taiwanese College Freshmen to Read and Interpret Graphics in English</i>	

10. Xiaozhou Zhou and Victoria Murphy	327-355
- <i>How English L2 Learners in China Perceive and Interpret Novel English Compounds</i>	
11. Hideki Sakai.....	356-384
- <i>Do Recasts Promote Noticing the Gap in L2 Learning?</i>	

Book Reviews

1. *Language Teacher Research in Africa*

Thomas S. C. Farrell and Leketi Makalela (Eds)

Reviewed by **Iris F. Levitis**..... 385-387

2. *An Introduction to Conversation Analysis*

A. J. Liddicoat

Reviewed by **Hayriye Kayi** 388-389

3. *Feedback in Second Language Writing: Contexts and Issues*

Ken Hyland & Fiona Hyland (Eds)

Reviewed by **Aysha Viswamohan**. 390-392

4. *Diagnosing Foreign language Proficiency: The Interface Between Learning and Assessment*

Charles J. Alderson

Reviewed by **Deepti Gupta**..... 393-395

***Asian EFL Journal* editorial information and guidelines**..... 396-405

Foreword

Welcome to the first quarterly issue of 2011 in which we again cover a wide range of topics investigated in many different locations in Asia. Several of these studies illustrate attempts to use new developments in language learning scholarship to improve curriculum.

Dr Mohammed Zaid (*Effectiveness of Organised E-mail Exchanges and Online Reading/Writing on College Students' Literacy Development and their Attitudes towards English: A Study from Saudi Arabia*) explores the effectiveness of organised emailing and webquesting on college students' attitudes towards learning English. His findings indicate that the model is especially effective when integrated with collaborative learning within a "problem-solving inquiry-based learning environment". Improved attitudes towards English were an additional bonus.

Amin Saleh Moheidat & Abdallah Ahmad Baniabdelrahman in *the Impact of Omani Twelfth-Grade Students' Self-Assessment on their Performance in Reading in English* investigate the impact of Omani twelfth-grade students' self-assessment on their performance in reading in English using one-minute papers and rating-scale sheets. The findings indicate that the students' self-assessment had a positive effect on their reading performance in English.

In *EFL in Higher Education: Designing a Flexible Content-Based Curriculum at University-Level*, Zubeyde Sinem Genc emphasizes the need to offer ELT students in higher education more than general proficiency in English. Her description of a content-based language program at a Turkish university should provide a useful comparison for curriculum developers in other contexts.

In *Effects of L2 proficiency and gender on choice of language learning strategies by university students majoring in English*, Adel Abu Radwan responds to a call by Rebecca Oxford to examine the relationship between LLSs and various factors in a variety of settings and cultural backgrounds. Radwan investigates the use of language learning strategies of 128 students majoring in English at Sultan Qaboos University (SQU) in Oman. The study extends our current knowledge of the relationship between the use of language learning strategies and gender and English proficiency. The findings allow Radwan to suggest that the EFL cultural setting may be one factor that determines the type of strategies preferred by learners.

In an investigation of tertiary-level EFL students, Ayten İflazoğlu Saban and Dilek Yavuz Erkan (*Writing Performance Relative to Writing Apprehension, Self-Efficacy in Writing, and Attitudes Towards Writing: A Correlational Study in the Turkish Tertiary-Level EFL Context*) report a negative correlation between writing apprehension and writing performance and between writing apprehension and writing self-efficacy. The surprising result that would warrant further investigation is the positive correlation between writing apprehension and attitude towards writing. This study again underlines the importance of raising students' levels of confidence to improve their performance.

In *the Role of Portfolio Assessment and Reflection on Process Writing*, Ahmad Sharifi and Jaleh Hassaskhah underline the close link between language teaching and testing, arguing that it is “impossible to work in either without taking the other into account”. Their present study investigates the effect of using a portfolio assessment technique and reflection activities on students' writings and process writing. Their findings point to a positive effect of the treatment and also reveal that the students' attitude was positive towards portfolio-based learning.

Sun Xiao-zhao & Du Juan (*Understanding the influence of L1 and lexical aspect in temporal acquisition: Quantitative and qualitative studies*) report a mixed-method comparative investigation of Chinese and Japanese EFL learners on the influence of L1 in temporal acquisition. Study one used a written error recognition and correction task whereas study two examined the meta-cognitive process by means of qualitative data obtained from retrospective interviews. L1 influence was found to be an active factor in Japanese learners' progressive marking performance but L1 influence was not found in Chinese learners.

Fatemeh Takallou (*The Effect of Metacognitive Strategy Instruction on EFL Learners' Reading Comprehension Performance and Metacognitive Awareness*) uses a multi-method approach to investigate metacognitive strategy instruction in relation to reading comprehension performance (on authentic and inauthentic texts) and metacognitive awareness. They found that instruction on ‘planning’ and ‘self-monitoring’ did have an impact on performance and that awareness of metacognitive strategies significantly

increased after instruction.

Tzung-yu Cheng (*the Ability of Taiwanese College Freshmen to Read and Interpret Graphics in English*) investigates a topic that is arguably not given enough attention in language learning publications: the ability of college freshmen to read and understand English graphics. The study finds that the ability to read English itself is a major cause of difficulty, concluding therefore that it is “important to train high school students to read and interpret graphics before they enter college”.

Xiaozhou Zhou and Victoria Murphy, in *How English L2 Learners in China Perceive and Interpret Novel English Compounds*, investigate the way English learners in China interpret novel English noun-noun compounds. Their study “extends the research area into the L2 domain with a view to comparing interpretations of L2 learners with those of L1 children”. Results pointed to overall competence in understanding and interpreting novel English compounds. The differences in performance success between the intermediate and advanced groups were not found to be statistically significant. Possible factors affecting participants’ interpretations and vocabulary teaching implications are discussed at the end of the study.

Hideki Sakai (*Do Recasts Promote Noticing the Gap in L2 Learning?*) reports the effects of implicit negative feedback in the form of recasts on noticing. The participants’ verbal reports about noticing their errors were elicited through stimulated recall. Results suggest that recasts did have an impact on L2 learners’ ability to notice their errors and that merely providing opportunities for production is not likely to have the same impact.

**Effectiveness of Organised E-mail Exchanges and Online
Reading/Writing on College Students' Literacy Development and their
Attitudes towards English: A Study from Saudi Arabia**

Dr Mohammed Zaid

King Khalid University, Abha, Saudi Arabia

zaid20sa@hotmail.com

Bio Data:

Mohammed A. Zaid is an associate professor of applied linguistics in the College of Languages and Translation, King Khalid University, Abha, Saudi Arabia. He obtained his MA and PhD from the University of Colorado at Boulder, USA. He has published research in the USA, UK, Germany, Egypt and Saudi Arabia. His research interests include L₂ acquisition, language teaching technology, reading, language and culture.

Abstract

Being a quasi-experimental study, this study explored the effectiveness of organised email exchanges and online reading utilising webquests on developing reading and writing of college students. The study explored the effectiveness of organised emailing and webquesting on college students' attitudes towards learning English in a Saudi university. The underlying premise is that e-learning can assist foreign language students in their progress in reading and writing. The study employed a four-stage model integrating organised emails and webquests in an inquiry-based learning approach. Findings indicate that the model is effective in ameliorating reading and writing of college students, especially when integrated with collaborative learning, all in a problem-solving inquiry-based learning environment. Findings also indicated improved attitudes towards English upon utilising this collaborative, enquiry-based learning model.

Introduction

Resources in online e-learning environments, from online support for traditional classes to online courses and full online degree programmes, can efficiently induce significant improvements in learning, especially the use of synchronous and asynchronous communication methods in higher education institutions (Andrews & Haythornthwaite, 2007). One such application is the electronic mail (e-mail for short) which has suffered from a paucity of research in the field of composition theory. Electronic mail is defined as “an electronic communications system that is used to send information from one person/site to another (one-to-one communication) or from one person to many people at the same time” (Cross and Raizman, 1986, p.3). In its pedagogical context, D’Souza (1992, p.22) defines email as a programme that “uses computer text-processing and communication tools to provide a high-speed information exchange service” which can have strong, effective applications in teaching reading and writing. This is because one of the most prominent features of e-mail communication lies in its ease of editing, storing, duplicating, and delivering or distributing (Huff, Sproull & Kiesler, 1986). This feature, when combined with its suppressing of social cues, lends itself to the flexibility both in the content scope in which e-mail functions and in the direction information flows. Pedagogical applications of emailing for literacy development have been theoretically contemplated (e.g. D’Souza’s, 1992; Warschauer, 1995).

McComb (1994) has argued that critical learning occurs when students are engaged in critical reflection on their personal, political, and social lives. E-mail communication seems to promote a “pedagogy that encourages students to be active creators of, rather than passive to, society” (p. 157). Furthermore, learning activities, from this perspective, come from students’ concerns and interests, and occur in reflection and interactions. It

has also been confirmed in Chiu's study (2005) that email communication, when pedagogically employed, can lead to better classroom interactions and creative learning. This can also improve and build up students' confidence and facilitate their writing skills in literacy classrooms. Hiltz, et al. (2007) relevantly indicate that the field of Asynchronous Learning Networks (ALN) integrates social and technical aspects, linking together teachers and learners in e-learning circles, especially if these groups are involved in collaborative social/ pedagogical processes emerging from the communication that is supported by the technology.

Further, it is surmised that e-mail technology can positively help students bridge the gap between "skill-getting and skill-using", borrowing the expression from Rivers (1975, p. 12). Omaggio-Hadley (1993) observed that the use of computers and computer-assisted instructional programmes in teaching foreign and second language composition is one relatively recent development in writing instruction, and it is so much so that preliminary experiments affirm this apt observation. Earlier, Smith (1999) found that, upon comparing two writing instruction methods - one computer-mediated and the other traditional, "the computer users improved significantly in their ability to read and express oral and written ideas" (pp. 80-81).

Studies by Neu and Scarcella (1991) Phinney (1991), and Thaipakdee (1992) all found that students had positive attitudes towards reading and writing with computers and less apprehension about writing, and more importantly, students developed better attitudes towards reading and writing with computers; this corresponded with better writing.

By the same token, Tao (1995, p.7) reviewed the literature and investigated "the still nebulous area of e-mail implications" in literacy learning. He indicated that:

"(1) in reducing the social cues and enriching functions of easy editing,

storing, and manipulating, e-mail lends itself to more user-control and user responsibility; (2) e-mail is capable of bringing traditionally peripheral persons into the instructional mainstream; (3) e-mail offers users chances to develop positive attitudes but can also touch off some undesirable behaviors; (4) due to its reduced sense of presence, e-mail communication exhibits some sense of anonymity and depersonalization; (5) e-mail is direct, straight forward and more self-centred; and (6) anecdotal reports suggest that e-mail might have effects on social behaviors such as collaboration and motivation, metacognitive aspects central to the learning process especially in writing”.

In addition, Stein and Stein (1995) revealed three rationales that function as the theoretical framework of e-mail-based literacy development and instruction:

Rationale 1: The e-mail training and assignments are embedded in the context of the overall language class objectives.

Rationale 2: Throughout the e-mail training, the instructor provides the participants with scaffolds of how to compose a message, how to send a message, how to receive a message and how to save a message by downloading it or printing it.

Rationale 3: E-mail instruction works best with training with awareness. Further e-mail applications e.g. delete a message, a folder, change user password send the same message to multiple users (with cc), are then introduced as students become expert e-mail users.

Important as this medium is, as having become our principal, steadfast form of organized communication, it has not yet been adequately studied (Hawisher & Moran, 1993; Tao and Reinking, 1996; Schaefer & Dillman, 1998; Blase, 2000; Rodrigues, 2007).

In the field of literacy education, research is on the increase in e-learning and online education (Anderson & Lee, 1995; Bolter, 1991; Reinking, 1994; Harris, 1994; Andrews

and Haythornthwaite, 2007); yet, literacy research in e-mail and its impacts on literacy acquisition seem at most scanty and hardly begun.

Given our recognition of the so-called mystique of e-mail, however, the empirical testimony for using e-mailing to enhance the processes of reading and writing online is scarce, and may perhaps be nonexistent among those learning English as a foreign language in our Arabic-speaking community. Even in non-Arab environments, only a few studies have been conducted in the United States delving into the impact of the electronic mail and the uncovered promises that it may have for writers in English (Anderson and Lee, 1995; Lee, 1994; Mabrito, 1991; McKay, 1989; Traw, 1994; Warschauer, 1995; 1996). Some of these studies explored the influence of e-mail as to cooperative learning and interaction in literacy (Allen & Thompson, 1995; Smolenskey, et al., 1990; Kern, 1995; Fey, 1994; Reiss, 1996), while other studies emphasized the effect of e-mailing on cognitive development as related to language (Kaufman, 1998). Hawisher & Selfe (2007) , therefore, call for a critical review of the 'pedagogical enthusiasm to new technological innovations' in order to shift researchers' focus from the technology itself to the ways in which such technologies shape and are shaped by students and teachers, rhetorical and pedagogical contexts, composing products and composing processes (Snyder, 2007; Whitworth, 2007; Hawisher & Selfe, 2007).

Furthermore, as tacitly voiced in Tao (1995, p.22) and further work (e.g. Rodrigues, 2007), the e-mail impact on the cognitive aspects of literacy acquisition and instruction still remains unclear.

The purpose of the present study was to look for empirical testimony for the most important role of e-mailing in enhancing the basic reading and writing skills through using grammar checkers, spellcheckers, dictionaries, etc. available in the compose box,

and developing intercultural communication in EFL learning in Saudi Arabia. As well as this, it seeks to integrate online reading using a webquests model with email exchanges in reading and writing learning to check their effectiveness on developing under-developed skills, as well as the students' attitudes towards EFL.

This study further seeks to recognize students' attitudes towards English as a foreign language. Many a researcher has claimed a strong relationship between students' attitudes towards language learning and achievement rates; they concluded that students' attitudes are an integral part of learning and that they should become an essential component of second/foreign language learning pedagogy (Weinburgh, 1998; Gardner, 1985; Gardner & Lambert, 1972). The premise that the importance of attitudes assumes in learning is that 'Cognitive theories of learning will be rejected unless a role is assigned to affectivity' (Brown, 1994: 134). However, Internet-based Instruction (IBI) for language learning was found to have varying effects on students' attitudes towards foreign/second language instruction. In Chen's study (2004), 1,026 freshmen and sophomore students in Taiwan developed positive attitudes towards educational technology use for EFL instruction. In a similar vein, Felix (2001) reported that on the whole, students were positively inclined to working with the web and found it useful, with the majority preferring to use the web as a supplement to face-to-face teaching. Furthermore, intermediate level community college ESL students and teachers expressed very positive attitudes toward using IBI (Schnackenberg, 1997). Some of this research indicated age and gender differences that explain variation in enhanced attitudes.

Context of the Study

Based on the experience of the researcher as an instructor of English language skills

courses, especially writing and reading, and on the low scores achieved on these two skills by the population of the study, the problem of underachievement in reading and writing skills appears prominently worthy of treatment. Students regard their reading and/or writing classes, in their traditional delivery methods, as courses in which they are offered reading passages in class or as homework, or they are assigned titles for essays and short paragraphs to write about comfortably at home, submitting samples of their writing to the instructor who, in his/her turn would instruct them in some of the skills of the academic year in a uni-directional fashion; i.e., with students; in this approach writing and/or reading is a product. The problem of the study is therefore summarised in the following research questions:

What is the effectiveness of an instructional programme grounded in email exchanges and reading online on developing reading and writing skills among students of English?

How does using e-learning in this programme affect the students' attitudes towards English?

Importance of the Study

This study is based on the premise that e-learning can assist foreign language students in their progress in reading and writing skills. Empirical testimony to support this observation is needed as voiced in some studies (Tao, 1995; Tao & Reinking, 1996; Anderson, 2002; Baron, 1998; Smith, 1999).

This study, therefore, describes the impact of e-mailing on how well participants will acquire basic reading and writing skills and subskills, predetermined as lacking in those students, and how this delivery medium would affect the students' attitudes towards learning English. Furthermore, this study is based on prior rationale research, which

supports the use of Internet-based instruction (IBI). Evidence from prior research suggests that this environment supports active, collaborative learning and the construction of knowledge (Killins, 2002). The study is important to the field of English language teaching for the following reasons:

1. This study is conducted in line with the international trend of computerized education followed in advanced countries in all stages of education; trends include computer-based learning and internet-based learning and instruction;
2. There is a paucity in research studies in Arabic literature, especially in Saudi Arabia, which tackle Internet-based instruction (IBI); research is very meager or non-existent here in Saudi Arabia.
3. The findings of the study may contribute to enhancing and developing new courseware that can be effective for teaching and learning reading and writing skills, given that the College of Languages and Translation will soon shift to an e-learning paradigm in the next few years.
4. The findings of the study may contribute to constructing new learning models and instructional designs appropriate for EFL skill instruction which could be new assets to the theory and practice of e-learning in English language Teaching (ELT).

Research Methodology

A. Subjects and Sampling:

A typical university classroom of IV-level students enrolled in the English department constitutes a representative population. Sampling was purposeful: thirty male students were selected on the basis of their basic computer skills (as measured by a computer

skills questionnaire¹) to participate in the experimental group. Other participants of the same level were randomly assigned to the control group.

B. Experimental Design:

This is a qualitative/quantitative study of the effectiveness of Internet-based learning and instruction. The study utilizes a model grounded in the WebQuest theory and involves online reading and e-mailing for developing reading and writing skills in EFL students. The experimental design used in this study is of the type: Pretest-Posttest Control Group Design (Gay, 1996).

This design involved one treatment group; the experimental group which received instruction into reading and writing skills for the academic year 2007-2008. An instructional design underlying the teaching method was devised based on the WebQuest model by Bernie Dodge and developed by the researcher and the incorporation of email as a useful technique for developing reading and writing skills (Dodge, 1997; 1998).

Data were analyzed using t-tests, mean scores, One-way Analysis of Co-variance (ANCOVA); other statistical methods used included averages, weighted percentages, Chi² and qualitative data analyses. Qualitative analyses were done using students' e-logs and the instructor's diary. E-logs are peer-review assessment sheets with items scaled from awesome to aweless on a five-point Lickert Scale to peer-evaluate learners' performance during the session and after. The teacher's diary is essentially a diary where one would record one's classes, the activities and language being taught, the problems

¹ This is a questionnaire aiming to sort students who are regular attendants of English majors at Level IV. The questionnaire is set in three sections; it starts with gathering demographics of the students, then, in section one, it seeks student information as to whether students have computers at home or not and what type they are, whether they are connected to the Internet or not, and their experience and attitudes towards computers. Section two of the questionnaire gleans information about students' computer skills, i.e. keyboarding skills, file management, word processing and operation systems. Section three seeks information as to skills of using the Internet and e-mail, i.e. basic knowledge of the Internet and frequency of using e-mailing for correspondence.

experienced and so on so that over time one can build up a reference book of problems and their solutions, useful activities and such like. After each session, the researcher jotted down whatever happened during the e-session in his diary to make amendments and modifications in the e-workbook which contains the session plans for e-learning.

Limitations of the study

The present study is constrained by the following delimitations:

- 1) The sample of the study, which consisted of male students only due to the nature of the segregated educational system in the country, included the IV Level Writing majors for the academic year 2007/2008;
- 2) The instruments of the study;
- 3) The statistical measures and instructional methods used in the study.

C. Procedures of the Study:

With students having attained Level IV, still some of the reading and writing skills are either missing or weakly developed. The researcher conducted a Reading/Writing Questionnaire to recognize the reading and writing skills direly needed by the students of the English department. This study sought to check the effectiveness of varied instructional delivery medium on learning outcomes and the students' attitudes towards learning English in this context. Therefore, an e-course utilizing webquests, an inquiry-oriented activity in which most or all of the information used by learners is drawn from the Web, and e-mail were developed based on the core skill areas and the sequence of the traditional course for developing basic reading and writing skills.

The Instructional Model

There have been so many models of incorporating e-learning in the EFL curriculum; in the field of ELT, many attempts have been exerted to transform the traditional classroom into computer-based, computer-mediated, and Internet-based media of instruction. Further, ELT methodologists in collaboration with technologists looked for the advantages of the World Wide Web not only as a medium of instruction, but also as a source of courseware and content for ELT, either in native speaker communities or in contexts where ELT is delivered as ESL or EFL (e.g. Dodge, 2003a; 2003b; Bartoshesky & Kortecamp, 2003; March, 2004; Menchaca & McVicker, 2003).

This is a procedural model that integrates the benefits of e-mailing and webquests for enhancing writing skills with online reading as an inquiry-based learning approach. The model is grounded in the webquest theory, and was developed in this study in order to develop reading and writing. A 'WebQuest' is a constructivist approach to learning, and is operationally described as follows:

‘Students not only collate and organize information they've found on the web, they orient their activities towards a specific goal they've been given, often associated with one or more roles modeled on adult professions. Since students have to participate in the elaboration of their learning strategies, the level of autonomy and creative production they attain is increased. With the proper guidance and "scaffolding" students can accomplish far more actual learning than in traditional transmission-of-knowledge situations that so often leave them wishing they were anywhere but in the classroom (Benz, 2001, p.12).’

E-mailing, however, is operationally defined as the first stage of the webquesting model

when students sign in to their accounts or their team accounts to get introduced to the topics of the webquests used in the literacy development classroom, or to get instructions as to their e-sessions. It is also the last stage of the model when students submit copies and/or samples of their work to other teams or to colleagues or to the instructor.

To summarise, this is a circular model for teaching online reading/writing in collaborative teams based on the notions of constructivism and inquiry-based learning. The design moves around six stages diagrammatically depicted in *Figure 1* below.

Many studies agree that e-mail communication seems to lend itself well to collaborations of various kinds (Fey, 1994; Mabrito, 1991; Schwartz, 1990, Selfe, 1990; Spitzer, 1989; Traw, 1994). Collaborations of any kind involve interactions with other people within society. Since social interactions are believed to affect literacy acquisition (Vygotsky, 1978), the possibility of collaboration offered by e-mail communication provides a viable means for understanding e-mail's potential in promoting literacy acquisition. This type of collaboration among students in sharing their perceptions and enhancing understanding of others provides the true context for meaningful writing.

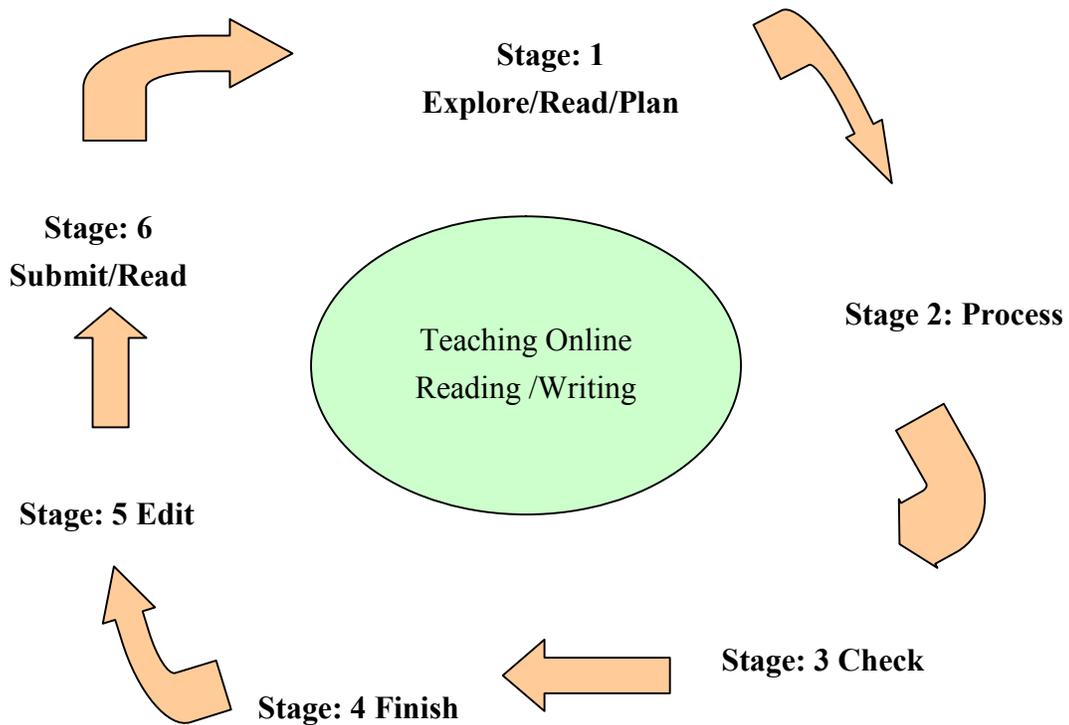


Figure 1: *the instructional model grounded in e-mail exchanges*

Stage 1: Explore/Read/ Plan

At this stage, students explore the Internet and read online using webquests to explore the topic of the writing. To plan their outline for their writing, they will explore the online materials they have collected, and in their respective teams, they brainstorm ideas for the paragraphs / essays as to what to cover, how to cover it, and how to devise an outline for the writing project. Students at this stage should filter the bulk of ideas they obtain and choose amongst them, e-mailing to one another what each of them has decided. Here at this stage, the instructor initiates the webquesting process, which should proceed as follows:

Introduction:

The introduction section provides background information and motivational scenarios such as giving students roles to play, e.g. "You are an underwater research scientist" or "You are an astronaut planning a trip to the moon." It also provides an overview of the learning goals for students. The goal of the introduction is to make the activity desirable and fun for students. When projects are related to students' interests, ideas, past experiences, or future goals, they are inherently more interesting. The goal of the motivational component is to engage and excite students at the beginning of each WebQuest.

Task:

The task is a formal description of what students will have accomplished at each stage of the webquests; e.g. Log online to www.google.com. Search for the entry 'astronauts' or 'voyage to the moon'.

Having been guided to the essay topic, the teams find resources for that particular topic on the Web. Then, the instructor devises an activity for the students that incorporates the information from the various sites. This task should be doable and interesting.

Stage 2: Process

This is the 'explain/expand/support' phase. Students in their respective teams seek to synthesise, expand, and provide justifications and in stage 4 edit their final product.

At the step of explanation, students analyse (a step back to stage 1 and forward to stage five) and synthesise their preliminary notions. Then in the step of expansion, they apply their agreed upon ideas and extend them. In the support step, they give reasons as to what they select and provide justifications for or against such ideas in team discussions.

The instructor provides ideas, feedback and support when necessary. This is a description of the steps learners should go through in accomplishing the task, with links embedded in each step. Step three of the webquesting model by Dodge (1997), known as Resources, is fitting here; this section of the WebQuest consists of a list of the resources (bookmarked Web sites, print resources, etc.) that the students will need to complete the task.

Stage 3: Check

This is the 'test/assess/evaluate' stage; in this stage, students apply the evaluation criteria for their finalised writing products. Evaluation criteria cover language, style of writing, critical thinking/critical reading, resource management, and content (quality/quantity). Students will peer-review their writings in teams according to a set of standards and indicators. There is also ongoing evaluation where teams exchange their writings with one another within their respective teams through email, and to other teams and the instructor via the secretary of the team, using their CHECK tools.

The instructor at this stage evaluates the students using a follow-up/ reporting form: here, the plan is that the instructor looks for how the students:

- Depict (decide, choose);
- Collect (download, classify, save);
- Analyse (Explain, give reasons, support);
- Synthesise (read extensively, read critically, author);
- Evaluate (self-review, peer-review, instructor-review);
- Finalise (edit, check for style, double-check for language and mechanics,

conclude)².

Stage 4: Finish

This is the 'Conclude/Suggest/Recommend' stage. Students, having finished their writing products, then write the last draft, and teams suggest to one another points where change is most likely to be advantageous (intra-team finishing and inter-team finishing); they would also recommend new ideas or language corrections where necessary.

The instructor at this stage moves about and suggests necessary changes; additionally, this can also be done via email with feedback sent to the students.

Stage 5: Style

This is the 'edit/change' stage, featuring new suggestions and recommendations from teams (intra/inter-team recommendations) and from the instructor sent over email. This phase of the model is associated with the conclusion step in Dodge's webquesting model, allowing for reflection by the students and summation by the teacher. Setting aside time for discussion of possible extensions and applications of the essay honours the constructivist principle: "We learn by doing -- but we learn even better by talking about what we did." During the concluding section of Webquests, the instructor can encourage the students to suggest ways of doing things differently to improve the essay.

Stage 6: Submit

Having edited the final version against the newly received feedback, students submit it to the instructor/peer for review. This is the stage now termed present/feedback stage. The

² These ideas are the researcher's, further developing the WebQuest paradigm.

cycle could be reinitiated cyclically with new ideas generated for new reading/writing topics or more extended development of the same essay, or an argument based on the older one.

Findings

Hypothesis one states that “There are no statistically significant differences between the mean scores of the experimental group subjects and those of the control group on pretesting as measured by a Reading/Writing skills test (1, 2, 3, 4, 5, and the total score).”

A t-test for two independent samples was utilized.

Table 1. t-test results for independent samples (experimental and control) on pretesting of reading skills (1,2,3,4,5, and the total score)

Skills	Subjects	Mean Score	Standard deviation	no	t-value
Skill 1 Previewing	<i>Experimental Group</i>	13.23	1.695	30	0.015
	<i>Control Group</i>	13.24	1.615	25	
Skill 2 Skimming	<i>Experimental Group</i>	12.70	1.745	30	0.039
	<i>Control Group</i>	12.68	2.056	25	
Skill 3 Scanning	<i>Experimental Group</i>	13.77	2.431	30	0.834
	<i>Control Group</i>	13.16	2.968	25	
Skill 4 Questioning & Reviewing	<i>Experimental Group</i>	12.27	2.377	30	0.400
	<i>Control Group</i>	12.52	2.293	25	

Skill 5 Note-taking	<i>Experimental Group</i>	13.23	2.112	30	1.764
	<i>Control Group</i>	12.36	1.411	25	
Total score on all skills	<i>Experimental Group</i>	65.20	5.623	30	0.845
	<i>Control Group</i>	63.96	5.168	25	

Table 2. *t*-test results for independent samples (experimental and control) on pretesting of writing skills (1,2,3,4,5, and the total score)

Skills	Subjects	Mean Score	Standard deviation	no	t-value
Skill 1 Making a time plan	<i>Experimental Group</i>	13.57	3.350	30	0.111
	<i>Control Group</i>	13.68	4.200	25	
Skill 2 Outlining & organization	<i>Experimental Group</i>	13.50	2.980	30	0.529
	<i>Control Group</i>	14.08	5.049	25	
Skill 3 Coherence & Cohesion	<i>Experimental Group</i>	12.17	2.984	30	1.393
	<i>Control Group</i>	13.52	4.204	25	
Skill 4 Style	<i>Experimental Group</i>	14.37	5.156	30	0.505
	<i>Control Group</i>	13.72	4.148	25	

Skill 5 Presentation	Experimental Group	13.4667	3.28773	30	0.133
	Control Group	13.6000	4.15331	25	
Total score on all skills	Experimental Group	67.07	9.063	30	0.585
	Control Group	68.60	10.380	25	

Results from the above tables indicate that the hypothesis is verified, as there were no differences between experimental subjects and control subjects that were of statistical significance.

To verify hypothesis two which states: "There are no statistically significant differences between the mean scores of the experimental group subjects and those of the control group on posttesting as measured by the reading/ writing skills test (1, 2, 3, 4, 5, and the total score)." A t-test for two independent samples was utilized. Tables 3 and 4 show the results of the t-test:

Table 3. t-test results for experimental and control subjects on posttesting reading skills

Skills	Subjects	Mean Score	Standard deviation	no	t-value
Skill 1 Previewing	Experimental Group	18.12	1.737	30	7.417**
	Control Group	14.50	1.878	25	
Skill 2 Skimming	Experimental Group	18.84	1.709	30	**12.857
	Control Group	13.10	1.573	25	

Skill 3 Scanning	<i>Experimental Group</i>	18.92	1.882	30	**12.617
	<i>Control Group</i>	12.90	1.605	25	
Skill 4 Questioning & Reviewing	<i>Experimental Group</i>	18.68	1.971	30	9.985**
	<i>Control Group</i>	12.66	2.495	25	
Skill 5 Note-taking	<i>Experimental Group</i>	19.24	1.788	30	9.211**
	<i>Control Group</i>	13.33	2.919	25	
Total score on all skills	<i>Experimental Group</i>	93.40	4.455	30	**21.143
	<i>Control Group</i>	67.13	4.743	25	

**significant at 0.01

Table 4. t-test results for experimental and control subjects on posttesting writing skills

Skills	Subjects	Mean Score	Standard deviation	no	t-value
Skill 1 Making a time plan	<i>Experimental Group</i>	18.2400	1.52414	30	**17.852
	<i>Control Group</i>	14.7667	1.83212	25	
Skill 2 Outlining & organization	<i>Experimental Group</i>	19.1200	1.61210	30	**9.159
	<i>Control Group</i>	13.2333	1.71561	25	
Skill 3 Coherence & Cohesion	<i>Experimental Group</i>	19.1200	2.06782	30	**14.750
	<i>Control Group</i>	13.0000	1.66633	25	
Skill 4 Style	<i>Experimental Group</i>	19.0000	1.87420	30	**14.815

	Control Group	12.7333	2.51661	25	
Skill 5 Presentation	Experimental Group	19.6000	1.88917	30	16.813**
	Control Group	13.5000	2.19848	25	
Total score on all skills	Experimental Group	94.0800	4.28858	30	**20.822
	Control Group	66.2333	4.69858	25	

**significant at 0.05

The above tables demonstrates that there are statistically significant differences between the mean scores of the experimental and control groups on reading and writing skills (1,2,3,4,5, and the total score) on posttesting to the advantage of the experimental group, thus the hypothesis is rejected as invalid.

A follow-up statistical test for post-comparisons was conducted for the writing test. The following table sums up the results:

Table 5. t-test results for a follow-up post comparisons for experimental subjects on re-administration 1 month later

Skills	Mean differences	SD for differences	Standard error for differences	t-value
Skill 1	.10000	.48066	.08776	1.140
Skill 2	.03333	.55605	.10152	.328
Skill 3	.03333	.49013	.08949	.372

Skill 4	.03333	.66868	.12208	.273
Skill 5	.20000	.61026	.11142	1.795
Total score	.33333	1.29544	.23651	1.409

The above table demonstrates that there are no statistically significant differences between post comparisons of the experimental subjects' mean score differences on the post-comparison follow-up t-test for the writing skills (1, 2, 3, 4, 5, and the total score) as upon re-administration of the same test one month later; this indicates reliability of the findings.

A follow-up statistical test for post-comparisons was conducted for the reading test. The following table sums up the results:

Table 6. t-test results for a follow-up post comparisons for experimental subjects on re-administration 1 month later

Skills	Mean differences	SD for differences	Standard error for differences	t-value
Skill 1	.16667	.64772	.11826	1.409
Skill 2	.26667	1.01483	.18528	1.439
Skill 3	.20000	.55086	.10057	1.989
Skill 4	.10000	.54772	.10000	1.000
Skill 5	.23333	.81720	.14920	1.564
Total score	.20000	.66436	.12130	1.649

The above table demonstrates that there are no statistically significant differences between post comparisons of the experimental subjects' mean score differences on the post-comparison follow-up t-test for the **reading** skills (1, 2, 3 , 4, 5, and the total score) as upon re-administration of the same test one month later; this indicates reliability of the findings.

To verify hypothesis three which states that “There are statistically significant differences between the attitudes of the experimental subjects on pretesting and posttesting”, the following table summarises the attitudes of the students towards learning English in an online environment:

Table 7. Students' Attitudes towards Learning English via Internet-based Instruction (IBI)

Attitudes towards Learning English via IBI	No (%)	Sometimes (%)	Yes (&)	Mean Score (SD)
Positive Attitudes	11.5	47.5	41	2.3 (0.7)
<i>Willing to Learn English online even if not required</i>				
<i>Increasingly absorbed in learning English online</i>	9	54	37	2.27 (0.6)
Negative Attitudes				2.18 (0.7)
<i>Study as required due to limited time</i>	18	47	35	
<i>Feel that extra work is limited in this environment</i>	47	37	16	1.79 (0.7)

As is shown in the above table, there are four items and for each one, experimental subjects were requested to choose one out of the following 3 options: No, Sometimes, and Yes. Option 1 was scored as 1, representing total disagreement and option 3 was scored as 3, representing total agreement. Responses to the first two items in the Attitudes Questionnaire show that the majority (88.5%) like to spend time improving their English even if they are not required to do so, and 91% become increasingly absorbed in learning

English as they do it online (mean score = 2.30, 2.27, respectively). This suggests that students have positive attitudes towards learning English.

Nevertheless, these students also have negative attitudes towards learning English as responses to the other two items show that 82% only study what is specifically required due to limited time, and 63% feel that doing extra work in English is not necessary (mean score = 2.18, 1.79, respectively). These findings seem to suggest that most subjects have ambivalent attitudes towards learning English in this environment. On the one hand, they like to improve their English, and become increasingly absorbed as they learn English; however, they are not willing to commit themselves when they are asked to spend extra time on English (Table 8). This finding was also confirmed by their low desire to use IBI to improve their English. All in all, the hypothesis is yet confirmed; students' attitudes towards learning English in an IBI environment has induced bettered attitudes towards learning the language.

Table 8. Spearman Correlation Coefficients for the Desire to Learn in an IBI Environment and their Attitudes towards English

Attitudes towards Learning English via IBI	No (%)	Sometimes (%)	Yes (&)	Mean Score (SD)
Positive Attitudes	11.5	47.5	41	2.3 (0.7)
<i>Willing to Learn English online even if not required</i>				
<i>Increasingly absorbed in learning English online</i>	9	54	37	2.27 (0.6)
Negative Attitudes				2.18 (0.7)
<i>Study as required due to limited time</i>	18	47	35	
<i>Feel that extra work is limited in this environment</i>	47	37	16	1.79 (0.7)

As depicted in Table 8, there is a strong correlation between the desire to use IBI as in this study's programme and their attitudes towards learning English.

Results from qualitative data gleaning techniques

Since the qualitative results of the implementation of this study relied on different techniques for gleaning subjective data, only marginal observations would be highlighted here.

E-logs

The following aspects of IBI-based process writing design had been checked over and over again in each session. Generally, the following webquest process aspects were observed to have been adequately attended to:

- Roles were well defined. It was clear who did what and when.
- Roles were integral to getting the job done, not just tacked on.
- Logistics were clear (e.g. it was clear how groups got formed)
- Enough resources were identified (Web or other) to convince one that the learners would have enough information to go on.
- Enough guidance was provided for activities in which learners interacted with each other (e.g. brainstorming) or with data (e.g. analyzing a photograph, interviewing an expert).
- There was enough specific guidance on how to produce/perform the task (e.g. suggested outlines, examples, formats).
- The Process matched the Task description.
- Consistent voice was used (addresses students as "you", not "the students".)
- Vocabulary was matched to the reading level of the audience.
- Bulleted and numbered lists were used to break up long paragraphs.
- Links were placed so as not to distract readers and cause them to click off to other

- sites prematurely.
- Long role-specific information was put onto separate pages.

Instructor's diary

This is the most lengthy and detailed section of qualitative data gathered during the implementation of the IBI instructional design. The diaries written were beyond the scope of reporting here; however, it would be helpful to note that the diaries were kept to tackle the daily circular evaluation of what happened in the IBI lab. Therefore, the researcher put into mind that diaries should report on several aspects, mainly:

Initiation of the session:

Mostly, there had always been a problem initiating the session related to starting time due to few students' fairly late arrival. That was a minor problem because a lot of work involving the model had to be done online at the participants' convenience, even at home. The computer lab capacity differed from one session to another. This was a frequent problem.

Organization of the class:

Collaborative working was executed by the instructor so that all group members would work together towards the instructional objectives of each session which were accessible to them in their group e-mails. Different inter-group roles such as group secretary, group leader, and group reporter were changed each session, which was a little bit time-consuming.

Discipline and Logistics:

Moving about in the lab was mostly difficult due to unavailability of enough chairs for all members in the groups. Also, some computers always needed to be checked before

getting to class each session. The subjects of the study were adults; therefore, it was pretty hard to interfere with their disciplinary conduct. That was left to them each time, trying not to impose certain disciplinary behaviours upon them unless it had to do with sabotage or uncleanliness in the computer lab.

Feedback:

As a way of getting started, observations from previous sessions were publicly noted, with group and individual feedback provided by the instructor in each session. Usually, students got on the right track easily after the first introductory sessions. Sometimes, private feedback was reported to students via their private emails or to groups via the group emails. Most students reported in their e-logs and in feedback emails to the instructor that they enjoyed learning in the lab. Several students noted that learning to read and write in an online environment was joyous and entertaining, with the least constraints of anxiety and time usually encountered in traditional learning environments.

Finalization of the session:

Students were told what to do in the final five minutes in each session. The instructor emailed the next timetable for the next session to groups. Assessments using the student assessment criteria developed first by the researcher and later by the students in their respective groups were used as group evaluation methods. Furthermore, e-portfolios were kept by individuals in their groups and as groups to be sent as cc to the instructor's email. Daily checks of these e-portfolios were done, mostly to check students' understanding of the tasks and their performances. Overall, the e-portfolios provided a vivid landscape of what was going on, indicating ongoing progress and development in the students' skills of process writing.

Discussion of Findings

Hypothesising that there are statistically significant differences between the mean scores of the experimental group subjects and those of the control group on pretesting as measured by Reading/ Writing Skills Test (1, 2, 3, 4, 5, and the total score), results showed that the hypothesis was confirmed, attesting to the evidence that all subjects in the control sample and the experimental sample, at the time the experiment was initiated, were all equal with regard to their competences in writing as investigated by the reading online/ writing test.

These results support the idea that the results of pretesting equating all subjects, control and experimental, may be buttressed by one's 'off-the-record' evaluation of the students, as possessing similar levels of reading / writing competency.

As for the hypothesis which reads: "There are no statistically significant differences between the mean scores of the experimental group subjects and those of the control group on posttesting as measured by online reading/writing skills test (1, 2, 3, 4, 5, and the total score)", as seen, the hypothesis was rejected as invalid now that there were statistically significant differences between the mean scores of the experimental and control groups on both reading and writing block skills (1, 2,3,4,5, and the total score) on posttesting to the advantage of the experimental group. This confirms our primary assumption that the instructional design merging online reading via webquests and email exchanges in instructional model described earlier in this paper is effective in developing the major skill areas of reading and writing competency. This finding meshes with previously established findings as regards the effectiveness of Internet-based Instruction (See the reviews by Cradle, 2003; (Andreas & Haythornthwaite, 2007). Findings from this study confirm previous research indicating the effectiveness of different e-learning

models; such IBI models of learning and instruction described the potential of email and webquests in online reading and writing to extend content knowledge and promote higher level thinking (e.g., March, 2000; Dodge, 2003a; Kortecamp & Bartoshesky, 2003; Menchaca & McVicker, 2003; Molebash & Dodge, 2003; Monroe & Orme, 2003; Draper, Smith & Sabey, 2004). Findings are also commensurate with the findings from Frazee (2004) who reported enhanced learning and time management in his study, given that the experimental students in this study did well on all aspects of writing skill areas, including time management.

Incorporating the use of e-mail with reported research findings on its efficacy as a medium for developing literacy has been an effective part of the instructional design tested in this study (Schwartz, 1990; Fey, 1994; Mabrito, 1991; Selfe, 1990). The findings are compatible with the established findings favouring the use of email as a medium and technique for developing writing skills (e.g. D'Souza, 1992, Anderson & Lee, 1995; Romiszowski & de Haas, 1989). Also, the collaborative groupings and teaming up of learners in the computer lab may have helped towards the success of the model.

Considering the last hypothesis which assumes a strong relationship between learning in an IBI environment and bettered attitudes towards English, the findings are commensurate with previous research which drew emphasis to the importance of attitudes in foreign language learning (Brown, 1994; Gardner, 1985), and the significance of IBI environments in improving attitudes towards language learning (Chen, 2004; Felix, 2001; Schnackenberg, 1997; Neu and Scarcella, 1991; Phinney, 1991; Thaipakdee, 1992).

Concluding Remarks

In conclusion, the research literature on emailing in writing and webquests in reading online is sparse and it is this researcher's opinion that more empirical studies examining the model are warranted. Hence, additional research is needed to examine the variations related to these important factors such as learning gains, more elaborate and controlled investigations into the effects of IBI on learning preferences changes.

For instance, one area for future research would include taking a more in-depth examination of the group process including individual factors that may affect collaborative work such as age, major, and preferences and expectations about teaching and learning methods (e.g. student reactions to the uncertainty that comes with ill-structured authentic, complex tasks).

In future studies, researchers who choose to use this same model could examine differences in engagement, including satisfaction and time spent on task, based on the area of focus for these specific online learning model tasks.

Further research using controlled studies on integrating webquests and e-mail correspondences need to further be conducted to check IBI effects on literacy development with various age groups and grades, on attitudes towards English and on the relationship between IBI and learning styles.

Conclusions extracted from this study confirm previously established research findings. Given the authentic classroom setting in which this study took place, the results are promising for educators interested in effective IBI design strategies and models as the one in this study which merges webquests and emailing, especially when integrated with e-mailing and collaborative learning, all in a problem-solving inquiry-based learning environment. Attitudes towards learning English can be enhanced in IBI environments,

especially when integrated with collaborative and enquiry-based learning.

References

- Allen, G. & Thompson, A. (1995). Analysis of the effect of networking on computer assisted collaborative writing in a fifth grade classroom. *Journal Of Educational Computing Research*. 12: 66-75.
- Anderson, D. (2002). Interfacing e-mail tutoring: Shaping an emergent literate practice. *Computers and Composition*, 19 : 71-87.
- Anderson, J. & Lee, A. (1995). Literacy teachers learning a new literacy. A study of the use of electronic mail in a reading education class, *Reading Research and instruction*, 34, 222-238.
- Andrews, R. & Haythornthwaite, C. (2007). *Handbook of e-learning research*. London: Sage.
- Baron, N. S. (1998). Writing in the age of e-mail: The impact of ideology versus technology. *Visible Language*, (32) 1 : 35-53.
- Bartoshesky, A. & Kortecamp, K. (2003). WebQuest: An instructional tool that engages adult learners, promotes higher level thinking and deepens content knowledge. In C. Crawford et al. (Eds.), *Proceedings of Society for Information Technology and Teacher Education International Conference 2003* (pp. 1951-1954). Chesapeake, VA: AACE.
- Benz, P. (2001). WebQuests for language learning. Online Resource. Retrieved May 18th, 2004 from: <http://www.ardecol.ac-grenoble.fr/english/tice/enwebquest2.html>
- Blase, J. D. (2000). A new sort of writing: E-mail in the English classroom. *The English Journal*, 90 (2): 47-51.

- Bolter, J. D. (1991). *Writing space: The computer, hypertext and history of writing*. Hillsdale, Nj; Lawrence Erlbaum Associates, publishers.
- Brown, H. D. (1994). *Principles of language learning and teaching*. New Jersey: Prentice Hall.
- Chen, P. (2004). *EFL student learning style preferences and attitudes toward technology-integrated instruction*. PhD Thesis. University of South Dakota DAI-A 64/08, 2813.
- Chiu, C. (2005). *Writing in English: Perspectives of an ethnic Chinese teacher and her students*. PhD Thesis, The University of New Mexico, AAT 3177082.
- Cross, T. and Raizman, M. (1986). *Networking: An electronic mail handbook*. Glenview: Scott, Foresman and Company.
- Dodge, B. (2003a). *WebQuests and other viruses*. Keynote presented at the Association for Educational Communications and Technology Conference, Anaheim, California, November 2003.
- Dodge, B. (1997). *Some Thoughts About WebQuests*. San Diego, CA: San Diego State University. Retrieved December 14, 2009 from:
http://webquest.sdsu.edu/about_webquests.html
- Dodge, B. (1998). *WebQuests: A strategy for scaffolding higher level learning*. On-line Resource. Retrieved December 14, 2009 from:
<http://edweb.sdsu.edu/webquest/necc98.html>
- Dodge, B. (2003b). A Webquests about Webquests. Online Resource. Retrieved December 14, 2009 from:
<http://webinstituteforteachers.org/~mrubin/webaboutweb/index.html>
- Draper, R.J., Smith, L.K., & Sabey, B.L. (2004). Supporting change in teacher education: Using technology as a tool to enhance problem-based learning. In N. Wentworth, R.

- Earle, & M.L. Connell (Eds.), *Integrating Information Technology into the Teacher Education Curriculum: Process and products of change* (pp. 25-42). New York: The Haworth Press.
- D'Souza, P. (1992). Electronic mail in academic settings: A multipurpose communications tool. *Educational Technology*, 32, 22-25.
- Felix, U. (2001). A multivariate Analysis of Students' Experience of Web Based Learning. *Australian Journal of Educational Technology*, 17(1), 21-36.
- Fey, M. (1994). Finding Voice through Computer Communication: A New Venue for Collaboration. *Journal of Advanced Composition* (14) : 221-238.
- Frazeo, J. (2004). *WebQuest design strategies: A case study measuring the effect of the jigsaw method on students' personal agency beliefs, engagement, and learning*. Ed.D. Thesis. University of San Diego and San Diego State University.
- Gardner, C. R. & Lambert, E. W. (1972). *Attitudes and Motivation in Second Language Learning*. Massachusetts: Newbury House Publishers.
- Gardner, R. C. (1985). *Social Psychology and Second Language Learning: The Role of Attitudes and Motivation*. London: Edward Arnold.
- Gay, L. (1996). *Educational Research: Competencies for analysis and applications*. London: Prentice-Hall.
- Harris, J. (1994). *Way of the ferret: Finding educational resources on the internet Eugene*. Oregon: International Society for Technology in Education.
- Hawisher, G. & Moran, C. (1993). Electronic mail and the writing instructor. *College English*, 55 (6): 627-643.
- Hawisher, G. & Selfe, C. (2007). On computers and writing. In Richard Andrews and Caroline Haythornthwaite (eds.), *Handbook of e-learning research*. London: Sage.

- Hiltz, S., Turoff, M. , and Linda Harasim (2007). Development and philosophy of the field of asynchronous learning networks: Facilitating collaborative learning. In Andrews R. and Haythornthwaite, C. (eds.), *Handbook of e-Learning Research*. London: Sage.
- Huff, C., Sproull, L, & Kiesler, S. (1989). Computer communication and organizational commitment: Tracing the relationship in a city government. *Journal of Applied Social Psychology*, 19, 1371- 1391.
- Kaufman, L. M. (1998). E-mail keypals in zone of proximal development.” *Paper Presented at the Annual Meeting of the Teachers of English to Speakers of Other Languages 32nd Seattle, A March, 17-12, 1998.*
- Kern, R. (1995). Restructuring classroom interaction with networked computers: Effects on quantity and quality of language production. *Modern Language Journal*, 79 (4):457-476.
- Killins, J. (2002). Educators collaborating across cultures in a computer-mediated communications environment. Unpublished Ed.D. Dissertation: Ontario Institute for studies in Education, University of Toronto, Canada.
- Lee, A. S. (1994). Electronic mail as a medium for rich communication: An empirical investigation using hermeneutic interpretation. *MIS Quarterly*, 1:43-57.
- Mabrito, M. (1991). Electronic mail as a vehicle for peer response. *Written Communication*, 8: 50 – 532.
- MacKay, W.E (1989) Diversity in the use of electronic mail: A preliminary inquiry. *ACM Transactions on Office Information Systems*, 6: 380- 397.
- March, T. (2000). WebQuests 101. *Multimedia Schools*, 7(5), 55-58.
- March, T. (2004). *Criteria for assessing best WebQuests*. Online Resource. Retrieved

December 14, 2009 from: <http://www.bestwebquests.com/bwq/matrix.asp>

- McComb, M. (1994). Benefits of Computer-Mediated Communication in College Courses. *Communication Education*, 43. 157-170.
- Menchaca, M. & McVicker, B. (2003). Scaffolding WebQuest creation for experienced teachers. In C. Crawford et al. (Eds.), *Proceedings of Society for Information Technology and Teacher Education International Conference 2003*. pp. 1025-1032. Chesapeake, VA: AACE.
- Molebash, P., & Dodge, B. (2003). Kick Starting inquiry with WebQuests and Web inquiry projects. *Social Education* 67(7), 158-163.
- Monroe, E. & Orme, M. (2003). *The nature of discourse as students collaborate on a mathematics WebQuest*. Comunicación presentada al NECC 2003.
- Neu, J. & Scarcella, R. (1991). Word processing in the ESL classroom: A survey of student attitudes. In P. Dunkel (ed.) *Computer-assisted language learning and testing*, pp. 169-187. Rowley, MA: Newbury House.
- Omaggio-Hadley, R. (1993). *Teaching language in context*, 2nd edition. Boston: Heinle & Heinle.
- Phinney, M. (1991). Computerized writing and writing apprehension in ESL students. In: P. Dunkel (ed.) *Computer-assisted Language Learning and Testing* pp.189-204. Rowley, MA: Newbury House.
- Reinking, D. (1994). *Electronic literacy (Perspective in Reading Research No. 4)*. Athens, GA: National Reading Research Center, Universities of Georgia and Maryland.
- Reiss, D. (1996). Electronic Toads: Computers and Writing in Introductory Literature. *Paper presented at the Annual meeting of the Conference on College Composition and Communication 47th Milwaukee, WI, March 27-30, 1996.*

- Rivers, W. (1975). *Teaching Foreign Language Skills*. 2nd edition. Chicago: UCP.
- Rodrigues, E. (2007). *Students' communicative practices around writing and digital technologies: An ethnographical study of an inner-city high school in the northeast of Brazil*. PhD Thesis, Michigan State University, AAT 3264222.
- Romiszowski, A. J. & de Haas, J. A. (1989). Computer mediated communication for instruction: Using e-mail as a seminar. *Educational Technology*. 29. 7-14.
- Schaefer, D. & Dillman, D. (1998). Development of a standard e-mail methodology: Results of an experiment. *The Public Opinion Quarterly*, 62 (3): 378-397.
- Schnackenberg, H. L. (1997). *Learning English electronically: Formative evaluation in ESL software*. ERIC Document Reproduction Service No. ED403877.
- Schwartz, J. (1990). Using an electronic network to play the scales of discourse. *English Journal*. 69. 16-24.
- Selfe, C. L. (1990). Technology in the English classroom: Computers through the lens of feminist theory. In C. Handa (Ed). *Computers and community teaching composition in the twenty-first century*. Portsmouth, NH: Heinemann.
- Smith, H. (1999). *Opportunities for information and communication technology in the primary school*. Stoke on Trent: Trentham Books Limited.
- Smolensky, M.W, Carmody, M.A. & Halcomb, C.G. (1990). The influence of Task Type, Group Structure and Extroversion on Inhibited Speech in Computer Mediated Communication. *Computers in Human Behavior*. 6. 261-272.
- Snyder, I. (2007). Literacy, learning and technology studies: Challenges and opportunities for higher education. In Andrews R. and Haythornthwaite C. (eds.), *Handbook of e-Learning Research*. London: Sage.
- Spitzer, M. (1989). Computer Conferencing: An emerging technology. In Hawisher G. F.

- & Selfe C. L. (Eds.). *Critical Perspectives on Computers and Composition Instruction* (pp. 187- 200). New York: Teachers College Press.
- Stein, L. D., & Stein, A. P, (1995). Electronic databases & information services: Integrating their use into the requirements of your syllabus. *Instructionally Speaking*, 6(2), 3-4.
- Tao, L. and Reinking, D. (1996). What research reveals about e-mail in education. *Paper presented at the Annual Meeting of the College Reading Association, 40th Charleston SC Oct 31-Nov.3, 1996.*
- Tao, L. (1995). *What do we know about e-mail: An existing and emerging literacy vehicle?* Paper presented at the Annual Meeting of the National Reading Conference. New Orleans, LA, November 29-December 2, 1995.
- Thaipakdee, S. (1992). *Relationship among writing quality, attitudes toward writing, and attitudes toward computers in a computer-mediated technical writing class for English as a foreign language students.* PhD Thesis. University of North Texas. *DAI*, A:53-1135-A.
- Traw. R. (1994). School/University collaboration via E-mail: A unique approach to caching reading and language arts. *Tech Trends*. 39. 28-31.
- Vygotsky, L.S. (1978). *Mind in society: The development of higher psychological processes.* Harvard University Press.
- Warschauer, M. (1995). *E-mail for English teaching.* Alexandria, VA: TESOL.
- Warschauer, M. (1996). Comparing face to face and electronic communication in the second language classroom. *CALICO Journal*, 13 (2), 7-26.
- Weinburgh, M. H. (1998). Gender, ethnicity, and grade level as predictors of middle school students' attitudes towards science. Online Resource. Retrieved February 11th

2011 from:

www.Ed.Psu.Edu/Ci/Journals/1998aets/S5_1_Weinburgh.Rtff

Whitworth, A. (2007). Researching the cognitive cultures of e-learning. In Andrews R. and Haythornthwaite C. (eds.), *Handbook of e-learning research*. London: Sage.

The Impact of Omani Twelfth-Grade Students' Self-Assessment on their Performance in Reading in English

Amin Saleh Moheidat & Abdallah Ahmad Baniabdelrahman

baniabdelrahman@yahoo.com

Bio Data:

Amin Saleh Moheidat is an MA holder in TEFL employed by the Ministry of Education in Jordan. He is currently seconded to the Ministry of Education of Oman. His research interests include methods of teaching English as a foreign language.

Dr. Abdallah Ahmad Baniabdelrahman is an associate professor in the Curriculum and Instruction Department at Yarmouk University, Jordan. He is currently the chair of the English Language Skills Department at The Preparatory Year Deanship at King Saud University in Saudi Arabia. He has published widely in the field of English language teaching.

Abstract

This study is aimed at investigating the impact of Omani twelfth-grade students' self-assessment on their performance in reading in English. The sample of the study consisted of two groups, an experimental group and a control group, of 39 students each. Both groups were selected in Moosa bin Nusseir School for General Education in Muscat. A reading test adapted from previous General Certificate Exams in the Sultanate of Oman was developed. To collect data about the students' progress, the researchers used the technique of students' self-assessment through one-minute papers and rating-scale sheets. The findings of the study revealed that the students' self-assessment had a positive effect on their performance in reading in English. In light of the findings, it is recommended

that students be trained on how to use self-assessment to improve themselves. It is also recommended that teachers be aware of the positive effect of students' self-assessment on students' performance in reading.

Keywords: Formative Assessment, Self-Assessment, Authentic Assessment, Autonomous Learning, Strategies of Reading

Introduction

In literature, the terms assessment and evaluation are sometimes used interchangeably (Collins & O'Brien, 2003). However, in this study they are used differently. Assessment, in the broad sense, means "any method used to better understand the current knowledge that a student possesses" (Collins & O'Brien, 2003, p.29). According to Crooks (2001), assessment is "any process that provides information about the thinking, achievement or progress of students" (p. 1).

Because assessment is important in teaching and learning, every teacher should assess his/her students' learning regularly. Some of the methods which teachers use to measure their students' learning are written tests, book reports, research papers, homework exercises, oral presentations, and question-and-answer activities. Therefore, teachers spend a great deal of their class time engaged in one type of assessment or another (Stiggins, 2001). On the other hand, assessment of students entails using a well organized system, namely tests, to make judgments about the students' achievement (Gronlund & Linn, 1990). These tests bring anxiety, fear, or disappointment to students which might negatively affect their language learning. Students might suffer from spending long hours of study and preparation for tests haunted by mixed feelings of hope and fear.

Unfortunately, they might be disappointed at the fact that what they concentrate on differs from what the teacher emphasizes on in his/her questions (Guskey, 2003; Shaaban, 2005).

Any learning system needs feedback (Davis, 1998). Sufficient data need to be collected about each student's learning proficiency to take whatever action needed to meet each student's learning needs and to adjust the system. To reach this end, different kinds of assessment could be used to provide reliable information about the students' learning progress. Assessment has "the most powerful influence on student learning" (George & Cowan, 1999, p.8). Therefore, teachers should always keep enough, and accurate, information on which they can build their judgment to improve their students' performance (Shaaban, 2005).

There are two types of assessment: formative assessment, that is, assessment for learning, and summative assessment, or assessment of learning (Stiggins, 2001). In fact, authentic assessment requires both. Summative assessment takes place at the end of a term or a course and is used to provide information about how much students have learned or how well a course has worked (Gipps, 1994). That is to say, a test is usually given at the end of a term, semester or year, the purpose of which is to measure proficiency (Boston, 2002).

On the other hand, formative assessment takes place during a course of teaching and is used essentially as feedback to the teaching-learning process (O'Malley & Pierce, 1996). In other words, it is an ongoing process of collecting information about the students' performance through various techniques of classroom assessment. The purpose of formative assessment is not only to measure proficiency, but also to improve it as well.

Formative assessments do not bombard students with questions to be answered within a

time limit. On the contrary, they “reflect the concepts and skills that the teacher emphasized in class, along with the teacher’s clear criteria for judging students’ performance” (Guskey, 2003, p.8). An assessment is summative when the intention is mainly to give a final judgment on students’ achievement. Formative assessment, or assessment for learning as it is sometimes called, is a continuous process in which the main purpose is beyond measurement; rather it is to help students improve. Summative assessment, or assessment of learning, on the other hand, tends to be an end point, usually expressed in grades and concerned with making judgments.

Students’ involvement in the teaching-learning process is important for there is much evidence to suggest that students’ self-assessment helps improve their performance (Brantmeier, 2005; Falchikov & Goldfinch, 2000; Shaaban, 2005). Self-assessment does not mean that students are allowed to assess themselves in the form of grades; instead they get continuous feedback on their progress to help both the students and the teacher.

One of the techniques used in assessment is the One-Minute Paper which provides very useful information. Angelo and Cross (1993) argued that the One-Minute Paper is the most used technique in assessment since it provides useful and quick feedback on student learning. They also claimed that over one semester, different studies proved that students in classes where minute papers were used out-gained those in classes without minute papers.

The information collected through formative assessment is used to detect the strengths and weaknesses of the learners for the purpose of improving proficiency (Collins & O’Brien, 2003; Shaaban, 2005). It provides useful information for both the teacher and the student upon which appropriate action can be taken (Guskey, 2003; Shaaban, 2005). Both forms of classroom assessment, formative and summative, are needed to determine

how much learning has occurred.

Black and Wiliam (1998b) defined the assessment process to include all activities undertaken in class, either by teachers to assess their students or by the students to assess themselves, which can be used as feedback to adjust the teaching-learning strategies. According to this definition, assessment includes teacher observation, classroom discussion, marking tests and collecting information from the students themselves about their own learning; namely *students' self-assessment*.

Students' self-assessment is considered to be one of the most important formative classroom assessment techniques. One of the purposes of this technique is to improve the quality of students' learning. It can also lead to modifications when teaching strategies have not met the required learning outcomes. Some educators have argued that students often find external assessment by teachers or supervisors unjust. Therefore, if students are given the chance to assess themselves, they will be more confident to give more accurate information about their progress, (Angelo & Cross, 1993).

By assessing their own learning, students can increase their awareness of what is happening in class (Gipps, 1994). Perhaps the most important factor of a successful teaching-learning process is active student involvement (Stiggins, 2001). This requires teachers to provide their students with feedback and teach them to use it effectively for learning. Consequently, students can learn how to assess themselves so that they can learn what they need to do in order to achieve success (Black & Wiliam 1998b). Hence, successful formative assessment depends on active student involvement.

Regarding reading comprehension, some researchers (The National Capital Language Resource Center NCLRC, 2004; Abraham, 2005) believe that it is a bit difficult to assess

it accurately. That is simply because “reading is a complex behavior composed of many skills” (Salvia & Ysseldyke, 1988, p. 354). It cannot be assessed in isolation from other skills (Rivas, 1999). In other words, a number of factors have to be taken considered when assessing reading, among which are the purpose of reading, the overall linguistic level of the student, and the role of previous knowledge. Assessment of reading ability depends mainly on the purpose of reading. “Reading for a purpose provides motivation – an important aspect of being a good reader (Grabe, 1991, p. 378).”

What to Assess in Reading?

So what do teachers assess in reading? They assess a number of sub-skills relating to reading. According to the Omani Ministry of Education (MOE) (2003-2004), teachers assess whether students are able to comprehend the general idea of a text; recognize the type of a text, e.g. interactive, informative, narrative, or evaluative; arrange the sequence of information in a text; use pre-reading activities to predict what a text would be about; guess the meanings of particular words from context; extract specific information from a text; and use different reading strategies, e.g. skimming, scanning, speed reading, paced reading, or timed reading.

The MOE seeks the implementation of all the above skills in its schools in order to improve students’ achievement in reading. The students’ command of those skills means that the MOE has succeeded in fulfilling its goals.

The MOE in the Sultanate of Oman has started a new evaluation system in General Education Schools: formative assessment which concentrates on self assessment, seeking improvement in students’ performance. This study is an attempt to investigate the impact of students’ self-assessment on their performance in reading in English. Moreover,

throughout the researchers' experience of teaching English, they have very often come across students whose reading skills were below average. Therefore, they decided to investigate the issue more closely.

The Educational System in Oman

The educational system in Oman is very similar to the educational systems in many other Arab countries. A child joins school at the age of 6 and graduates from school after 12 years of study.

A lot of money is spent each year on education in the Sultanate of Oman, yet there is a great dissatisfaction with the educational system. Efforts to improve the system have given much attention to the role of assessment in students' performance. Many educators have started to think of new measures, other than the traditional ones, to monitor the students' performance in English on the basis that traditional measures have failed to achieve significant learning outcomes.

In the past four years or so, the educational system in the Sultanate of Oman has seen great advances in understanding of how students learn. At the same time, tremendous developments have been done to support learning and to lead to better outcomes.

To meet the demands of the 21st century, considerable changes have been undertaken to reform the educational system. The MOE has launched the Basic Education System in 1998, seeking improvement in the quality rather than quantity of teaching and learning (Ministry of Education in Oman, 2005). This new system is composed of two phases: 'Basic Education' and 'General (Secondary) Education'. Basic education consists of two cycles: Cycle One which includes Grades 1 – 4 and was implemented in 1998 and Cycle Two which includes Grades 5 – 10 and was implemented in 2001. The secondary

education phase, also called General Education, extends for 2 years and includes Grades 11 - 12 (Sasidharan, 2005). Within this new system learners are given a more active role in the learning process than ever before.

Teaching and Assessing English in Oman

Because of the importance of English the ministry of education in Oman (MOE) introduced a new educational system in which English started to be taught from grade 1 instead of 4. This new system has brought to light new continuous assessment techniques which are used now to assess students' achievement. The Educational Evaluation Department in the MOE in Oman has introduced a new system of Continuous Assessment (CA) in the secondary level in the beginning of the academic year 2004 – 2005. This new system of CA is currently implemented along with the other formal assessment techniques used, namely semester tests. CA is meant to continuously measure the students' needs through “assessment of attainment..., measurement of the value of teaching methods and procedures, and diagnosis of individual or group difficulties” (Sasidharan, 2005, p. 60).

There are two English textbooks that are taught to Grade 12 students; ‘Core English’ and ‘Elective English’. Core English is compulsory to all twelfth-grade students, whereas Elective English is taught only to those who choose it, for twelfth-grade students have the chance to choose Elective English amongst other subjects.

The ‘Core English’ *Our World Through English (OWTE)* consists of two books: the *Activities Book (AB)* and the *Pupil's Book (PB)*. Both books are produced by the English Language Curriculum Department (ELCD). Moreover, ‘Core English’ is taught through six periods per week of 45 minutes each. On the other hand, the ‘Elective English’

consists of one textbook, Reading and Writing Targets 3, which is published by Express Publishers. It is taught through three periods per week of 45 minutes each.

In spite of all the efforts done by the Omani Ministry of Education to improve students' abilities in reading, the researchers have very often come across students whose reading skills were below average. Therefore, they decided to investigate the issue more closely.

Statement of the Problem

Although classroom assessment has been given much attention in the evaluation system in general education schools in the Sultanate of Oman, the performance of students is still not at the desired level. The researchers believe that neglecting the students' role in assessments may be one of the main factors behind this low performance. Moreover, students' self-assessment, in particular, has so far been completely disregarded, even though it may have a powerful and positive effect on students' learning.

In general education schools in the Sultanate of Oman, students are solely assessed by teachers; which is supposed to improve the students' performance. However, students' performance is far from what is expected. Class size has its negative effect on learning and teaching. "How much is a teacher successful?" is a typical question in such situations. In classes greater than 20 students, teachers cannot always be clear of what their students' needs are. That is simply because learning depends on what teachers and students do in class; teachers have to cater for 35 students or more in each class in order to help them become better learners (Black & Wiliam, 1998b). That is why students should play their role to help themselves become better learners. Thus, self-assessment increases the role of the student; however, the teacher remains the main source of assessment (Borg, 2006).

Purpose and Questions of the Study

This study addresses the following three questions:

1. Is there any significant effect of Omani Twelfth-Grade male students' self-assessment on their performance in reading in English?
2. Is there any significant difference in variance of achievement in the two groups?
3. What size is the effect of self-assessment on reading?

Hypotheses of the Study

This study hypothesized that:

1. There is no significant difference ($\alpha = 0.05$) between the mean scores of the Omani twelfth-grade male students who apply self-assessment and the mean scores of those who do not apply self-assessment in the English reading performance test ($H_0: \mu_E = \mu_C$).
2. There is no significant difference in variance of achievement in the two groups ($H_0: \sigma_E^2 = \sigma_C^2$).
3. The difference between the two means is not practically significant.

Significance of the Study

Studies investigating the impact of students' self-assessment on their performance in reading in English in Oman are rare and non-comprehensive to the best of the researchers' knowledge. There were two studies investigated by Al-Jardani, Khalid and

Al-Balushi (2002). Al-Jardani investigated the evaluation of the effectiveness of self-assessment in teaching English to young learners, whereas Al-Balushi investigated the EFL teachers' perceptions of alternative assessment tools used in the first cycle of Omani Basic Education Schools. Therefore, this study is one of the first studies to deal with this technique. If the results of the study are in favor of the experimental group, then it will shed some light on the importance and the value of this technique in improving students' performance in reading in English. Thus, the Educational Evaluation Department in the MOE in the Sultanate of Oman will benefit from it and it might encourage them to launch their 'Self-Assessment Project', which, as they say, "... is ... now very much in the documents and on the agenda" (O'Reilly & Al-Lawatia, 2005, p. 23).

It might also urge English teachers in the Sultanate of Oman to think about trying this particular classroom assessment technique in their classrooms. It could help them decide on the suitable strategies to use for particular lessons.

Moreover, students themselves can benefit from the results of this study. The results can help students detect their strengths and weaknesses. As a result, they can guide themselves to overcome their weaknesses.

The lack of research investigating the impact of students' self-assessment on their performance in reading in English in Oman also raises the need to conduct further studies.

Population and Sample of the Study

The population of the study consisted of all male twelfth-grade students in general education schools in the Muscat Region, Sultanate of Oman during the academic year 2006/2007. More specifically, there were 4010 students. The sample of the study

consisted of two 12th grade sections (39 each) from Moosa bin Nusseir Secondary school in Muscat. One section was assigned as the control group and one as the experimental group.

Design and Procedures

This study is a quasi experimental study. Two methods of assessment were compared.

They were:

- a. Self-Assessment method
- b. Traditional method of assessment

One section of the sample was assigned as the experimental group and the second section as the control one. The experimental group consisted of 39 students. The students in that section were taught by one of the researchers, using the technique of students' self-assessment, along with other traditional techniques of classroom assessment. Certain methods to measure how well students have learned the taught material were used. A One-Minute Paper at the end of each reading class period was used (Appendix 2) and a Rating-Scale Sheet at the end of each topic/unit (Appendix 3) as quick diagnostics that helped students reflect on their learning and gave the researcher immediate feedback. The control group also consisted of 39 students. The students in that group were also taught by the same researcher, using only the traditional techniques of classroom assessment.

In order to examine the equivalence of the experimental and control groups, the researchers developed and administered a reading test before the subjects were exposed to the treatment. Table 1 shows the results of the *t-test* of the students' scores in the test of equivalence.

Table 1:
Results of the t-test of the Students' Scores in the Reading Comprehension Test before the Study Started

Source	N	Mean	SD	St Err	DF	T-Value	P
Experimental Group	39	48.00	26.49	4.24	76	-0.191	0.85
Control Group	39	49.13	25.54	4.09			

Table 1 shows that there was no statistically significant difference at $\alpha=0.05$ between the mean scores of the experimental and the control groups in the reading comprehension test before the study started (t-value = -0.191, P = 0.85). This means that the two groups were equivalent. The mean of the control group was 49.13 with a standard deviation of 25.54, and the mean of the experimental group was 48.00 with a standard deviation of 26.49.

Instrument of the Study

Since this study aimed at investigating the effect of students' self-assessment on their performance in reading in English, the researchers developed a reading test which was conducted before and after the treatment as an instrument to achieve the purpose of the study (**Appendix 4**). This test was adapted from previous GCE in the Sultanate of Oman. The researchers made the necessary changes to fit the purpose of the study. The test in its final form consisted of four different reading questions. Question One consisted of 6 items, worth 2 marks each. The students had to match each item to the left with the suitable text to the right. There were 4 extra texts to the right. Question Two consisted of 6 items, worth 3 marks each.

The students had to read the 6 short texts given and then to write one word to complete the sentences which follow each text. Question 3 consisted of 9 items, 2 marks each. The students had to read the given text and then to answer the 9 multiple items which followed. Question 4 consisted of 4 items, 3 marks each. The students had to read a given text then to tick the items given according to whether they were TRUE, FALSE, or DOESN'T SAY. 'Doesn't Say' means that the piece of information is not mentioned in the text. The total score of the test was out of 60, which was converted later to 100 to suit the statistical calculations. The time allotted for the test was 60 minutes.

Validity of the Instrument

To guarantee the validity of the instrument, it was given to a jury of TEFL specialists; three professors in the English Department at Sultan Qaboos University, two school supervisors at Muscat Directorate of Education, and three experienced EFL teachers working in the schools of the MOE, Sultanate of Oman. The jury examined the test and provided their comments and suggestions for modifications to fit the purpose of the study. They evaluated the test according to language, clarity, appropriateness of the questions to the level of the students, relevance to the skill it meant to test, and timing. The jurors' comments and recommendations were appreciated and taken into consideration. Necessary modifications were made accordingly. The test in its former shape consisted of 5 questions of 30 items. The majority of the jurors suggested that the number of questions should be reduced to 4 questions in order to fit the allotted time for the test. Other modifications concerning structure, words choice, and layout of the test

were recommended. Moreover, question one had two distractors in the box. The jurors recommended that one more distractor be added. In its final form, the test consisted of 4 questions of 25 items as mentioned in the instrument section above.

Reliability of the Instrument

To establish the reliability of the instrument, a pilot study was conducted in a neighboring school, Al Khaudh School for General Education, three weeks before the pre-test was given to the two groups. Thirty-five male students of twelfth-grade sat for the exam. The researcher conducted the test for the first time on Sunday, 7th November, 2006. Before the test started, the researcher gave some explanatory notes and instructions to the students about the test. The students were given enough time to answer the four questions of the test. The researcher noticed that 60 minutes was enough even for average students.

During the test, the researcher noticed that the students showed considerable interest in answering the questions. It could have been because the test was some kind of practice for them since the final examinations were approaching. That is why they took it seriously. To guarantee authentic results, two teachers, as invigilators, were in the room, including the researcher. At the end of the test, the papers were collected and marked.

Three weeks later, precisely on the 2nd of December, 2006, the same test was conducted to the same subjects for the second time. The same procedures were followed to conduct the test for the second time. This was necessary to ensure the

validity of the results. To show the internal consistency of the test, the reliability coefficient was computed. Its value was calculated and turned to be $r = (0.83)$.

Data Collection

To collect the data for this study, the researchers adapted a One-Minute Paper in which students had to give short answers to three questions given (Appendix 2). The One-Minute Paper had a typical format. This format was that students were given three short questions. They had to reflect on their learning by answering those questions in brief. The researchers had the questions ready on paper beforehand. This One-Minute Paper was adapted from Angelo & Cross (1993). Some modifications were made to meet the need of the study. This One-Minute Paper helped students to reflect on the reading lessons which were taught by the researcher. It also gave the teacher immediate feedback about the students' learning. The researchers decided to use the One-Minute Paper at the end of each class period because they wanted to check the students' understanding of what had been taught in that particular class period. Then at the end of each topic/unit, they decided to distribute a Rating-Scale Sheet to check the students' understanding of the topic as a whole (Appendix 3). Both, the One-Minute Papers and the Rating-Scale Sheets, were pre-printed in order to save time. The researchers gave their students time to respond. They sometimes asked them to discuss their thoughts in pairs. The purpose of this instrument was to let the students 'think aloud' which provided both the teacher and the students with true data about the students' learning outcomes. Therefore, after analyzing the students' feedback, if the researcher felt that the class had failed to understand a certain point, he could try

another teaching strategy for the next class to teach that particular point again. To achieve this goal, the researcher collected the students' feedback and sorted it into major topics in order to have a clear idea of what to discuss in the next class. For example, some students stated that they faced problems with referencing tasks. Others said that they faced problems with guessing meaning of new vocabulary from context. A third group declared that they encountered difficulty in finding the main idea of a given paragraph. The researcher had to take the necessary procedure in order to help those students to overcome their problems.

To ensure the appropriate use of the One-Minute Paper, the researchers trained the subjects of the experimental group for two weeks, before the treatment, on how to self-assess so that they were able to use the One-Minute Paper correctly and effectively. That was necessary so that they were able to give accurate information about their learning. The researchers undertook two actions to ensure that the use of the One-Minute Paper was going smoothly. The first action was to prepare printed sheets so that they could distribute them to the students more easily. The second action was to emphasize the reason behind using the One-Minute Paper and what it is hoped to achieve.

The researchers conducted a test twice on both groups; before the treatment to check the equivalence of the two groups, and after the treatment to check if there was any significant difference between the two groups concerning performance in reading attributed to the treatment.

The study lasted for 13 weeks, starting from the 17th of February 2007 to the 16th of May of the same year, during which the researchers used the technique of students' self-assessment in teaching the subjects of the experimental group to

monitor their progress in reading. Nine reading texts of different topics were included in the study. At the end of the study, a post-test was administered to both groups to investigate the students' achievement in reading. The test was marked by the researchers and the results, along with the results of the pre-test, were analyzed, using the Statistical Package of Social Sciences (SPSS). Means and standard deviations were calculated to find out whether there was any significant difference between the reading achievements of the experimental group compared to the control group.

Data Analysis

To analyze the collected data:

1. A Two tailed t-test of independent samples was used to compare the means of the two groups.
2. An F-test was used to measure the effect of students' self-assessment on the achievement in individual differences.
3. The effect size equation was used to check the significance of the practical effect of students' self-assessment along with the statistical significance through the

level of improvement in standard deviation:
$$\Delta = \frac{\bar{X}_E - \bar{X}_C}{\left(\frac{\sigma_E + \sigma_C}{2}\right)}$$

Table 2 below presents the means and standard deviations of the students of the experimental and control groups after the treatment.

Table 2

Results of the t-test of the Students' Scores of the Two Groups after Treatment

Group	N	Mean	SD	DF	t-Value	P	F
Experimental	39	65.74	18.31	76	2.23	*0.0287	1.77
Control	39	54.87	24.34				

* $P < 0.05$

Table 2 shows that there are differences between the mean scores of the students of the experimental and control groups. The mean score of the experimental group is (65.74) with a standard deviation of (18.31) while the mean score of the control group is (54.87) with a standard deviation of (24.34). In order to test if the difference is significant, a t-test of independent samples was run. The results show that there is a significant difference between the reading achievement of the students of the experimental group over the students of the control group after treatment due to the method of assessment at $\alpha = 0.05$ ($T = 2.23$, $P = 0.0287$). This result supports the alternative hypothesis of the study ($H_1 : \mu_E \neq \mu_C$) which says that there is a significant difference between the two groups. Therefore, the null hypothesis ($H_0 : \mu_E = \mu_C$) is rejected.

Comparing the average scores of the subjects involved in the study to the scores of the students in the control group on the same test, F was found to be around 1.77. This indicates that the implementation of students' self-assessment had a positive effect on the students' performance in reading at $\alpha = 0.05$. This result supports the alternative

hypothesis of the study ($H_1 : \sigma_E^2 \neq \sigma_C^2$). Therefore, the null hypothesis ($H_0 : \sigma_E^2 = \sigma_C^2$) is rejected.

The effect size equation was used to check the significance of the practical effect of students' self-assessment along with the statistical significance through the level of improvement in standard deviation. The effect size was calculated $\Delta = \frac{\bar{X}_E - \bar{X}_C}{\left(\frac{\sigma_E + \sigma_C}{2}\right)}$ and

found to be 0.5, which is significant at $\alpha = 0.05$.

Results and Discussion

The findings of the study proved that due to the implementation of the self-assessment technique there was a statistically significant difference between the mean scores of the performance of the subjects of the experimental group over the period of the study compared to the mean scores of the subjects of the control group. Moreover, the results of the study revealed that there is significant difference in variance of achievement between the two groups. The effect size of self-assessment on reading is practically significant.

It took students about two weeks from the first introduction of the One-Minute Papers and the Rating-Scale Sheets to learn how to use them effectively to assess themselves. At the beginning, many students were reluctant to express themselves. They wrote very few and vague comments. Some weak students even did not like the idea of revealing their failure in comprehending certain texts. Sometimes they wrote that they understood when they really did not, which was similar to the “nod” mentioned earlier. The researchers had to spend some time with those students to explain to them that the technique aimed at helping the teacher to determine the students' problems so that he can handle them properly and to be aware of the students' strengths in order to reinforce them. However,

as time went on, they became more confident and so their comments became clearer, more accurate, and more to the point.

The results showed that almost all students understood the aim behind the use of the technique of self-assessment which, in turn, helped the researchers to collect accurate information about his students. This in itself was an important achievement. Angelo and Cross (1993) argued that it is important for students to understand the purpose of a continuous assessment technique and to be aware of their role within it in order to be able to be active participants.

The second point worth noting is that the researchers noticed that the use of the One-Minute Papers and the Rating-Scale Sheets was very useful, especially for students who were shy and reluctant to voice what they did not understand.

One more significant advantage to self-assessment was that it helped to show the students that there were other students who shared such problems with them. This meant that they were not alone in the arena facing certain difficulties. This fact helped reluctant students to freely reveal their problems which were kept as sacred secrets for quite some time.

During the study, the researchers found out that students' self-assessment had several advantages. He found out that it did lead to improvement in students' learning, as the results of the post-test show. It encouraged students to participate more openly which was clear from the researchers' day-to-day observation. It directed students to give more accurate and honest answers, especially when they felt that they would not be judged by marks. It also helped students to direct their own efforts more effectively which was clear in the paper work which they had submitted, especially towards the end of the study. More importantly, it made students more involved and motivated in the learning process

which was very clear in their participation inside and even outside the classroom. They started to feel that they have a role to play and when they play it right they gain more. That is to say, when they were able to assess themselves correctly and give accurate feedback to the teacher, the teacher played his role in helping them with their problems. Consequently, they started to overcome their reading problems and their reading comprehension improved. That is why they started to participate in the morning broadcast more effectively; they started preparing class magazines on their own; they even took the risk to prepare and organize a book exhibition in the school; they became more active in the English Club activities that were held inside or outside the school; and they started to visit the library more often.

There were 22 failures in the experimental group before treatment. However, the number decreased into 9 failures, whereas, the number of failures in the control group decreased by 1 only (from 21 failures to 20). Moreover, the increase in the scores of the individual low achievers in the experimental group was higher than in the case of the control group. The increase in the experimental group ranged between 16 and 35 marks, whereas it ranged between 3 and 12 marks only in the control group. This indicates that the low achievers in the experimental group made a lot of progress due to the implementation of the technique of students' self-assessment.

Comparing the raw scores of the high achievers in both groups, before and after the treatment, it is noticed that the subjects of the experimental group gained much progress in comparison to the subjects of the control group. However, comparing the increase in the scores of the high achievers and the low achievers of the same group, the experimental group, the researchers found out that low achievers gained a much higher increase in their scores. This indicates that low achievers could benefit more from the

implementation of the technique of students' self-assessment, not to mention that the high achievers could benefit from it as well.

References

- Abraham, E. (2005). Habits die hard: Inculcating reading habits. In *Proceedings of the Fifth National ELT Conference, 24-29*. Sultan Qaboos University, Muscat 20-21, April, 2005.
- Al Abdelwahab, S. (2002). *A qualitative investigation of portfolio self-assessment practices in an intermediate EFL classroom, in Saudi Arabia*. ProQuest Dissertation and Theses 2002. Section 0168, Part 0279 312 pages; [Ph.D. Dissertation]. United States, Ohio: The Ohio State University. Publication Number: AAT 308 1894.
- Al Belushi, K. (2002). *EFL Teachers' Perceptions of Alternative Assessment Tools Used in the First Cycle of Omani Basic Education Schools*. Unpublished Thesis, Sultan Qaboos University, Sultanate of Oman.
- Al Issa, A. (2006). The cultural and economic politics of English language teaching in Sultanate of Oman. *Asian EFL Journal*, 8 (1). Retrieved March 2, 2007, from: http://www.asian-efl-journal.com/March_06_asmi.php
- Al Mazroui, M. (2005). *Teaching English to Higher Grade Students*. Muscat Message, Ministry of Education. Muscat: Al Nahda Printing Press. 36-38.
- Al-Shaihani, S. (2002). *The Effect of Using Metacognitive Strategies on the Reading Comprehension of Omani First Secondary Female Students*. Unpublished Master Thesis, University of Jordan, Jordan.
- Al Zidjali, F. (2005). Professional development: A self-Based approach. *Muscat Message*,

- Ministry of Education. Muscat: Al Nahda Printing Press, 3-8.
- Angelo, T., & Cross, K. (1993). *Classroom Assessment Techniques: A Handbook for College Teachers* (2nd ed.). San Francisco, CA: Jossey-Bass.
- Atkinson, P. (2003). *Self-assessment leads to more progress*. Retrieved May 10, 2007, from <http://www.literacytrust.org.uk/Database/assessment.html>
- Bartholomae, D. & Petrosky, A. (Eds.). (2002). *Ways of Reading* (6th ed.). Boston: Martin's. Retrieved February 7, 2007, from: <http://www.csupomona.edu/~lrc/crsp/techniques.html>
- Ben Zid, M. (2005). How a poem should be read to attain relevance: Stylistic, intuition and the buried life of words. *In Proceedings of the Fifth National ELT Conference*, 40-45. Sultan Qaboos University, Muscat 20-21, April, 2005.
- Black, P.; Harrison, C.; Lee, C.; Marshall, B. & William, D. (2004). Working inside the black box: Assessment for learning in the classroom. *Phi Delta Kappan*, 86 (1), 9-21.
- Black, P., & Wiliam, D. (1998a). Assessment and classroom learning. *Assessment in Education*, 5 (1), 7-74.
- Black, P. & Wiliam, D. (1998b). Inside the black box: Raising standards through classroom assessment. *Phi Delta Kappan*, 80 (2), 139-148.
- Boersma, G., Brahmstedt, J., Bruyn, M., & Causing, K. (1995). Improving student self-evaluation through authentic assessment. Master of Arts in Teaching and Leadership Action Research, St. University. Theses- Master Theses (042), 1-138. (ED393885).
- Boersma, K., Dye, T., Hartman, E., Herbert, L. & Walsh, T. (1997). *Improving Writing Skills through Student Self-Assessment*. Published Master's Action Research Project, Saint Xavier University and IRI / Skylight, Illinois, USA, 1-72. (ED 411 517).

- Bogart, J. (n.d.). *Writing Portfolios: What teachers learn from student self-assessment*. Retrieved March 20, 2007, from:
<http://www.evergreen.edu/washcenter/resources/acl/e2.html>
- Borg, S. (2006). *Classroom Research in English Language Teaching in Oman*. Muscat, Ministry of Education, Sultanate of Oman: Oriental Printing Press.
- Brantmeier, C. (2005). Nonlinguistic variables in advanced second language reading: Learner's self-assessment and enjoyment. *Foreign Language Annals*, 38, (4), 494-504.
- Brantmeier, C. (2006). Advanced L2 learners and reading placement: Self-assessment, computer-based testing, and subsequent performance. *System*, 34, 15-35.
- Bruce, L. (2001). Student self-assessment: Making standards come alive. Retrieved July 9, 2006, from:
<http://www.ascd.org/readingroom/classlead/0109/2Sept01.html>.
- Cassidy, S. (2006). Learning style and student self-assessment skill. *Education and Training*, 48 (2-3), 170-177.
- Collins, J. & O'Brien, N. (2003). *The Greenwood Dictionary of Education*. USA, Westport: Greenwood Press.
- Crooks, T. (2001). The validity of formative assessment. Paper presented to *The British Educational Research Association Annual Conference, University of Leeds*, 13-15 September, 2001. Retrieved February 24, 2007, from:
<http://www.leeds.ac.uk/educol/documents/00001862.htm>

- El-Kilabi, A. (2005). The place of reading comprehension in second Language acquisition. In *Proceedings of the Fifth National ELT Conference, 198-207*). Sultan Qaboos University, Muscat 20-21, April, 2005.
- Falchikov, N. & Goldfinch, J. (2000). Student peer assessment in higher education: A meta-analysis comparing peer and teacher marks. *Review of Educational Research, 70* (3), 287-322.
- George, J. & Cowan, J. (1999). *A Handbook of Techniques for Formative Evaluation: Mapping the Students' Learning Experience*. London: Kogan Page Limited.
- Gipps, C. (1994) *Beyond Testing: Towards a Theory of Educational Assessment*. London: Falmer Press.
- Grabe, W. (1991). Current developments in second language reading research. *TESOL Quarterly, 25* (3), 375-406.
- Gronlund, N. & Linn, R. (1990). *Measurement and Evaluation in Teaching*. New York: Macmillan.
- Guskey, T. (2003). How Classroom assessment improves learning. *Educational Leadership, 60*, (5), 6-11.
- Harrison, C. (2001). Reflective pupils lift exam grades. *The National Literacy Trust 2006*. Retrieved January 23, 2007, from <http://www.literacytrust.org.uk>
- Hussein, N. (2005). Understanding words in context. *Muscat Message*, Ministry of Education. Muscat: Al Nahda Printing Press, 50-52.
- Liebovich, B. (2000). Children's Self-Assessment. Issue in Early Childhood Education: Curriculum, Teacher Education, and Dissemination of Information. *Proceedings of the Lillian Katz Symposium* (Champaign, IL, November 5-7, 2000). PS 030 765, 237-247. (ED 470 895).

- Little, D. (2005). The common European framework and the European language portfolio: Involving learners and their judgments in the assessment process. *Language Testing*, 22 (3), 321-336.
- Manuputty, R. (2000). An Investigation of Student Self-Assessment and Writing Progress in an English Composition Course at Pattimura University, Ambon, Indonesia. [M.A. Dissertation]. Canada: Simon Fraser. Publication Number: AAT MQ61465. *ProQuest Dissertation and Theses 2000. Section 0791, Part 0279* 132 pages.
- Marcotte, A. (2006). Incremental and Predictive Utility of Formative Assessment Methods of Reading Comprehension. [Ph.D. Dissertation]. United States, Massachusetts: University of Massachusetts Amherst. Publication Number: AAT 3215909. *ProQuest Dissertation and Theses 2006. Section 0118, Part 0535* 128 pages.
- Meldrum, R. (2002). The student experience of peer- and self-assessment as a social relation. Paper presented at the *Learning Communities and Assessment Cultures Conference* organized by the EARLI special interest group on assessment and evaluation, University of Northumbria, 28-30 August. Retrieved February 26, 2007, from <http://www.leeds.ac.uk/educol/documents/00002256.htm>
- The Ministry of Education (MOE). (2003-2004). *Our World through English, The English Language Course for the Sultanate of Oman, Teacher's Book 3S (B) (1st ed.)*. Muscat: Muscat Printing Press.
- Mission of Oman to the United Nations MOUN. (2000). Education. Retrieved March 3, 2007, from <http://www.etectonics.com/oman/gi/education.asp>
- Morris, B. (2004). Informing Reading in an Area of Accountability: The Relationship

- between Formative Assessment and High Stakes Testing. [Ph.D. Dissertation].
 United States, Virginia: University of Virginia. Publication Number: AAT 3131462.
Proquest Dissertation and Theses 2004. Section 0246, Part 0535 180 pages.
- The National Capital Language Resource Center (NCLRC). (2004). *Teaching reading: Goals and techniques for teaching reading*. Retrieved January 3, 2007, from <http://www.nclrc.org/about.html>
- O'Malley, J., & Pierce, L. (1996). *Authentic Assessment for English Language Learners: Practical Approaches for Teachers*. Boston: Addison-Wesley Publishing Company.
- O'Reilly, M. & Al-Lawatia, N. (2005). On the road to self-assessment. *Muscat, ELT in Oman*, Ministry of Education. Muscat: Al Nahda Printing Press L. L. C. PP. 22-24.
- Peat, M. (2000). Online assessment: The use of web based self-assessment materials to support self directed learning. In A. Herrmann and M.M. Kulski (Eds.), *Flexible Futures in Tertiary Teaching*. Proceedings of the 9th Annual Teaching Learning Forum, 2-4 February 2000. Perth: Curtin University of Technology. Retrieved December 17, 2006 from <http://lsn.curtin.edu.au/tlf/tlf2000/peat.html>
- Rassekh, S. (2001). Case study on gender-sensitive educational policy and practice in the Sultanate of Oman. A paper presented at *The Management of Curriculum Adaptation for Curriculum Specialists in the Gulf Region*: Muscat, 17-21 February 2001. Retrieved March 2, 2007, from: <http://www.google.co.uk/search?q=educational+system+in+oman&hl=ar&start=20&sa=N>
- Rivas, R. (1999). Reading in recent ELT coursebooks. *ELT Journal*, 53 (1), 12-21.
- Sasidharan, P. (2005). Continuous assessment: A review. *Muscat Message*, Ministry of Education. Muscat: Al Nahda Printing Press, 60-62.-

- Sert, N. (2006). EFL student teachers' learning autonomy. *Asian EFL Journal*, 8 (2), 180-201.
- Shaaban, K. (2005). Assessment of young learners. *English Teaching Forum* 43 (1), 34-40.
- Shatery, M. & Inder, D (2005). Addressing the reading difficulties of FL learners: A problem-solution approach. *In Proceedings of the Fifth National ELT Conference*, 135-140. Sultan Qaboos University, Muscat 20-21, April, 2005.
- Shen, H. (2002). Motivational and Self-Regulated Learning Components in Relation to Language Learners' Self-Assessment, Reading Strategy Use and Reading Achievement. [Ed.D. Dissertation]. United States-Washington: Seattle Pacific University; 2002. Publication Number: AAT 3148896. *Proquest Dissertation and Theses 2002. Section 1043, Part 0525* 174 pages.
- Sheng, H. (2000). A cognitive model for teaching reading comprehension. *English Teaching Forum* 38 (4), 12-15.
- Siddiqui, N. (2005). How do language proficiency, schema and reading strategies influence reading comprehension at higher levels in an ESL context? *In Proceedings of the Fifth National ELT Conference*, 171-178. Sultan Qaboos University, Muscat 20-21, April, 2005.
- Stiggins, R. (2001). *Student-Involved Classroom Assessment (3rd ed.)*. N.J., Upper Saddle River: Merrill Prentice-Hall.
- Thompson, G., Pilgrim, A., & Oliver, K. (2005). Self-assessment and reflective learning for first – year university geography students: A simple guide for simply misguided? *Journal of Geography in Higher Education*, 29 (3), 403-420.
- UK Centre for Legal Education (UKCLE). (2004). The impact of formative assessment

on student learning. Paper presented at the *Society of Legal Scholars Conference*,
September, 2004. Retrieved April 25, 2007, from
<http://www.ukcle.ac.uk/research/projects/bone.html>

Wragg, E. (2001). *Assessment and Learning in the Primary School*. London: Routledge
Falmer.

المراجع العربية

وزارة التربية والتعليم في سلطنة عمان. (2005). *35 عاما من التعليم*. مسقط: وزارة التربية والتعليم، المديرية العامة للعلاقات
العامة.

Appendix 2

One-Minute Paper

Dear Student,

Self-assessment is a vital component in learning. You are kindly requested to frankly comment on your own learning. You are encouraged to share responsibility for your own learning. Teaching cannot be effective unless the teacher comes to know your strengths and weaknesses. Your responses will help me find out how the course is going on and give me an idea on whether any changes are needed. This one-minute paper is mainly prepared to help you, and it won't take much of your time. The more you are open and accurate in answering these questions, the more progress in teaching and learning there will be.

1. What are the most important things you have learnt in today's class period?

.....

2. Which area(s) of the lesson did you fail to grasp?

.....

3. Which point(s) of the lesson is/are still not clear enough in your mind?

.....

I appreciate your cooperation.

Adapted from Angelo, T. & Cross, K. (1993). Classroom Assessment Techniques: A Handbook for College Teachers (2nd ed.). San Francisco, CA: Jossey-Bass.

Appendix 3

Rating-Scale Sheet (Self-assessment for Reading)

Name:	Date:..... / / 2007
<p>Before I read “.....,” I:</p> <p>Yes No</p> <p>_____ _____ thought about the title and what it suggested the text was about.</p> <p>_____ _____ previewed the whole text or parts of it.</p> <p>_____ _____ thought about the subject or situation.</p> <p>_____ _____ set a purpose for my reading.</p>	
<p>While I was reading “.....,” I:</p> <p>Yes No</p> <p>_____ _____ developed a dialogue with the writer (e.g., What is the writer communicating? What is the main idea? What do I already know about this?).</p> <p>_____ _____ visualized what places, people, events might look like.</p> <p>_____ _____ connected my personal experience to what I was reading.</p> <p>_____ _____ made inferences from textual clues given by the writer.</p> <p>_____ _____ tried to distinguish between fact and opinion.</p> <p>_____ _____ predicted and then checked what the writer might say next.</p> <p>_____ _____ went over the parts I found confusing.</p> <p>_____ _____ checked words that I did not know the meaning of from context.</p>	
<p>After I read “.....,” I:</p> <p>Yes No</p> <p>_____ _____ determined my initial impression of what I had read.</p> <p>_____ _____ discussed what I had read and my impressions with someone.</p> <p>_____ _____ reflected on what I had read.</p> <p>_____ _____ reviewed and summarized what I had read and learned.</p> <p>_____ _____ made notes in my notebook.</p> <p>_____ _____ developed a more thoughtful interpretation of what I had read (considered why the writer wrote the text, what was being presented, and how it was constructed).</p> <p>_____ _____ evaluated what I had read and supported my judgments with references to the text.</p>	

Adapted from: Saskatchewan Education. (1998). English Language Arts 20: A Curriculum Guide for the Secondary Level. Regina, SK: Saskatchewan Education.

Retrieved February 9, 2007 from,

<http://www.sasked.gov.sk.ca/docs/ela20/teach5.html>

**Appendix (4)
Reading Test
2006 / 2007**

Page 1

Student's Name:

Grade 12 /

TIME: ONE HOUR

This paper has 4 pages.

TOTAL MARK: 60

READING 1 (Items 1 – 6) ... (12 points)

Match the six texts on the left with the texts in the box. Write the correct letter, as in the example. There are three extra texts in the box.

EX. *She finished Grade 12 in 1999.* F

1. Maria Borgia was born in _____
1958 in a village in the south
of Italy.

2. Three months ago, Leila _____
started going to French
classes in the evenings.

3. After leaving university, my _____
sister got a job with an
archaeologist.

4. After visiting her uncle in _____
Dubai, Salma got into her car
and drove back to Muscat.

5. I know that many people _____
complain that Muna talks
too much.

6. Police in Singapore are _____
looking for a German woman

A The journey was fine but, she got home late because she had to wait a long time at the border.

B She's such a careless driver. I'm amazed she's never had a serious accident.

C She started with a four-week training course. Then she joined a team who were studying the ruins of Troy.

D But actually I like her. What she says is always very interesting, and she has a great sense of humour.

E They don't have a photo of her, but they say that she is tall and thin, has short blond hair and wears glasses.

F Then she went to Sultan Qaboos University to study Law.

G She's really enjoying herself. She's met some very nice people, and she's already learnt to speak the language quite well.

H If I phoned her or wrote her a letter, there could easily be a misunderstanding.

I So, she stayed in cheap hotels, and instead of renting a car, she travelled everywhere by bus or train.

J However, when she was only six months old, her family moved to a large industrial city in the north.

in her late thirties.

READING 2 (Items 7 - 12) ... (18 points)

Page 2

Read each short text, and then write **ONE** word or number to complete the sentences.

In the 1970s, the populations of Nizwa and Sohar were almost the same, but since then many people have moved to Sohar looking for work, while Nizwa has been much less successful. Both cities have, of course, much larger populations than the country's third city, Sur.

.....
7. The city with the largest population is _____.

The tourist guide told us that the trip would start at seven o'clock. So we got up early, and waited patiently at the entrance of the hotel. However, the bus was two hours late, so we didn't reach the fort until twelve o'clock.

.....
8. The bus left the hotel at _____ o'clock.

We are worried about Salim's behaviour during the last few weeks. He doesn't pay attention in class, and he's always late in completing his project work. Would it be possible for us to have a meeting to discuss this?

.....
9. The letter is to a _____.

The 'Ultra-Lex' is full of useful information for you. It tells you the meaning of a word; how it is spelt; how it is pronounced; and how it is used. It also provides real-life examples of more than 10,000 words in the English Language. Available as a book or as a CD.

.....
10. This is an advertisement for a _____.

We are arriving in Singapore harbour tomorrow afternoon. The passengers are very excited, as they haven't been on land for two weeks now. Some of them got very sick during the big storm that we went through last Friday. We had to work very hard to help them.

.....
11. Where is the writer? On a _____.

I didn't take my umbrella with me because the weather's usually dry at this time of the year. But after half an hour, a light drizzle started. Then it got heavier and heavier, and by the time I arrived, I was soaking wet. Just my luck!

.....

12. The word 'soaking' means _____.

READING 3 (Items 13 - 21) ... (18 points)

Page 3

Read the text and then, for each item, choose the correct option: A, B or C.

In October 1767, Isabel Grameson decided to leave her home in Ecuador and travel to Brazil to find her husband. Seven years before, he had left on a journey of exploration along the Amazon River. For years, she had heard nothing from him, until finally news came that he was on the north coast of Brazil and that he was seriously ill.

When they heard of her decision to leave, Isabel's family were very worried. She had no experience of travelling, and the journey across the South American continent would be extremely dangerous. Finally, it was agreed that her brother would travel with her, as well as a doctor and two Indian servants.

At first, as they walked down the eastern slopes of the Andes mountains, everything went well. After twelve days, they reached the Bobonaza River. Here the Indians built a small boat, which they piloted down the river. This wasn't easy because the water was rough, and soon they hit a large rock in the middle of the river. The boat sank, all their food and clothing was lost, and the doctor and both Indians drowned.

Isabel and her brother managed to swim to safety, but now they were stuck on a flat muddy piece of land, with no food, no shelter and no boat. So they agreed to try to walk through the forest. This was not a wise decision. They were soon completely lost and very hungry. At night, they were kept awake by the noise of animals in the forest and bitten by thousands of ants, flies, and mosquitoes. Finally, after six days with no food, Isabel's brother died. At first, she thought she would stay where she was and die, but then she decided that she had to survive. She kept on walking and was able to find some wild fruit to eat. After four more days, she was found by an Indian family, who nursed her back to health and then took her to the nearest town. A month later, she was finally re-united with her husband.

13. Isabel's husband was _____.

A. a doctor

B. a writer

C. an explorer

14. She had not seen him since _____.

A. 1757

B. 1760

C. 1767

15. At the beginning of her journey, she travelled _____.

A. by horse

B. by boat

C. on foot

16. In the accident on the River Bobonaza, _____ people died.



Title:

EFL in Higher Education: Designing a Flexible Content-Based Curriculum at
University-Level

Author:

Dr. Zubeyde Sinem GENC

zsgenc@yahoo.com

Bio Data:

Dr. Zubeyde Sinem GENC is currently Assistant Professor in the ELT Department at Uludag University, Turkey. She received her Ph.D. from the Indiana University of Pennsylvania (IUP), USA. She taught EAP at IUP and graduate courses in the MATESOL program at Southern Illinois University. Her research interests are second language teacher education, TESOL theory, curriculum development and second language acquisition.

Abstract

The current trend of globalization and the developments in information technology have boosted the role of English, which has become a universal language of knowledge and communication. In this demanding and challenging information era in which we live, EFL instruction at higher education institutions needs to offer the students more than general proficiency in English. They need to make the connection between English and students' future careers. However, it is not an easy task to design a curriculum at university level to address these issues. This article describes a content-based language program developed for teaching English as a

foreign language (EFL) at a Turkish university. It aims to illustrate the rationale and the process of designing a flexible content-based curriculum for university students. After a brief introduction to the place of English at Turkish universities, it discusses the limitations and inadequacies of the previous foreign language curriculum as well as the principles that the new program is based on at Uludag University. The curriculum presented in the article is intended to be a model for teaching EFL or other foreign languages at higher education institutions.

Key words: content-based instruction, university-level EFL instruction, curriculum development, Turkish EFL context

Introduction

Today, we witness an increasing awareness of the importance of learning foreign languages to successfully manage the current trend of globalization in the world (Oleksak, 2007). It is quite obvious that developments in information technology since the second half of the twentieth century have facilitated the access to knowledge produced elsewhere, which has contributed greatly to the globalization of knowledge, communication, social norms, values, and production/consumption habits. As a result of this, we find ourselves living in an information era where knowledge is power. The most crucial vehicle in the process of producing knowledge, reaching the produced knowledge and utilizing knowledge is, of course, language. This strong connection between knowledge and language makes it necessary to learn English as a foreign language (EFL) in those societies that speak other languages because English has become a universal language of knowledge and communication.

One of the basic components involved in this process of globalization in all societies is, of course, university. Universities, which are expected to be in

continuous interaction with universal knowledge, have the responsibility to educate students who are knowledgeable, motivated and equipped with the necessary skills to reach new sources of information for life-long professional development (Delanty, 2002; Fuller, 2003; Kuklinski, 2001). While designing their instructional curriculum, universities need to consider the connection between knowledge and language mentioned above, and enact dynamic and necessary reforms in order to reflect this onto the education they offer. In this sense, it becomes unavoidable for non-English speaking societies to incorporate an EFL program into their higher education in order to share the knowledge produced and to reach knowledge produced elsewhere. Integrating EFL teaching across the curriculum at higher education institutions is a challenging task and may have various formats. With the globalization and the increasing demands of national and international competitive job markets, EFL instruction at universities has to provide more than general proficiency in English. Students need to be trained to use English language for the special purposes required by their future careers. In other words, EFL instruction at universities needs to establish the strong connection between the target language and the subject matter of students' future careers.

These issues raise the question of whether the EFL program in one's own context incorporates the crucial features mentioned above. Unfortunately, the program at the higher education institution where I teach was far from having the characteristics of such a program. Indeed, we can talk about mainly two dimensions of the motivation for designing a new curriculum. The first reason related to the ineffectiveness and the failure of the existing EFL program. The inadequacy of the program could be observed in terms of the content, materials, texts and resources,

and in terms of the number and quality of teaching staff. The second reason for the search for a new program was related with the improvements and the educational reform that the university administration has planned to do. The ultimate goal of the reform was to increase the quality of education, prepare the university for the international accreditation process, fulfill the requirements of European Union University Quality Culture, participate in student exchange projects with European countries, and acquire a respected and preferred university identity. In this reform process, teaching English was considered crucial for the students' future careers. Students should be competent in reading, writing, understanding and speaking English, especially in topics related to their major fields of study. Competence is defined as "the ability to produce and understand the sentences of a given language, and identify ambiguous and deviant sentences" (Johnson & Johnson, 1999, p.75). The notion of competence needs to be clarified for the purpose of this article, especially after we discuss and consider English as an international language (EIL) in the Introduction section. Recent research questions the validity of the notion of ideal native speaker competence that Chomsky (1965) has proposed. Nunn (2007) redefines competence for EIL as a holistic, global and international concept, and argues that institutions which aim educational excellence in an international field need to establish a balance between international and local definitions of competence. Alptekin (2002) finds native speaker competence utopian, unrealistic and constraining, and indicates "for language to be authentic in its routine pragmatic functioning, it needs to be localized within a particular discourse community" (p. 61). The importance of community in defining the notion of competence has been emphasized by Nunn (2007) as well, who states "Competence

is linked to the notion of community. In any local context competence needs of students can be related to the different communities with and within which they will need to communicate” (p. 101). It is pointed out that a more holistic explanation of language competence would help grasp the characteristics of various local and international communities that the users of EIL communicate in. The implications of this framework for the definition of competence on language pedagogy have been discussed (Acar, 2007; Alptekin, 2002; Hyde, 1998; Nunn, 2005). Alptekin (2002) states “a new pedagogical model is needed to accommodate the case of English as a means of international and intercultural communication,... the EIL pedagogy should be one of global appropriacy and local appropriation” (p.63). Learners are expected to have the competence to use language effectively and appropriately in various contexts (Spolsky, 1989). Developing learners’ ability to use a language effectively for their communicative purposes in a variety of professional and social contexts requires learning environments which emphasize communicative, authentic and meaningful use of target language, and active learner involvement (Finney, 2002).

Upon the observations of problems in the existing program (see the Section below) and the educational reform desired in this area, the university administration appointed me to develop a new program to teach EFL to university students, which would be gradually implemented in all the departments.

EFL at Turkish Universities and Uludag University

There are about 150 universities in Turkey (as of December 2010). The majority of them are state universities. In the 1990s, we saw an increase in the number and

quality of private universities as well. Only in a few of the state universities, such as the Middle East Technical University and Bosphorus University, is the medium of instruction English. These universities have been established in cooperation with universities and organizations in the United States. In private universities such as Bilkent, Beykent, Isik and Izmir University of Economics, English is also used as the medium of instruction. In other state or private universities, not all but some departments and faculties offer English-medium instruction. Therefore, we can talk about mainly three types of universities in regard to the English instruction offered:

- Universities that use English as a medium of instruction in all their academic programs
- Universities that use English as a medium of instruction in some of their academic programs
- Universities that use only the native language in all their academic programs

The majority of the universities use the first language (L1), Turkish, in their academic programs, and English is taught as a foreign language in separate courses. This is, of course, natural if we consider the rationale for conducting higher education in the native language. However, it is obvious that universities using English as medium of instruction have more advantages because students graduating from these universities are equipped with a richer education if we take into consideration the connection between knowledge and language discussed in the introduction. This reality gives the L1-instructed universities the responsibility of teaching English through separate English language courses, which are limited in terms of weekly hours, their content and goals.

Uludag University is one of the largest state universities in Turkey as it consists of eleven faculties, fourteen vocational schools and about forty thousand students. The medium of instruction and study at Uludag University is Turkish. In order to fulfill the foreign language requirement, a separate EFL course was offered for only two hours a week during the first year. However, this basic language course fell short of meeting students' needs in learning the English language. The major problems can be stated as follows:

1. EFL teaching meant just a two-hour English course which was offered during the first year. There was not any other course or activity related to EFL in the following years of university education.
2. Students were not grouped according to their proficiency levels in English. Students at different proficiency levels followed the same syllabus, which consisted of beginner – low intermediate level materials.
3. Classrooms were too crowded. There were over eighty students in one class, seated in a conference room.
4. There was not adequate equipment or other materials, except for the textbook, that would support language teaching. Occasionally, instructors brought a tape recorder; however, only students sitting at the front could hear it. Students sitting in the back row could not even hear the instructor.
5. This two-hour English course aimed to teach basic general English for all students in different fields of study, and it mostly emphasized grammar and vocabulary knowledge. However, students needed to build on it and learn

English for professional purposes at advanced levels so they could use English effectively when they start their careers after graduation.

6. Teaching staff were not adequately qualified and did not have the vision or motivation to implement more challenging syllabi.

During the initial phases of the educational reform which aims to make the university one of the most promising and serious ones in the country, it was emphasized that students graduating from different departments should be able to use English effectively in their future careers. Thus, teaching professional English became the major objective of the new EFL program and it was believed that students needed to come to a certain proficiency level before they started using English for professional purposes; i.e. for their future careers. For this purpose, the following beliefs and assumptions underlined the new EFL curriculum:

- Developing students' proficiency and ability to use English effectively in a variety of contexts related to their future careers,
- Increasing the quality of education by creating opportunities for students to utilize resources and materials written in English in their own majors.
- Increasing students' opportunities in the job market in terms of technology and know-how, and the economy and commercial areas.
- Competing with more prestigious universities in recruiting successful high quality students who are tempted to choose universities that offer all English-medium instruction.
- Adopting English as a prerequisite for the accreditation process implemented by the university.

- Increasing opportunities for student and staff exchanges with universities in other countries
- Encouraging students and staff to take part in international joint projects in their fields.

When we examine the motivation and the assumptions behind the attempt to design a new EFL curriculum in this context, we can see that a content-based language program would be suitable to provide opportunities for all students with different majors of specialty to use English effectively in their future careers. The new program was expected to be flexible so that it could be implemented in all majors, but adopting the content, materials and tasks of specific majors. Content-based instruction (CBI) where content learning and language learning are integrated presents a valuable option for teaching foreign languages at universities. Content-based language instruction, an influential approach in language pedagogy, aims to develop competence in a foreign language and improve the knowledge of a subject matter at the same time.

Content-Based Language Teaching

CBI allows for creating environments where learners can be effectively involved in using the target language for various communicative and meaningful purposes. CBI is defined as “the integration of language teaching aims with subject matter instruction” (Snow, 2001, p. 303). Grabe and Stoller (1997) indicate that combining the development of language with content knowledge enhances the learning process. They argue that content-based language instruction is supported by research done within the framework of educational and cognitive psychology and point to

Anderson's research (1990, 1993) on learning and processing which emphasizes the importance of coherent and meaningful information for better learning and recall. Similarly, Singer (1990) emphasizes that when learners are exposed to thematically organized materials they learn and remember better the information presented. In CBI, materials are naturally presented around certain topics focusing on coherent and meaningful information as well as relevant language learning activities in a context. CBI continually provides opportunities for learners to use their knowledge of content area and target language in increasingly complex tasks. Integration of language and content provides a meaningful context for learners to foster their academic and cognitive development as well as the skills and proficiency in the target language.

The development of CBI is derived from immersion programs in Canada to teach French as a second language to English speaking children in schools. We observe three types of immersion programs: early immersion where the first three or four grades of schooling are done completely in the second language; delayed immersion where the fourth and fifth grades receive instruction in the second language, and late immersion where students in the seventh and eighth grades receive all instruction in the second language. The programs are also categorized as total immersion or partial immersion depending on the amount of instruction done in the second language (Genesee, 1985). In total immersion, all the instruction is conducted in the second language; while partial immersion means 30-70% of instruction in the target language. The immersion model has been adapted in the United States to enrich the school programs in terms of educational, cultural and linguistic levels, to establish a racial balance and to achieve bilingualism in minority populations. Many programs at secondary school and university levels have been

designed to address the needs of learners with limited English proficiency (Cantoni-Harvey, 1987; Crandall, 1987; Crandall & Kaufman, 1998; Met, 1998; Snow & Brinton, 1988; Snow & Kamhi-Stein, 1997; Wegrzecka-Kowalewski, 1997). Models of content-based programs are distinguished from each other by the setting, by the instructional level and by the degree of emphasis on language and content (Snow, 2001). Some of these models offer the majority of the education through the foreign language while the amount of time the foreign language is used for instruction might be much less in other models of content-based language teaching. They are also different from each other in terms of the degree of emphasis on language and content since some are more content-driven and others are more language-driven. Another variable is the setting; that is, whether the target language is used naturally in the environment or it is taught as a foreign language in educational institutions.

Among various models of content-based language programs, theme-based, CALLA and Language-Content-Task models constitute the basis of the new EFL program designed for Uludag University.

Theme-Based Model is organized around selected topics from one content area or from across the curriculum to develop students' general academic skills. Content is chosen from various topics according to students' interests. Stoller and Grabe (1997) propose the Six T's Approach for theme-based instruction, which include theme, topics, texts, threads, tasks and transitions. Teachers extract language learning activities from selected themes and topics (Snow, 2001). It is usually found in EFL contexts and taught by an EFL teacher or team taught with a content specialist (Davies, 2003).

CALLA (Cognitive Academic Language Learning Approach) is a strategy-based instruction that combines academic language development, content area instruction and

explicit learning strategy instruction for intermediate and advanced ESL students. It aims to improve academic language skills of ESL learners who have developed social communicative skills in the target language, ESL learners who have academic language skills in their native language but need to transfer them to the target language, or bilingual students who have achieved social communicative skills but lack academic language skills in either language. In CALLA approach, content determines the academic language selection and the learning strategies to be taught. Learner autonomy is encouraged through the development of independent learning skills.

LCT (Language-Content-Task) integrates language, content and relevant tasks into the curriculum. The language component focuses on semantics, syntax, pragmatics and functions while content incorporates curriculum topics; tasks aim to improve procedural knowledge (Short, 2000). For example, when the **content** is about plant and animal cells, and the **task** is to use a microscope to view, describe and compare plant and animal cells, the **language** component of the lesson emphasizes vocabulary for cell parts, grammatical structures for describing and comparing, and the development of reading skills in the target language.

These models are by no means the only possibilities for designing a content-based language curriculum. Brinton, Snow and Wesche (1989) mention the possibility of new formats and different combinations of models and point out that the configuration of a model may differ significantly depending on the setting of the program. They indicate that the features of different models may blend together. Indeed, we can observe numerous examples in recent years that introduce well-designed and successfully-implemented content-based courses and programs in various settings. A good example is given by Adamson (2006), who used a

combination of teacher transmission and student collaboration for a content-based course on sociolinguistics at college level in an EFL context. Another recent example took place in Japan. Lingley (2006) integrated a task-based approach with content-based instruction for intermediate-level EFL learners to overcome the difficulties in teaching a content-based course on Canadian Studies.

The New Content-Based EFL Curriculum

The new content-based EFL curriculum at Uludag University is based on the following general principles:

- a) Gradual introduction / integration of content-based instruction
- b) Text and task authenticity
- c) Flexibility
- d) Learner and teacher autonomy
- e) Integration of technology
- f) Teacher involvement in the program design and development

The new curriculum has two major parts:

- Intensive English Program (IEP): An introduction of CBI is done in the Intensive English Program for students at intermediate and upper intermediate levels. Students at beginner and pre-intermediate levels attend courses to improve their proficiency in English.
- Content-Based Program (CBP): When students finish the Intensive English Program successfully, they start their academic departments. At this stage, a combination of Theme-Based, CALLA and Language-Content-Task models is implemented for integrating content learning with language learning.

The new curriculum incorporates a combination of several models because it is

spread throughout the higher education of students and involves several different stages. Each part is explained in detail indicating the goals and the content of courses. The major parts of the curriculum can be seen in the following figure:

<p style="text-align: center;">A- INTENSIVE ENGLISH PROGRAM (IEP)</p> <p>(Students attend IEP before they start their intended academic departments)</p> <p>a) For students at beginner (30 hours/a week), elementary (30 h/w) and pre-intermediate (25 h/w) levels:</p> <p>Intensive English Courses: Language courses taught by language specialists to improve overall proficiency of students</p> <p>b) For intermediate (25 h/w) and upper intermediate (20 h/w) students:</p> <p>1) Intensive English Courses: Language courses taught by language specialists to improve overall proficiency of students</p> <p>2) Introduction to CBI: Language-driven courses focusing on language learning strategies and on various themes taught by language specialists only.</p> <p>Project work is included.</p>
<p style="text-align: center;">B- CONTENT-BASED PROGRAM (CBP)</p> <p>(Students take the following content-based courses after they start their intended academic departments)</p> <p>1) Theme-Based Courses: Advanced Writing Advanced Reading and Speaking</p> <p>Language-driven courses taught by a language specialist and supported by content specialist</p> <p>2) Language-Content-Task Courses: English for Professionals I English for Professionals II English in the Workplace</p> <p>3) CALLA Course: Research Paper</p> <p>Content-driven courses taught by content specialist and language specialist together</p>

Figure 1. The Content-Based EFL Curriculum at L1-Medium University

A- Intensive English Program (IEP)

The IEP is offered in a separate building at the School of Foreign Languages. All students attend this program for one year before they start their intended academic departments. The goal of this program is to bring students at different proficiency levels to upper-intermediate level in English and to offer an introduction to content-based instruction through topics and projects they undertake in and outside class. They are prepared to take content-based courses when they start their intended academic departments.

The School of Foreign Languages administers two exams at the beginning of the academic year. The first one is a proficiency exam that is administered to all students. Students who pass the proficiency exam are considered proficient in English and allowed to take content-based English courses while attending their own academic departments. The School of Foreign Languages has the right to waive the proficiency requirement if a student verifies her/his proficiency with a passing grade in other exams such as TOEFL, IELTS, CPE, and FCE. The second exam is a placement exam administered for students who cannot achieve a passing grade in the proficiency exam. The results form the basis for placing the students at different proficiency levels. Students are grouped in different levels according to the results of the placement exam. A level coordinator is assigned for each proficiency level to keep the courses offered integrated and balanced. For each class, there is a class tutor who usually teaches the greatest number of hours in that particular class. The courses offered in the IEP focus on target language skills such as reading, writing, listening, and speaking in English, and vocabulary and grammar. Students at the intermediate and upper intermediate levels are exposed to

topics of their future departments and assigned to do projects of their interests. The School of Foreign Languages has a library and a computer lab for students. For teachers, there are a resource room, a materials development room and a testing room.

Brown (1995) suggests that curriculum planners should avoid taking the position of an expert who knows what is good for the program and to seek the views of various interest groups. The curriculum proposed here includes different committees of EFL teachers responsible for important aspects of the program. Teachers take part in different committees responsible for different aspects of the program. These committees include testing, program development, materials development, professional development and student life improvement:

Testing Committee: This group of teachers prepares and administers all the exams and tests given to the students.

Program Development Committee: This committee is responsible for designing the weekly program and tracking whether teachers have any problems in the implementation of the curriculum.

Materials Development Committee: This committee's responsibility is to develop or adopt materials appropriate for each level.

Professional Development Committee: This committee is responsible for diagnosing the issues and topics that teachers need and want to focus on and for arranging seminars, workshops and other academic events that help teachers develop professionally.

Student Life Improvement Committee: This group of teachers deals with activities, issues and problems of students and aims to improve students' academic

and social lives at the university.

The curriculum of the IEP was planned to involve both in-class work and out-of-class work. These operate as follows: **In-class work** consists of courses in reading and vocabulary, writing, grammar, listening and speaking taught in class. Students at each level take four mid-term exams and numerous quizzes in one semester. At the end of the academic year, successful students are allowed to take the proficiency exam. If they get a passing grade in the proficiency exam, they register at their intended academic departments. Those who are not successful in the coursework are not allowed to take the proficiency exam. They can attend the summer school and take the proficiency exam given at the end of the summer school. If they are successful, they can go on with their education in their departments; if not, they have to repeat the IEP. **Out-of-class work** encourages students to do extensive reading, individual and group projects that supplement in-class topics and issues. Besides, students are encouraged to study in the computer lab where they have access to the Internet and to authentic materials and visual activities. The progress of each student is tracked by class tutors. If needed, extra tutoring hours for low level students are organized. They are not allowed to take advanced level content-based courses unless they succeed in the Intensive English Program and pass the proficiency exam.

Content-Based Program (CBP)

The second part of the curriculum is the content-based EFL program. After students successfully finish the IEP at the School of Foreign Languages and pass the proficiency exam, they take content-based courses in English starting in the

freshman year while taking their academic classes. Each of these courses is for four hours a week. Students need to pass a course in order to take another one in the following semester. Otherwise, they have to repeat the same course. In other words, the courses are a prerequisite for one another. The CBP aims to enhance students' use of English, oral and written, in their future careers. In this sense, the materials covered are chosen in accordance with students' future careers. For example, if they are medical students, topics covered in all the activities and materials are on medicine.

The content and the materials for the courses are determined by the language instructors together with the faculty of the specific department. Although the courses are conducted mainly by the departments' teaching staff, they cooperate to organize the courses together. The ultimate goal of the courses is to equip students with the language skills needed to read, write, understand and speak in English in the areas and topics related to their future careers. The CBP emphasizes teaching English for professional purposes. The common goal is to integrate content teaching with language teaching. The aim is to use the subject matter of the students to develop students' academic skills, to improve their foreign language skills and at the same to teach the subject matter to them. The content-based curriculum focuses on preparing students to have strong proficiency in English and to perform successfully in content areas using English. Considering the objectives of the program, the student population and the setting, our model is more language-driven in the beginning and gradually becomes more content-driven. The foreign language is used in the content-based language classes. The ultimate goal of the Program is to equip students with the language skills that will enable them to use

English effectively in a range of written and oral contexts related to their professional lives, and to help them use all kinds of resources in English to become knowledgeable and intellectual individuals in the subject matters in which they specialize. The first two courses treat language skills separately. However, the integration of language skills is achieved in the courses students take afterwards such as Professional English and English at the Workplace. Besides the integration of skills, there is a focus on authentic tasks students might encounter in their real professional lives. The required content-based courses originally included the following ones:

Advanced Writing Skills in English: The goal of this course is to reinforce students' writing skills, to do remedial structure and vocabulary work, and to improve their summarizing and paraphrasing skills focusing on different modes of written language and various topics related to their subject matters. The focus is on the language more than the content. This course is offered for four hours a week during the first and second semesters in their departments.

Advanced Reading and Speaking Skills in English: In this course, students are expected to improve their reading strategies to understand advanced texts on the topics related to their fields and be able to conduct debates, participate in discussions and do individual or group oral presentations on the subjects they read. It is a language-driven course although a considerable amount of time is spent on the content as well. This course is offered for four hours a week during the third semester.

English for Professional Purposes I: This course is content-driven and follows the basic principles of content-based instruction. In that sense, students

focus on the content by reading, listening, speaking and writing about their subject matters. Task types in this course include *information management* where students sift information into different categories, or *hands-on activities* where students manipulate information through games and experiential activities (Nunan, 2001). Professional authenticity of the tasks and the texts covered is crucial. Students improve their terminology and ability to follow the literature on the topics related with their future profession. This course is offered for four hours a week during the fourth semester of their education.

English for Professional Purposes II: Since this course is a continuation of the course offered in the previous semester, the same principles for the task types and texts are implemented in this course. Students are expected to reinforce their content knowledge and language skills to handle the written and oral texts in English related to their future profession. This course is offered for four hours a week during their fifth semester in the department.

English in the Workplace: The goal of this course is to equip students with advanced oral and written communication skills in English and the comprehensive content knowledge they will need in their relations with foreign individuals and institutions after they graduate and start pursuing their professions. They are expected to carry out tasks similar to the ones they will face in their future careers. This course is offered for four hours a week during their sixth semester.

Research Paper in English: Students are equipped with the necessary research skills to prepare and write a research paper on a topic of their interest related to their future careers. The majors of the students determine the requirements for such a paper because content-based academic writing is

incorporated within this course (Tutunis, 2000). The outcome of the course is a kind of graduation thesis on the content of students' specific fields. Since this course is CALLA-oriented, students are encouraged to focus on the development of academic English as well as the content area and learning strategies. Strategy instruction is an integral part of this course to. It is offered for four hours a week during the seventh and eighth semesters.

The content and materials to be used in the courses are determined together with the faculty of each department. The courses "Advanced Writing Skills, and Advanced Reading and Speaking Skills in English" are more language-driven and taught mainly by language instructors, whereas the other courses are increasingly more content-focused and taught by the faculty in departments with the help of language instructors. Language instructors receive guidance and help from the faculty on the topics and materials to be covered in the courses. In turn, content instructors (the faculty of the departments) are supported by language instructors in terms of the language work to be done in the courses. The common purpose of these courses is to improve students' language skills in order to use them for professional purposes.

The program outlined above has been approved by the University Senate. However, the Higher Education Council requested some modifications regarding the first year (first and second semesters) and the last year (seventh and eighth semesters). The courses "Advanced Writing Skills and Research Paper in English" have been removed from the program indicating that students are required to take the two-hour foreign language course in the first year and they are required to prepare a thesis in Turkish in the last year of their education. The Council thought

students would be overloaded with two English courses in the first year and another research paper in English in the last year. The final version of the content-based EFL program has been approved and recommended by the Higher Education Council as a model curriculum for all other universities with L1-medium instruction in Turkey since 2002-2003 academic year.

Going beyond the Curriculum

The guiding principle of the whole program is to encourage both learners and teachers to be autonomous. One of the important concepts that the program tries to establish is learner autonomy. Learners need to go beyond classroom instruction, take responsibility for their own learning and plan for their learning process in the future after they leave the intensive English program at the School of Foreign Languages. For this purpose, students are trained to acquire and use effective language learning strategies. Teacher autonomy is as important as learner autonomy. Raya (2007) indicates that teacher autonomy is maintained when institutions promote teacher involvement in decision-making process, increase opportunities for peer collaboration and discussion of real school problems, and encourage teacher responsibility and choice in the teaching process. In other words, for teachers, being autonomous means to be able to make their own informed decisions in their teaching, to be able to address needs and problems as they arise in the classroom and to be able to go beyond the written program/curriculum if/when necessary. For this purpose, teachers in this program were introduced to various ways for professional development to achieve autonomy. These included action research, peer and self-observation, reflective journals and utilizing technology to improve

their teaching. Seminars and workshops were arranged to present information on teacher autonomy.

Feedback from Students and Instructors

This section includes the first set of feedback on the education at the School of Foreign Languages. Since the other parts of the curriculum had not started fully at the time of collecting data, we do not have feedback for the other parts of the curriculum. At the end of the first year, instructors and the students were asked to provide feedback about the curriculum. The following issues were observed to be addressed immediately:

Students split into two major groups in terms of their level of motivation. One group was motivated and really wanted to learn English. They believed English was necessary for their future careers and for better positions they planned to have in the future. The other group's aim was just to pass the English courses and concentrate more on their own majors. Students thought that they would not need English later in their education because the medium of instruction was not English. In order to address this problem, an orientation week was planned to increase students' interest and motivation in learning English. The benefits of learning English were explained explicitly to students.

It has been observed that students' grammar, vocabulary and reading skills were quite good. However, they were very weak in speaking, listening and writing skills. This was the general profile of most EFL learners. In order to solve this problem, the number of class hours in speaking, listening and writing was planned to be increased; and students were encouraged to do out-of-class projects in these

problem areas.

Students did not show much interest in the topics covered in reading textbooks. For this reason, the books were supplemented with extra reading materials that students chose.

Writing was a difficult skill for most students. The students needed to go beyond paragraph writing and produce longer pieces of various genres. A workbook of extra materials for writing was planned to be prepared. Self-assessment was introduced to help the students become more autonomous and confident in their writing.

Preparing instructional materials for the content-based courses is quite challenging. In EFL contexts, students have limited opportunities to be exposed to the authentic language. A number of criteria for selecting the materials was planned to be incorporated into the curriculum. The important criteria for material selection were task and text authenticity. Materials should be:

- authentic and real-life situations that students would face during their professional experiences in the future.
- diverse in terms of sources and genre.

For this purpose, relevant websites, terminology glossaries, current news, source books, slides and worksheets were planned to be used more effectively. Designing appropriate instructional activities and tasks related with the materials was a challenging area that needed special attention.

Another important issue relates to the use of educational technology. For this purpose, the number and use of computer labs were planned to be increased in order to promote computer-assisted language learning (CALL) and expose students

to authentic materials through the internet. Teachers were asked to do research on the topic and share the results with their colleagues during a workshop. They were also encouraged to improve themselves on this topic since CALL was crucial for students to be autonomous in learning a second language. Use of e-learning was planned to be explored in detail so that a combination of e-learning and face-to-face sessions could be conducted.

Conclusion

The content-based curriculum has its rewards and challenges. It is rewarding because it provides a flexible framework to design a curriculum for students of different majors at a university setting. At the same time, it is challenging because it requires a lot of thinking and effort to coordinate the program as a whole and to find materials and implement the courses in each specific department.

With these points in mind, the program presented in this paper aims to give an idea about the possibilities of designing similar or different content-based EFL programs at universities. Students who finish the content-based language program successfully are expected to be proficient in both the target language and the subject matter they have studied for their profession. They can update their knowledge since they get access to the latest developments in their fields through the sources most of which are written in English. As an international language, English is perhaps the most effective tool for the globalization of the world, bringing countries closer and closer. This program is expected to create opportunities for a more interesting and motivating environment to learn and use English to fulfill real purposes and to develop much wider knowledge in students'

future careers. It also aims to develop valuable academic skills such as note taking, summarizing, and extracting key information from texts and to reinforce higher-level thinking skills as students gather information from different sources, evaluate and synthesize. The ultimate achievement is, of course, to equip our students with proficient skills in English, which will hopefully help them find a better job, and use English in their future working places efficiently.

References

- Acar, A. (2007). Standards and competence in English as an International Language pedagogy. *Asian EFL Journal*, 9(4), 39-53.
- Adamson, J. (2006). From EFL to content-based instruction: What English teachers take with them into the sociolinguistics lecture. *Asian EFL Journal*, 9. Retrieve June 15, 2007 from http://www.asian-efl-journal.com/pta_feb97_ja.pdf.
- Alptekin, C. (2002). Towards intercultural communicative competence in ELT. *ELT Journal*, 56(1), 57-64.
- Anderson, J. R. (1990). *Cognitive psychology and its implications*. (3rd edition). New York: W. H. Freeman.
- Anderson, J. R. (1993). Problem-solving and learning. *American Psychologist*, 48, 35-44.
- Brinton, D. M., Snow, M. A., & Wesche, M. B. (1989). *Content-based second language instruction*. New York: Newbury House.
- Brown, J. D. (1995). *The elements of language curriculum*. Boston, MA: Heinle & Heinle.
- Cantoni-Harvey, G. (1987). *Content-area language instruction: Approaches and strategies*. Reading, MA: Addison-Wesley.
- Chomsky, N. (1965). *Aspects of the theory of syntax*. The Hague: Mouton.
- Crandall, J. A. (Ed.). (1987). *ESL through content-area instruction*. Englewood Cliffs, NJ: Prentice Hall Regents.

- Crandall, J. A., & Kaufman, D. (Eds.). (1998). *Content-based instruction in tertiary settings*. Alexandria, VA: TESOL.
- Davies, S. (2003). Content-based instruction in EFL contexts. *The Internet TESL Journal*, 9(2). Retrieved October 12, 2005 from <http://iteslj.org/Articles/Davies-CBI.html>.
- Delanty, G. (2002). The governance of universities: What is the role of university in the knowledge society? *Canadian Journal of Sociology*, 27(2), 185-198.
- Finney, D. (2002). The ELT curriculum: A flexible model for a changing world. In J. C. Richards & W. A. Renandya (Eds.) *Methodology in language teaching: An anthology of current practice*, (pp. 69-79). Cambridge: Cambridge University Press.
- Fuller, S. (2003). The university: A social technology for producing universal knowledge. *Technology in Society*, 25, 217-234.
- Genesee, F. (1985). Second language learning through immersion: A review of U.S. programs. *Review of Educational Research*, 55(4), 541-561.
- Grabe, W. & Stoller, F. L. (1997). Content-based instruction: Research foundations. In M. A. Snow & D. M. Brinton (Eds.), *The content-based classroom: Perspectives on integrating language and content*, (pp. 5-21). New York: Longman.
- Hyde, M. (1998). Intercultural competence in English language education. *Modern English Teacher*, 7(2), 7-11.
- Johnson, K. & Johnson, H. (1999). (Eds.), *The Encyclopedic dictionary of applied linguistics: A handbook for language teaching*. Blackwell Publishing.
- Kuklinski, A. (2001). The role of universities in stimulating regional development and educating global elites. *Higher Education in Europe*, 26(3), 437-445.
- Lingley, D. (2006). A task-based approach to teaching a content-based Canadian Studies course in EFL context. *Asian EFL Journal*, 8(1), 97-122.
- Met, M. (1998). Curriculum decision-making in content-based second language teaching. In J. Cenoz & F. Genesee (Eds.), *Beyond bilingualism: Multilingualism and multilingual education* (pp. 35-63). Clevedon, UK: Multilingual Matters.

- Nunan, D. (2001). Syllabus design. In M. Celce-Murcia (Ed.), *Teaching English as a second or foreign language* (pp. 55-65). Boston: Heinle & Heinle.
- Nunn, R. (2005). Competence and teaching English as an International Language. *Asian EFL Journal*, 7(3), 61-74.
- Nunn, R. (2007). Redefining communicative competence for international and local communities. *Asian EFL Journal*, 9(4), 77-110.
- Oleksak, R. (2007). Ensuring America's place in the global economy by building language capacity in the schools. *Foreign Language Annals*, 40(1), 5.
- Raya, M. J. (2007). Developing professional autonomy: A balance between license and responsibility. Retrieved July 25 2008 from <http://www.learnerautonomy.org/raya2007.pdf>.
- Short, D. J. (2000). Using ESL standards for curriculum development. In M. A. Snow (Ed.) *Implementing ESL standards for pre-K-12 students through teacher education* (pp. 103-136). Alexandria, VA: TESOL.
- Singer, M. (1990). *Psychology of language: An introduction to sentence and discourse processing*. Hillsdale, NJ: L. Erlbaum.
- Snow, M. A. (2001). Content-based and immersion models for second and foreign language teaching. In M. Celce-Murcia (Ed.), *Teaching English as a second or foreign language* (pp. 303-318). Boston: Heinle & Heinle.
- Snow, M. A. & Brinton, D. M. (1988). Content-based language instruction: Investigating the effectiveness of the adjunct model. *TESOL Quarterly*, 22(4), 553-574.
- Snow, M. A. & Kamhi-Stein, L. D. (1997). Teaching academic literacy skills: A new twist on the adjunct model. *Journal of Intensive English Studies*, 11, 93-108.
- Spolsky, B. (1989). *Conditions for second language learning*. Oxford: Oxford University Press.
- Stoller, F. L. & Grabe, W. (1997). A six-T's approach to content-based instruction. In M. A. Snow & D. M. Brinton (Eds.), *The content-based classroom: Perspectives on integrating language and content*, (pp. 78-94). New York: Longman.

Tutunis, B. T. (2000). Content based academic writing. *The Internet TESL Journal*, 6(7). Retrieved August 25, 2001 from <http://iteslj.org/Articles/Tutunis-ContentBased.html>.

Wegrzecka-Kowalewski, E. (1997). Content-based instruction: Is it possible in high schools? In M. A. Snow & D. M. Brinton (Eds.) *The content-based classroom: Perspectives on integrating language and content* (pp. 319-323). New York: Longman.

Effects of L2 proficiency and gender on choice of language learning strategies by university students majoring in English

Adel Abu Radwan
adel_radwan@yahoo.com

Bio Data:

Dr. Radwan received his Doctorate in applied linguistics from Georgetown University in Washington, DC. He worked as an adjunct professor at George Mason University in Virginia, USA. He is currently an assistant professor at Sultan Qaboos University in Oman, where he teaches courses in psycholinguistics, language acquisition, and theoretical linguistics. Dr Radwan's chief interest is investigating the role of metacognition in language learning and pedagogy.

Abstract

This study investigates the use of language learning strategies by 128 students majoring in English at Sultan Qaboos University (SQU) in Oman. Using Oxford's (1990) Strategy Inventory for Language Learners (SILL), the study seeks to extend our current knowledge by examining the relationship between the use of language learning strategies (LLS) and gender and English proficiency, measured using a three-way criteria: students' grade point average (GPA) in English courses, study duration in the English Department, and students' perceived self-rating. It is as well a response to a call by Oxford to examine the relationship between LLSs and various factors in a variety of settings and cultural backgrounds (see Oxford, 1993). Results of a one-way analysis of variance (ANOVA) showed that the students used metacognitive strategies significantly more than any other category of strategies, with memory strategies ranking last on students' preference scale. Contrary to the findings of a number of studies (see e.g., Hong-Nam & Leavell, 2006), male students used more social strategies than female students, thus creating the only difference between the two groups in terms of their strategic preferences. Moreover,

ANOVA results revealed that more proficient students used more cognitive, metacognitive and affective strategies than less proficient students. As for study duration, the results showed a curvilinear relationship between strategy use and study duration, where freshmen used more strategies followed by juniors, then seniors and sophomores, respectively. Analysis of the relationship between strategy use and self-rating revealed a sharp contrast between learners who are self-efficacious and those who are not, favoring the first group in basically every strategy category. To find out which type of strategy predicted learners' L2 proficiency, a backward stepwise logistic regression analysis was performed on students' data, revealing that use of cognitive strategies was the only predictor that distinguished between students with high GPAs and those with low GPAs. The present study suggests that the EFL cultural setting may be a factor that determines the type of strategies preferred by learners. This might be specifically true since some of the results obtained in this study vary from results of studies conducted in other cultural contexts. Results of this study may be used to inform pedagogical choices at university and even pre-university levels.

Keywords: language learning strategies, metacognitive strategies, cognitive strategies, gender, proficiency, self-efficacy.

Introduction

Recently, the field of English as a Foreign Language (EFL) has witnessed a gradual shift among language educators towards student-centered approaches, leading to numerous studies investigating the impact of socio-cultural, psychological, cognitive, and affective variables on achievement in learning second/foreign languages (Nunan, 1988; Brown, 2000). Educators and researchers alike have considered these variables to be the source of discrepancies among second language learners in their learning outcomes. Increased interest in the role of these variables has led to numerous studies investigating individual learning styles and language learning strategies (LLS) and their relationship to success in learning a second language (Green & Oxford, 1995; Griffiths & Parr, 2001; Khalil, 2005;

Hsiao & Oxford, 2002; Wharton, 2000). Applied research on learning strategies had two major goals:

(1) [to] identify and compare the learning strategies used by more and less successful language learners, and (2) [to] provide instruction to less successful learners that helps them become more successful in their language study. (Chamot, 2001, pp. 25-26)

Overall, research on the use of learning strategies (see e.g., Dreyer & Oxford, 1996; Grenfell & Harris, 1999; Harris, 2003; Park, 1997; Wharton, 2000) suggests that language learners, whether consciously or unconsciously, utilize a variety of learning strategies. Successful language learners, however, employ more effective and diverse language learning strategies than less successful learners. Chamot (2004 p. 14) stated that strategic language learners possess “metacognitive knowledge about their own thinking and learning approaches, a good understanding of what a task entails, and the ability to orchestrate the strategies that best meet both the task demands and their own strengths.” Accordingly, in order to help second language learners in general and less successful learners in particular, researchers have recommended integrating strategy training into language curricula (Chamot & Kupper, 1989; Tyacke, 1991).

Despite the preponderance of research on language learning strategies within English as a second language context, there is an apparent paucity of this type of research within the Arabic EFL context. A very small number of studies (see, e.g., Al-Otaibi, 2004; El-Dib, 2004; Khalil, 2005; Kaylani, 1996; Shmais, 2003) examined the use of learning strategies by students in the Arab World, with only two studies (Al-Otaibi, 2004; El-Dib, 2004), investigating the use of LLSs in the Arab Gulf states of Saudi Arabia and Kuwait respectively. However, no research on LLSs has been conducted within the context of the

Gulf state of Oman. In this regard, and as Park (1997) remarked, there is a need for additional research in this area to determine whether the patterns of strategy use that exist in a particular linguistic setting are unique to that setting or common to all linguistic and cultural contexts.

In Oman, although Arabic is the official language, English has a special status; all government publications and correspondences are normally written in both languages. Moreover, English functions as a common language for communication with the large population of expatriates working in Oman. More importantly, English is a compulsory subject from the first grade, and it is the primary medium of instruction at the majority of universities. Despite its pivotal role, students at SQU, due to their limited proficiency in English, do not usually perform well in the English prerequisites, which negatively affects their performance in content-based courses. Since there is a substantial body of evidence to support the positive contribution of learning strategies to improvement in learning a foreign language, an examination of how students in the Omani context utilize these strategies is very critical.

Hence, the aim of the current study is to fill the gap in this area of research by exploring the use of language learning strategies used by Omani students, and it is, as well, a response to a call to examine the relationship between LLSs and various factors in a variety of settings and cultural backgrounds (see Oxford, 1993). More importantly given the small number of studies that have examined the correlation between strategy use and self-rating (one of the variables examined in this study), there is clearly a strong need for further research in this area. Hence this study investigates the use of learning strategies by students majoring in English at Sultan Qaboos University in Oman, exploring in particular the relationship between language learning strategies and a number of variables,

including gender, and language proficiency as measured by students' GPAs, study duration, and their perceived self-rating. Research indicates that these four variables, which are believed to have considerable influence on the process of language learning, contribute to considerable variability in strategy preferences (see, e.g., Green & Oxford, 1995; Lan & Oxford, 2003; Magogwe & Oliver, 2007).

2. Literature Review

Language learning strategies (LLS) are defined as “the conscious thoughts and actions that learners take in order to achieve a learning goal” (Chamot, 2004, p. 14). Through repeated use, these strategies become automatic. However, learners, if required, can call them to conscious awareness (Chamot, 2005). This, as Littlejohn (2008) points out, requires learners to develop some degree of meta-awareness that would enable them to think about their thinking, and then analyze any learning task and eventually choose the appropriate strategy required to accomplish that task.

Interest in LLSs emerged from studies that attempted to investigate the behavior and qualities of a good and successful language learner, with a view to teaching these qualities to less successful learners in order to make them more effective second language learners (Chamot et al., 1999; Grenfell & Harris, 1999; Harris, 2003). Research into LLSs started with the identification and description of learning strategies used by language learners (see e.g., Oxford, 1990; Rubin, 1987; Stern, 1975). Later, research explored the correlation between these strategies and other learner variables such as proficiency, gender, motivation, self-efficacy, self-rating, cultural background, and the like (see e.g., El-Dib, 2004; Green & Oxford, 1995; Hong-Nam & Leavell, 2006; Khalil, 2005; Magogwe & Oliver, 2007; Nisbet, Tindel & Arroyo, 2005; Shmais, 2003). More recently,

research investigated how other variables such as the task itself and the target language affect the selection and use of learning strategies (Chamot & Keatley, 2004).

Although researchers have proposed different classifications and conceptualizations of language learning strategies (see e.g., O'Malley & Chamot, 1990; Schmidt & Watanabe, 2001), Oxford (1990) developed the most comprehensive, detailed and systematic taxonomy of strategies to date. Contrary to O'Malley & Chamot (1990) who divided LLSs into three categories: cognitive, metacognitive, and social-affective, Oxford (1990) classified them into six groups: memory, cognitive, compensatory, metacognitive, affective, and social strategies. Based on this broad classification, Oxford (1990) designed a strategy assessment survey, the Strategy Inventory for Language Learning (SILL) to collect information about learners' use of language learning strategies. This survey was checked for reliability and validity, producing a high reliability coefficient (.86-.95 Cronbach's α) (Khalil, 2005). The fact that numerous studies established a significant relationship between strategies and language proficiency as measured in a variety of ways (grades, TOEFL scores, self-ratings, etc.) gives the instrument a high validity according to Oxford & Burry-Stock (1995). Woodrow (2005 p.91), however, questioned the reliability of the instrument, pointing out that while the scale has a high overall reliability, there is no "evidence to support the sixfold classification of LLSs in the SILL in the form of subclass reliabilities."

Despite this criticism, the SILL has been widely used to assess strategy use and to explore the effects of various variables on strategy preferences (see, e.g., Djigunovic, 2000; Dreyer & Oxford, 1996; Khalil, 2005; Park, 1997; Yang, 1999). In general, studies using the SILL have invariably shown significant variation in strategy preferences due to gender, and proficiency differences. Since this study explores the effects of these factors

on strategy preferences, the following discussion will be limited to studies that examined these variables.

Several studies have established the existence of gender differences in the use of language learning strategies. Green & Oxford (1995) found that females use strategies more frequently than males. Moreover, gender differences are reflected in the type of strategy used by males and females. Female learners tend to use more social learning strategies (Ehrman & Oxford, 1989), more conversational and input strategies (Oxford & Nyikos, 1989), and more memory and metacognitive strategies (Khalil, 2005) than their male counterparts. Contrary to these findings, Shmais (2003) did not report any differences in strategy use among university-level students as a result of gender difference. This could be attributed to the fact that the sample for this study was university English majors who are typically more aware of the process of learning a foreign language and of the strategies required to obtain proficiency than other groups. Similarly, Wharton (2000) did not reveal any effects for gender in both the number and types of strategy used by bilingual foreign language learners in Singapore. Again, this might be attributable to the language learning abilities of bilingual learners which may have nullified any gender difference.

Language learning strategies research has consistently established a positive link between language proficiency and strategy use (e.g., Khalil, 2005; Magogwe & Oliver, 2007; Park, 1997; Shmais, 2003), suggesting that more proficient learners usually use more strategies than less proficient learners. Researchers have utilized a multitude of ways to determine students' proficiency in the foreign language including standardized tests such as TOEFL (Arroyo, 2005), students' GPAs in English courses (Shmais, 2003), language achievement tests (O'Mara & Lett, 1990), language course grades and

placement examinations (Mullin, 1992), teachers' judgments about their students (Magogwe & Oliver, 2007), duration of study (Khalil, 2005), and self-ratings (Oxford & Nyikos, 1989). In the current study, proficiency has been determined using multiple measures, including students' GPAs in English courses, duration of study in the English program, and students' self-rating. Lack of an appropriate standardized language assessment test and the relatively large sample size were the main reasons for using these various measures. It is worth noting that students' GPAs are by far the most accurate indicator of students' proficiency in English, as they represent students' performance in the English courses taken at the Department as part of their degree requirements, including language skills course as well as literature and linguistics courses). In this regard, grading students' work in the English Department focuses on language and is normally based on given descriptors of high, good, middle and poor proficiency at various linguistic levels.

Investigating the relationship between strategy use of Korean university students and language proficiency, Park (1997) found a significant relationship between SILL learning strategies and English proficiency as measured by students' TOEFL scores. Additionally, the study showed that cognitive and social strategies were more predictive of TOEFL scores than other strategies. Similarly, Lan & Oxford (2003) found significant effects for language proficiency on Taiwanese elementary school EFL learners' use of metacognitive, cognitive, compensatory and affective strategies. Contrary to these studies, Nisbet, Tindel & Arroyo (2005) showed a minimal correlation (4% of the variation in TOEFL score) between learning strategies and proficiency and that only one category of learning strategies (metacognitive strategies) was significantly correlated with TOEFL score. Likewise, Shmais (2003) revealed that students with high proficiency, as measured

by their GPAs, differed from less proficient learners only in their use of cognitive strategies.

According to Chen (1990), the relationship between strategy use and proficiency does not always involve a simple linear correspondence between them. The study revealed a pattern whereby more proficient learners used fewer communication strategies although they used them more effectively than less proficient learners. Similarly, though Magogwe & Oliver (2007) revealed a trend in strategy use consistent with previous research, i.e. overall strategy use increases with proficiency, they showed that this relationship is a rather a curvilinear one, where proficiency influenced strategy use at the primary level but not at the secondary or the tertiary level. More importantly, as Mahlobo (2003) and Halbach (2000) point out, though many studies revealed strong correlations between strategy use and proficiency level, no claims of causality can be established in this type of research, i.e. “it cannot be determined whether the language proficiency comes before, after or concurrently with strategy use” (Magogwe & Oliver, 2007, p. 340).

A small number of studies investigated the link between the duration of English study and strategy use. Griffith (2003) reported a positive relationship between students' level in a private language school in New Zealand and frequency of language learning strategy use. Likewise, Oxford & Nyikos (1989) found that years of study significantly affect the use of learning strategies. In a study of adolescent learners of French L2, Ramierz (1986) reported similar results. Comparing high school students with university students, Khalil (2005) found that university students used more strategies than high school learners. This might be a result of the increased demands which proficient learners encounter while communicating in the target language.

Language learning strategies research has also examined the relationship between self-

efficacy beliefs and self-rating and strategy use. Bandura & Schunk (1981p. 31) defined self-efficacy as “people’s judgment of their capabilities to organize and execute courses of action required to attain designated types of performances.” Research in this area suggests that self-efficacy beliefs correlate positively with increased strategy use. For example, Pajares & Schunk (2001) found that learners who believed they were capable of performing certain tasks used more cognitive and metacognitive strategies than those who did not. According to Ching (2002), this result may be due to the fact that highly efficacious learners are more committed to learning L2 and working harder to avoid failure, and they usually link failure to insufficient effort or skills. In another study, Magogwe & Oliver (2007) demonstrated a slightly different pattern. The study showed that there was a positive, significant but weak relationship between self-efficacy beliefs and use of language learning strategies, which probably justifies further research in this area to examine thoroughly the effects of this variable.

3. The Study

The present study is a response to recommendations by many researchers (e.g., Green & Oxford, 1995; Park, 1997, among many others) for additional research to examine, using a reliable and valid instrument, the relationship between language learning strategies and several factors, which are believed to influence the process of learning a foreign language, in a variety of settings worldwide. Hence, the study investigates the use of language learning strategies by university students majoring in English at Sultan Qaboos University in Oman. More specifically, it explores the effects of gender, and proficiency, as measured by students’ GPAs, duration of study, and self-rating on reported strategy use by these students. The study attempts to answer the following research questions:

1. What are the most frequent language learning strategies used by Omani students majoring in English at SQU?
2. Are there any differences in strategy use as a result of gender differences?
3. Are there any differences among learners in strategy use due to proficiency differences as measured by their GPAs, duration of study, and self-rating?
4. Which strategies are predictive of (correlated with) L2 proficiency?

3.1 Participants

The questionnaire used in this study was distributed to 147 students majoring in English at Sultan Qaboos University. The questionnaire was distributed to regular classes that represent the different study durations (freshmen, sophomores, juniors, and seniors). Only 128 students returned their questionnaires completely answered, meeting all the study requirements. Since the questionnaires were distributed to regular classes and due to the demography of the English Department at SQU where the female-male ratio is approximately 2 to 1, favoring female students, the sample was not fairly balanced, consisting of 39 males and 89 females, whose ages ranged from 18 to 23 at the time of data collection. The subjects were freshmen (30), sophomores (21), juniors (39), and seniors (38). By the time of this study, all participants had received a minimum of 8 years of English as a foreign language instruction in the pre-University stage. Despite the relatively long period of study, the students' command of English is generally poor. As part of the students' bachelor degree requirements, they were required to complete 85 credit hours in English language and literature. It should be noted that, before joining the English Department, all applicants must pass an English language placement test, which assesses listening, reading, writing, and grammar. On the basis of their scores, students

who pass the test are either placed in a non-credit intensive English program for one whole semester, or are required to register for credit courses in the University’s Language Center, taking six courses in general language skills, writing, and reading, two courses each.

Since there is no standardized English test available in the University to administer to all students participating in this study, the students’ GPAs in English courses were used as a measure of their proficiencyⁱ. Students’ GPAs were divided into two groups B-and-Above and C-and-below. As for self-rating, students were asked to rate their English (listening, writing, reading, and speaking) as excellent, good, fair, or poor; however, only a very small number of respondents rated themselves as “excellent” and “poor”, thus resulting in the elimination of these two categories, for their data could not be used to form two independent groups for statistical purposesⁱⁱ. Thus, only two self-rating categories were retained: “Good” and “Fair.” Table (1) demonstrates the characteristics of the sample population.

Table 1. *Demographic description of participants*

	<i>n</i>	%
Gender		
Male	39	30.5
Female	89	69.5
Study Level		
Freshman	30	23.4
Sophomore	21	16.4
Junior	39	30.5
Senior	38	29.7
GPA		

B and above	57	44.5
C and below	71	55.5
Self-rating		
Good	48	47.5
Fair	80	62.5

3.2 Instrument

The study used Oxford’s (1990) Version 7.0 of the SILL, designed for EFL/ESL learners. Due to the high reliability of this survey, it has been used widely in more than 50 studies, assessing the frequency of strategy use by students from different linguistic and cultural backgrounds. The SILL uses a five-point Likert-type scale ranging from 1 (“Never or almost never true of me”) to 5 (“Always or almost always true of me”) (see the appendices for details of the questionnaire). The taxonomy of strategies consists of 50 statements about strategies used by language learners covering six broad categories of strategies, each represented by a number of items. Consider the following examples:

1. **Memory Strategies** (items 1-9): These strategies help learners remember, store and retrieve new information when there is a need for communication. This is achieved through using words in sentences, connecting words to mental picture of a word, grouping, and reviewing lessons frequently (e.g., representing sounds in memory, grouping, using physical responses).
2. **Cognitive Strategies** (items 10-23): These help learners understand and produce new language through practicing, summarizing, reasoning deductively, and analyzing (e.g., repeating, taking notes).
3. **Compensatory Strategies** (items 24-29): These strategies enable learners to use the language to overcome any limitations and gaps in their linguistic

knowledge through guessing, making up new words, and using circumlocution and synonyms (e.g., language switching, making gestures, and seeking help).

4. **Metacognitive Strategies** (items 30-38): These help learners control their own cognition and enable them maximize learning through monitoring their language use, planning, coordinating the learning process, and looking for opportunities to use the language (e.g., linking new information with old information, self-monitoring, planning, evaluating, and seeking practice opportunities)
5. **Affective Strategies** (items 39-44): These strategies help learners through lowering their anxiety levels, increasing their motivation, and controlling their emotions (e.g., discussing feelings with others, using music to lower anxiety).
6. **Social Strategies** (items 45-50): These help learners to interact, communicate, cooperate, and empathize with others to maximize learning (e.g., developing cultural understanding, cooperating with others).

This SILL questionnaire is used to identify the level of strategy use for each strategy or group of strategies. Along with the survey, Oxford (1990) developed a scale, which reflects the level of strategy usage: (1) high usage (3.5-5.0), (2) medium usage (2.5-3.4), and (3) low usage (1.0-2.4). The SILL was accompanied with a background questionnaire to collect demographic information about the students (see the appendices). Information collected included students' gender, GPAs in English courses, and duration of study in the English Department. The respondents were also asked to rate their English proficiency. The present study used the English version of the SILL with translation of difficult words into Arabic. The main reason for using the English version is that at the

time of the study the students had finished at least one year of intensive English in the Language Center at the University, thus, gaining a good command of the language. Moreover, English is the official language of instruction at SQU, and students are generally familiar with filling out all types of forms in that language. According to Shmais (2003), it is estimated that around 50 major studies utilized the English as a foreign language version of the questionnaire. Several researchers, however, (see, e.g., Khalil, 2005) used a translated version of the questionnaire to “avoid any problems participants could encounter in understanding the items and response scale” as a result of limited English proficiency (Khalil, 2005p. 110).

3.3 Data Collection Procedure

The SILL and background questionnaires were distributed to regular classes representing the different levels (freshmen, sophomores, juniors, and seniors) during the regular class meetings. The class instructors, who were informed about the nature of the questionnaire and its administration procedure, supervised the distribution process. Both questionnaires took an average of 35 minutes to finish under complete conditions of anonymity and confidentiality. Of the 147 distributed questionnaires, a total of 137 questionnaires were answered completely. All questionnaires that were not fully answered were disregarded. Moreover, nine students who categorized their English proficiency as “excellent” or “poor” were removed as explained before, thus leaving the total number of participants at 128 students.

Data analysis was carried out using the SPSS 15 statistical program to obtain descriptive and inferential statistics. First, means and standard deviations of the data were computed. Then, to determine any variation in strategy use due to gender, and English

proficiency (measured by GPA, duration of study, and self-rating), several analyses of variance (ANOVA) were conducted to determine whether there were any significant differences among learners with regard to strategy use. Where significant differences were located, the post-hoc Scheffé and LSD tests were used to determine the location of these differences. Finally, a stepwise backward regression was used to determine which of type of strategy was predictive of success in language learning.

4. Results

4.1 Overall strategy use

To answer the first research question about the most frequently used strategies, the students' data were submitted to a one-way analysis of variance (ANOVA). Analysis results revealed statistically significant differences ($F = 27.047$, $p = .000$) in the overall use of strategies by all participants. Table (2) presents the rank ordering of the strategies according to their frequency of use. As can be seen in the table, only metacognitive strategies ranked high in use ($M = 3.5$ - 5.0). The other strategies fell within the medium usage range ($M = 2.5$ - 3.4). These were compensatory strategies ($M = 3.38$), followed by cognitive strategies ($M = 3.34$), social strategies ($M = 3.24$), affective strategies (3.14) respectively, and finally the least preferred strategies, memory strategies ($M = 2.99$).

Table 2. *Descriptive statistics for the various learning strategies and F-test*

Variable	Mean	SD	Minimum	Maximum	Rank	F	Significance
Metacognitive	3.85	0.691	1.78	5.00	1	27.047	0.00
compensatory	3.38	0.608	1.33	4.83	2		
Cognitive	3.34	0.509	1.86	4.57	3		
Social	3.24	0.743	1.5	5.00	4		

Affective	3.14	0.685	1.5	5.00	5
Memory	2.99	0.560	1.33	4.33	6
Total	3.32	0.690	1.33	5.00	

In order to locate the multiple differences among the various strategy groups, a Scheffé post-hoc test was used. Multiple comparisons revealed the following significant differences between the different groups of strategies: (1) memory and cognitive in favor of cognitive ($p = .002$), (2) memory and compensatory in favor of compensatory ($p = .000$), (3) memory and metacognitive in favor of metacognitive ($p = .000$), (4) cognitive and metacognitive in favor of metacognitive ($p = .000$), (5) compensatory and metacognitive in favor of metacognitive ($p = .000$), (6) social and metacognitive in favor of metacognitive ($p = .000$), and (7) affective and metacognitive in favor of metacognitive ($p = .000$). Overall, these results show that metacognitive strategies were significantly used by L2 learners more than any other strategy. Table (3) shows the Scheffé test results.

Table 3. *Scheffé results for multiple comparisons among various strategy groups*

Strategy	Metacognitive	Compensatory	Cognitive	Social	Affective	Memory
Metacognitive		.000	.000	.000	.000	.000
Compensatory			.999	.692	.140	.000
Cognitive				.893	.315	.002
Social					.934	.098
Affective						.602
Memory						

Table (4) (see the appendices) ranks strategy use by individual strategies mean scores in a descending order from most to least used. The most frequently used strategy by all participants was a metacognitive strategy, “*I pay attention when someone is speaking in English*” ($M = 4.44$). In contrast, the least preferred strategy was affective, “*I write down my feelings in a language learning diary*” ($M = 2.16$). Among the top ten most used strategies were six metacognitive, two affective (“I encourage myself to speak in English even when I am afraid of making a mistake”, ($M = 3.77$), and “I try to relax whenever I feel afraid of using English”, ($M = 3.73$), one compensatory “If I can’t think of an English word, I use a word or a phrase that means the same thing”, ($M = 4.17$), and one cognitive “I watch English Language TV shows spoken in English or go to movies spoken in English”, ($M = 4.05$).

4.2 Use of strategy by gender

The second research question deals with the relationship between gender and the use of language learning strategies. Results of a one-way analysis of variance (ANOVA)ⁱⁱⁱ did not reveal any significant differences in the overall strategy use between male and female students ($F = .719, p = .39$). However, the results for strategy categories showed that males students, surprisingly, used significantly more social strategies than female students ($F = 3.811, p = .05$). As for the other categories, although there were no significant differences between the two groups, male students used slightly more memory, cognitive, and metacognitive strategies than female students, see Table (5).

Table 5. *Variation in strategy use by gender*

Strategies	Male		Female		<i>F</i>	Significance
	Mean	SD	Mean	SD		
MET	3.89	0.66	3.83	0.71	.196	0.66
COM	3.36	0.58	3.38	0.62	.039	0.85
COG	3.37	0.46	3.33	0.53	.186	0.67
SOC	3.43	0.75	3.15	0.73	3.81	0.05
AFF	3.14	0.54	3.15	0.74	.007	0.93
MEM	3.10	0.46	2.94	0.60	2.19	0.14
Total	3.39	0.41	3.31	0.52	.719	0.40

4.3 Strategy use by proficiency

Language proficiency in language learning strategies research has been determined in a multitude of ways. For instance, while a number of researchers used standardized tests to determine proficiency level (see e.g., Mullin, 1992; Nisbet, Tindall & Arroyo, 2005; Park, 1997), others relied on duration of study in English medium programs as a measure of proficiency (see e.g., Khalil, 2005; Shmais, 2003; Magogwe & Oliver, 2007). Yet, another group of researchers relied on perceived proficiency (self-efficacy and self-rating) as a measure of students level in the foreign language (see e.g., Oxford & Nyikos, 1989; Magogwe & Oliver, 2007; Shmais, 2003). A fourth measure used students' grade point average (GPA) in English courses to place them in different English proficiency categories (see, e.g., Shmais, 2003). In this study, multiple measures of proficiency were utilized, including students' GPAs (B and above, C and below), duration of study (freshmen, sophomores, juniors, and seniors), and self-rating (Good, Fair).

4.3.1 Use of strategy by students' GPAs

The third research question deals with the relationship between language proficiency and use of language learning strategies. The students were grouped into two groups: proficient students averaging B and above, which is relatively speaking close to 80%, and less proficient students, averaging C and below. It should be noted that for students to remain in good standing in the English department at SQU, their GPAs must not drop below C.

A one-way analysis of variance (ANOVA) revealed statistically significant differences between proficient students and less proficient students in the overall use of strategies ($F = 8.142, p = .005$). Overall, the proficient students used more strategies than the less proficient in all categories of strategies, see Table (6). In addition to that total use of strategies, significant results were obtained for cognitive strategies ($F = 9.350, p = .003$), metacognitive strategies ($F = 7.184, p = .008$), and affective strategies ($F = 4.350, P = .039$). The other groups of strategies did not show any significant differences between the two groups though the results, as can be seen in the students' means, clearly favor the more proficient students, see Table (6).

Table 6. *Variation in strategy use by students' GPAs*

Strategies	Proficient		Less Proficient		<i>F</i>	Significance
	Mean	SD	Mean	SD		
MET	4.03	0.59	3.71	0.74	7.18	.008
COM	3.47	0.52	3.30	0.67	2.59	.111
COG	3.49	0.48	3.22	0.51	9.35	.003
SOC	3.36	0.67	3.14	0.79	2.90	.091
AFF	3.28	0.70	3.03	0.66	4.35	.039

MEM	3.09	0.58	2.91	0.54	3.16	.078
Total	3.47	0.45	3.22	0.50	8.142	.005

4.3.2 Strategy use by study duration

To determine whether duration of study in the English Department has any effect on use of language learning strategies, students' answers to the questionnaire were submitted to a one-way ANOVA. As shown in Table (7), the freshmen group consistently used more strategies than any other group. However, data analysis revealed a significant difference among the four groups only in the use of affective strategies ($F = 2.82, P = .042$). To locate where the difference was, the Least Significance Differences test (LSD) was used^{iv}. The test showed that the Freshmen group used significantly more affective strategies than both the sophomore and senior groups ($p = .009, p = .03$, respectively)

Table 7. *Strategy use by study duration*

Strategies	Freshman		Sophomore		Junior		Senior		F	Significance
	Mean	SD	Mean	SD	Mean	SD	Mean	SD		
MET	4.04	0.57	3.78	0.59	3.89	0.74	3.70	0.73	1.52	.212
COM	3.60	0.51	3.31	0.74	3.27	0.55	3.35	0.63	1.95	.126
COG	3.45	0.45	3.15	0.51	3.36	0.52	3.33	0.54	1.46	.228
SOC	3.43	0.73	3.02	0.73	3.28	0.67	3.15	0.81	1.54	.207
AFF	3.39	0.65	2.89	0.51	3.20	0.74	3.03	0.69	2.82	.042
MEM	3.08	0.49	2.83	0.65	2.96	0.65	3.04	0.56	0.98	.403
Total	3.50	0.37	3.18	0.48	3.34	0.50	3.29	0.54	1.99	.119

4.3.3 Strategy use by self-rating

The background questionnaire included a self-rating scale used to determine the students' own judgment of their proficiency in English. The students had to rate their proficiency in English using one of four options: excellent, good, fair, poor. Only a very small number of students rated themselves as "excellent" or "poor", resulting in, as mentioned before, eliminating them from the study. Students' data were submitted to a one-way analysis of variance (ANOVA) to determine if there were any differences between the two groups, "good" and "fair". ANOVA results showed that, for all strategies, those students who rated their English language proficiency as "good" used significantly more strategies than the other group, see Table (8). Moreover, the "Good" group used every category of strategy more frequently than the "Fair" group.

Table 8. *Variation in strategy use by students' self-rating*

Strategies	Good		Fair		<i>F</i>	Significance
	Mean	SD	Mean	SD		
MET	4.02	0.61	3.57	0.73	13.75	.000
COM	3.52	0.55	3.13	0.63	13.46	.000
COG	3.48	0.46	3.11	0.51	17.24	.000
SOC	3.42	0.71	2.93	0.69	14.66	.000
AFF	3.29	0.69	2.90	0.60	10.27	.002
MEM	3.14	.0.52	2.75	0.54	16.36	.000
Total	3.49	0.45	3.08	0.46	24.14	.000

4.4 Strategies predictive of L2 proficiency

To answer question four and find out which category of learning strategies is predictive of (correlated with) L2 proficiency as measured by students' GPAs, a backward stepwise

logistic regression analysis was performed on students' data. The six categories of strategies were specified as the predictor variables with students' GPAs as the criterion (dependent) variable. A test of the full model with all six predictors against a constant-only model was statistically reliable, $\chi^2(1, 128) = 9.172, p = .002$, indicating that the predictors, as a set, reliably distinguished between students with high GPAs and those with low GPAs. The regression model revealed that only cognitive strategies were significantly correlated with students' GPAs ($B = 1.120, p = .004$). In the second step of the analysis, the compensatory strategies variable was removed, followed by memory strategies, affective strategies, social strategies, and metacognitive strategies respectively. Table (9) shows regression coefficients, Wald statistics, exponential B (odds ratio), and 95% confidence intervals for odds ratio for each of the six predictors in the first step of the regression.

Table 9. *Backward logistic regression analysis of students' GPAs as a function of strategy use*

Variables	<i>B</i>	Wald Test	Odds Ratio	95% confidence Interval for Odds Ratio	
				<i>Lower</i>	<i>Upper</i>
Memory strategies	-1.161	.120	.852	.343	2.114
Cognitive Strategies	.966	2.463	2.626	.786	8.770
Compensatory strategies	.032	.007	1.033	.495	2.154
Metacognitive strategies	.310	.477	1.363	.566	3.281
Affective Strategies	.225	.415	1.252	.632	2.481
Social Strategies	-.214	.354	.808	.400	1.632

5. Discussion

Data analysis reveals several significant findings. First, in general Omani students favored metacognitive strategies over all other strategies. Second, there were no significant differences between males and females in the overall use of strategies, although analysis results showed that male students used significantly more social strategies than their female counterparts. Third, the more proficient students differed from the less proficient learners in several ways: (1) they used more overall strategies; (2) they used more cognitive, metacognitive, and affective strategies than the less proficient learners. Fourth, freshmen in general used more strategies followed by juniors, seniors, and sophomores. However, the differences between the four groups were significant only with regard to affective strategies, where freshmen used more of these strategies than both sophomores and seniors. Fifth, self-rating was evidently the strongest factor distinguishing between students. Results demonstrated that students who perceived themselves as proficient users of the language (the “Good” group) used significantly more strategies than the other group. Finally, though students with high GPAs differed from students with low GPAs in the overall use of strategy, and in the use of cognitive, metacognitive, and affective strategies, only cognitive strategies, in a regression model analysis, was predictive of students’ GPAs.

5.1 Overall Strategy use

In general, students in this study reported medium to high use of SILL learning strategies with a preference for metacognitive strategies, which reflects the students’ endeavor to become proficient in the target language. Among the top ten strategies used by all participants, six belong to the metacognitive strategies set. These strategies are essential

for successful language learning, since they, as pointed out by Oxford (1990), help learners coordinate and maximize their own learning process through monitoring and evaluating language use, planning, focusing, organizing, and seeking opportunities to use the language. Considering that maintaining a good GPA throughout the course of study at the Department is a requirement for continuation in the program, no wonder that the majority of the students are instrumentally motivated to learn the language. Accordingly, the use of the various strategies subsumed under the metacognitive heading seems for all of them to be an indispensable requirement if they are to graduate from the department with a degree to qualify them to teach English. The relative high use of metacognitive strategies has also been reported in other studies, including Hong-Nam & Leavell, 2006; Magogwe & Oliver, 2007; Nisbet, Tindall & Arroyo, 2005, among others⁷.

Among the five least favored strategies (low use: 2.4 or below) were one compensatory strategy (I make up new words if I did not know the right ones I English), three memory strategies (I use rhymes to remember new English words; I use flashcards to remember new English words; I physically act out new English words), and one affective strategy (I write down my feelings in a language learning diary), occupying the final position. Given that writing diaries is not an exceptionally popular practice in the Arab World, students' disdain from this strategy seems justifiable. The fact that memory strategies were the least favored strategies is quite surprising considering that the educational system in most of the Arab countries emphasizes rote memorization. This relatively surprising result may reflect students' displeasure with the conservative educational methods and techniques and their quest for alternative strategies that depart from the conventional didactic strategies to more communicatively oriented strategies. Moreover, this result, obtained also by other researchers (e.g., Al-Otaibi, 2004; Khalil, 2005; Hong-Nam & Leavell,

2006), may underscore the students' recognition that excelling in learning a foreign language requires actively involving themselves in the learning process, seeking opportunities to use the language, cooperating with their peers, etc.

5.2 Strategy use by gender

Contrary to most research findings (e.g., Hong-Nam & Leavell, 2006; Khalil, 2005; Oxford, 1990; Oxford & Ehrman, 1995), male students in this study used more learning strategies than did female students though the differences between the two groups were not significant in most cases. The only significant difference between males and females was in their use of social strategies. While one expects female students to use more social strategies than male students as they generally excel at establishing strong relationships and building vast social networks (Khalil, 2005), this was not borne out in this study. A logical explanation for this result can be attributed to the cultural background of these students. Omani society is organized into tribes and until recently the tribes, which consist of large extended kin groups that interact frequently with each other, were of major political and social importance (Wilkinson, 1987). Men in particular have to develop extremely good social skills to operate in this context, and even though Oman now has a centralized and modern government, the tribal units are still central to the organization of Omani society. Moreover, the conservative nature of culture, customs, and habits prevents females in the Arab World socializing and establishing relationships outside their immediate circles, which is a prerequisite for excelling in acquiring a foreign language within any communicatively oriented approach to language learning.

5.3 Strategy use by English proficiency

Research examining the use of learning strategies by different proficiency groups showed a linear relationship between the two factors (e.g. Green and Oxford, 1995; Khalil, 2005; Wharton, 2000). The present study revealed a complex picture due to the multitude of measures used to assess language proficiency. In terms of students' GPAs, the study showed that the more proficient students used significantly more overall strategies, cognitive strategies, metacognitive strategies and affective strategies than the less proficient learners, concurring thus with result obtained by Nisbet, Tindall & Arroyo (2005). These results show that the proficient learners seem to be more aware of their language needs; thus, they tend to utilize strategies that will help them master the target language through practicing, reasoning, analyzing, as well as strategies that allow them to control their own learning through planning and evaluating learning. Moreover, these learners exercise a great deal of control over their emotions and attitudes through lowering their anxiety levels and increasing their motivation levels. In this regard, Oxford & Nyikos (1989, p. 295) remark that "language proficiency can be either effects or causes of strategy use." They add that "use of appropriate strategies leads to enhanced actual and perceived proficiency, which in turn creates high self-esteem, which leads to strong motivation, spiraling to still more use of strategies, great actual and perceived proficiency, high self-esteem, improved motivation, and so on."

Many studies showed a positive relationship between strategy use and study duration. Khalil (2005), for example, revealed that university-level students reported higher use of almost all strategy categories than did high school students, which suggests, as pointed out by Magogwe and Oliver (2005p. 346) that "many strategies may be developmentally acquired." Hence, the longer the duration of language study is, the more are the strategies

used by learners. This study, however, showed a curvilinear relationship between duration of study and strategy use. In the overall use of strategies, freshmen showed a high use of strategies ($m= 3.5$), sophomores, on the other hand, demonstrated medium use ($m= 3.18$), and so did juniors ($m= 3.34$) and seniors ($m= 3.29$). A similar result was obtained by Phillips (1991). Likewise, Hong-Nam & Leavell (2006, pp. 410-11) explain a similar result by indicating that once learners reach advanced proficiency levels, “their need to consciously administer and [become] deliberate about their learning choices becomes less necessary.” Moreover, “[a]dvanced learners’ habitual and successful application of language strategies may be soon internalized that they do not report what has become for them an automated process;” accordingly they reported less strategy use than the freshmen group.

As for the different categories of strategies, learners differed significantly in their use of affective strategies. It might be the case that the freshmen students in this study realized that learning a second language requires exercising considerable control over their emotions, motivation, and attitudes; thus, they might have worked on lowering their anxiety levels and increasing their motivational levels. This type of control over emotions is critical for these learners considering that their performance in their first year determines their status in the Department. Failure to control their emotions might lead to poor performance, which might eventually result in dismissal from the program. As for the other strategies, the different groups reflected the same pattern and order of strategy use, whereby metacognitive strategies occupied the top of the list followed by compensatory strategies, cognitive strategies, social strategies and affective strategies. This means that, for all groups, the duration of study effects were relatively marginal.

As for the third proficiency parameter (self-rating), students’ perception of their

linguistic capabilities seems to be the strongest factor distinguishing learners in their use of strategies. Increase in students' self-rating translated into significant increases in their use of all types of strategies. Pajares and Schunk (2001) explain such a result by suggesting that high self-efficacy beliefs are associated with high achievement and indicated that high strategy use is an attribute of a good language learner. In this study, learners with high self-rating reported high strategy use of three categories (metacognitive, compensatory, and cognitive); on the other hand, the fair learners reported high use of only metacognitive strategies. This result seems to show that the higher the students' self-perceived proficiency, the more likely they were to exercise control over their learning than the students with lower self-perceived proficiency. Accordingly, they are more likely to use strategies that would help them organize, plan, monitor, focus, and evaluate their learning. Moreover, they are more likely to use compensatory strategies than the other group, since these strategies generally help learners use the language to overcome any limitations and gaps in their linguistic knowledge through guessing, making up new words, and using circumlocution and synonyms.

5.4 Strategies predictive of success in L2

The regression model used to analyze students' data revealed that cognitive strategies are significantly correlated with proficiency, measured using students' GPAs in English courses. This result concurs with Park (1997) who showed that social and cognitive strategies predicted TOEFL scores among Asian students. Examples of the cognitive strategies include item 13 "*I use the English words I know if different ways,*" item 20 "*I try to find patterns in English,*" and item 22 "*I try not to translate word for word.*" These

strategies and many others subsumed under the cognitive heading help students practice, analyze, revise their language, create structures, and look for opportunities to use the language. These strategies are critical for success in learning English, and they are the type of strategies language learners in the English Department usually focus on.

6. Conclusion

In general, while some of the results reported here are consistent with the findings of previous research on strategy use (Green & Oxford, 1995; O'Malley & Chamot, 1990; Park, 1997), the current study reveals a more complex pattern of strategy use than has been observed in previous studies. Like previous research (e.g., Khalil, 2005; Shmais, 2003), the current results demonstrate unequivocally that English students at SQU were aware of the significance of learning strategies to the development of their proficiency in English; the students used learning strategies with a medium to high frequency, with metacognitive strategies ranking highest among all strategies. In all comparisons, metacognitive strategies, which help student direct, organize, and plan their learning, were favored by all students over all other strategies. A similar result was obtained by Shmais (2003) and Hong-Nam & Leavell (2006), showing that students overall prefer metacognitive strategies over other strategies, and the least preferred strategies were affective and memory strategies. The use of metacognitive strategies must be supported in curricula design, especially through the beginning stages of learning a second/foreign language, where obtaining some type of declarative knowledge is critical in order to create “heightened understanding of the *what* and *how* of successful language learning” (Hong-Nam & Leavell, 2006, p. 412).

Unlike many previous studies (see e.g., Khalil, 2005), this study did not reveal any

significant differences between male and female students except in their use of social strategies, unexpectedly favoring male students. While one would expect female students to utilize their superior social skills to establish social networks that would assist them in the process of learning, male students showed higher preference for these strategies than their female counterparts, which means they were more likely to interact, cooperate, and empathize with others. This, as explained previously, is probably due to disparity in social expectations placed on both groups.

With regard to the relationship between strategy use and proficiency, the results showed that proficiency had a main effect on the overall strategies used by learners as well as on three categories, namely cognitive, metacognitive, and affective strategies, favoring proficient students. As for the effects of study duration, it only had a significant effect on the use of affective strategies, showing higher use of these strategies by freshmen students. Of all factors examined in this study, self-rating was the strongest predictor of differences among learners. Unlike Magogwe & Oliver (2007), who showed a weak relationship between self-efficacy beliefs and the overall use of language learning strategies, this study showed that students with high self-rating reported more strategy use than those with fair self-rating. Future research in this area should focus on the interaction between self-rating and proficiency to investigate various combinations and how they relate to strategy use; for example, how someone with low proficiency and high self-rating will compare to someone who has low high proficiency but low self-rating and so on.

The results of the present study highlight the importance of incorporating strategy training into L2 classroom instruction and into curriculum design. Accordingly, both students and teachers are required to develop awareness of these various strategies

through appropriate instruction or training for both groups. Becoming aware of their preferred learning strategies might help learners become more independent and effective in approaching the task of learning a second language. Oxford (2001 p. 1, as cited in Nisbet et al., 2005) emphasizes the importance of such autonomy by stating that learning strategies “are aimed at self-management in language learning and self-reliance in language use.” Moreover, explicit training in strategy use might be necessary so as to allow students at different levels and those with different proficiency levels and learning styles to practice a wide range of these strategies that are “appropriate to different instructional task and activities that constitute an essential part of the classroom L2 experience” (Khalil, 2005 p. 115). This stipulates that learners at different levels require variable degrees of teacher’s intervention in the learning process. Accordingly, for less proficient learners, high levels of explicit instruction in strategy use might be essential to raise learners’ awareness of the significance of developing their strategic competence. Thus, the teacher’s role is critical. “Effective teachers should consider each learning task from a novice’s perspective and scaffold the learning process through purposeful strategy choice” (Hong-Nam & Leavell, 2006, p.412). Cohen, Weaver, & Li (1995, p.29) emphasized that approach by pointing out that “explicitly describing, discussing, and reinforcing strategies in the classroom can have a direct payoff on student outcomes.” As learners develop some explicit knowledge of various strategies, and as they become relatively autonomous, the teacher’s role changes to a facilitator who selects the strategies, which are appropriate for various tasks and are suitable for different individuals. This eventually might help learners gain more independence and develop confidence, which, in turn, will enhance their linguistic abilities. To achieve this, it is critically important to provide teachers with the proper training in strategy assessment and instruction, through

the systematic introduction and reinforcement of learning strategies that help learners improve their proficiency in the target language (Cohen, Weaver & Li, 1995).

In terms of curriculum design and material preparation, researchers have often recommended that strategy training be integrated into language curriculum (see e.g., Khalil, 2005; Oxford, 1990; Tyacke, 1991). Therefore, teachers and materials' developers should incorporate a variety of tasks and activities, which target strategies that teachers view as critical for success in learning a second language. The fact that students with different proficiency levels utilize different learning strategies must guide the development of instructional materials (Chamot & O'Malley, 1996). These materials, according to Khalil (2005, p. 115), should provide "students with further opportunities to practice a wide variety of strategies that are appropriate to different instructional tasks and activities that constitute an essential part of the classroom L2 experience." Moreover, instructional materials might be developed to target certain strategies which research finds essential for success in learning a second language. In this regard, Ellis and Sinclair (1989) advocated the use of organization strategies, risk-taking strategies, and personal strategies in content-based instruction. Chamot and O'Malley (1996) also developed instructional materials which incorporated explicit instruction in learning strategies.

In conclusion, a number of variables were considered in assessing Omani students' use of learning strategies. While these variables may explain to some extent discrepancies in strategy use among various groups of learners, other factors that might affect the use of LLS such as the role of beliefs, social and cultural background, motivation, attitude, personality, etc. must be as well studied to find out their interaction with strategy use. Moreover, further research in this area is critical to the development of teacher training and student instruction in order to base these two components on firm theoretical and

empirical foundations.

7. Limitations of this study

One of the limitations of the study is lack of a standardized English proficiency test. To determine students' proficiency, the study utilized different constructs such as students' GPAs, students' self-rating, and duration of study, which may account for the apparent contrasts in the results reported for each measure. Another limitation is the complete reliance on SILL to determine strategies used by students. While this quantitative measure is very beneficial, the students "may not remember the strategies they have used in the past, may claim to use strategies that in fact they do not use, or may not understand the strategy descriptions in the questionnaire items (Chamot, 2004 p.15). Accordingly, the SILL should be supplemented with other techniques such as think-aloud protocols concurrent with a specific learning task, written diaries, stimulated recall interviews, and other methods which might provide richer and more sample-specific data.

References

- Al-Otaibi, G. (2004). *Language learning use among Saudi EFL students and its relationship to language proficiency level, gender and motivation*. Doctoral Dissertation: Indiana University of Pennsylvania, Indiana, PA.
- Bandura, A. & Schunk, D. (1981). Cultivating competence, self-efficacy, and intrinsic interest through proximal self-motivation. *The Journal of Personality and Social Psychology*, 41, 586-598.
- Brown, D. (2000). *Principles of language learning and teaching*. NY: Longman.
- Chamot, A. (2005). *The Cognitive Academic Language Learning Approach (CALLA)*:

- An update. In P. Richard-Amato & M. Snow (Eds.), *Academic success for English language learners: Strategies for K-12 mainstream teachers* (pp. 87-101). White Plains, NY: Longman.
- Chamot, A. (2004). Issues in language learning strategy research and teaching. *Electronic Journal of Foreign Language Teaching*, 1(1), 14-26.
- Chamot, A. (2001). The role of learning strategies in second language acquisition. In M.P. Breen (Ed.), *Learner contributions to language learning: New directions in research*, pp. 25-43. Harlow, England: Longman.
- Chamot, A., & Keatley, C. (2004, April). *Learning strategies of students of less commonly taught languages*. Paper presented at the 2004 Annual Meeting of the American Educational Research Association, San Diego, CA.
- Chamot, A., Barnhardt, S., El-Dinary, P., & Robbins, J. (1999). *The learning strategies handbook*. White Plains, NY: Addison Wesley Longman.
- Chamot, A. & O'Malley, J. (1996). Implementing the cognitive academic language learning approach (CALLA). In R. Oxford (Ed.), *Language learning strategies around the world: Cross-cultural perspectives* (pp. 61-74). Honolulu: University of Hawaii at Manoa.
- Chamot, A. & Kupper, L. (1989). Learning strategies in foreign language instruction. *Foreign Language Annals*, 22(1), 13-24.
- Chen, S. (1990). A study of communication strategies in interlanguage production by Chinese EFL learners. *Language Learning*, 40, 155-187.
- Ching, L. (2002). Strategy and self-regulation instruction as contributors to improving students' cognitive model in an ESL program. *English for Specific Purposes*, 13, 261-289.

- Cohen, A.D., Weaver, S., & Li, T-Y. (1998). The impact of strategies-based instruction on speaking a foreign language. In A. Cohen, *Strategies in learning and using a second language* (pp. 107-156). London: Longman.
- Djigunovic, J. (2000). Language learning strategies and affect. *CLS Occasional Papers*, 59 (Autumn), 1-28.
- Dreyer, C. & Oxford, R. (1996). Learner variables related to ESL proficiency among Afrikaans speakers in South Africa. In Oxford, R. (Ed.), *Language learning strategies around the world: Cross-cultural perspectives* (pp. 61-74). Honolulu: University of Hawaii at Manoa.
- Ehrman, M. & Oxford, R. (1989). Effects of sex differences, career choice, and psychological type on adult language learning strategies. *The Modern Language Journal*, 73, 1-12.
- El-Dib, M. (2004). Language learning strategies in Kuwait: Links to gender, language level, and culture in a hybrid context. *Foreign Language Annals*, 37, 85-95.
- Ellis, H., & Sinclair, B. (1989). *Learning to learn English: A course in learner training*. Cambridge, UK: Cambridge University Press.
- Green, J. & Oxford, R. (1995). A closer look at learning strategies, L2 proficiency, and gender. *TESOL Quarterly*, 29(2), 261-297.
- Grenfell, M. & Harris, V. (1999). *Modern languages and learning strategies: In theory and practice*. London: Routledge.
- Griffiths, C. (2003). Patterns of language learning strategy use. *System*, 31(3), 367-383.
- Griffiths, C., & Parr, J. (2001). Language-learning strategies: Theory and perception. *ELT Journal*, 53(3), 247-254.
- Halbach, A. (2000). Finding out about students' learning strategies by looking at their

- diaries: a case study. *System*, 28, 85-96.
- Hatch, E. & Lazaraton, A. (1991). *The research manual: Design and statistics for applied linguistics*. New York: Newbury House Publishers.
- Hon-Nam, K. & Leavell, A. (2006). Language learning strategy use of ESL students in an intensive English learning context. *System*, 34, 399-415.
- Hsiao, T. & Oxford, R. (2000). Comparing theories of language learning strategies: a confirmatory factor analysis. *The Modern Language Journal*, 86(3), 368-383.
- Kaylani, C. (1996). The influence of gender and motivation on EFL learning strategy use in Jordan. In R.Oxford (Ed.), *Language learning strategies around the world: Cross-cultural perspectives* (pp.75-88). Honolulu, HI: University of Hawaii Press.
- Khalil, A. (2005). Assessment of language learning strategies used by Palestinian EFL learners. *Foreign Language Annals*, 38(1), 108-119.
- Lan, R. & Oxford, R. (2003). Language learning strategy profiles of elementary school students in Taiwan. *IRAL*, 41, 339-379.
- Littlejohn, A. (2008). Digging deeper: learners' disposition and strategy use. In Cane, G. (Ed.), *strategies in language learning and teaching*. Singapore. RELC.
- Magogwe, J. & Oliver, R. (2007). The relationship between language learning strategies, proficiency, age and self-efficacy beliefs: A study of language learners in Botswana. *System*, 35, 338-352.
- Mahlobo, E. (2003). The relationship between standardized test performance and language learning strategies in English as a second language: a case study. *Journal for Language Teaching*, 37(2), 164-176.
- Mullin, P. (1992). Successful English language learning strategies of students enrolled at the Faculty of Arts, Chulalongkorn University, Bangkok, Thailand. *Dissertation*

- Abstracts International*, 53(06), 1829A. (UMI No. 9223758).
- Nisbet, D., Tindall, E. & Arroyo, A. (2005). Language learning strategies and English proficiency of Chinese university students. *Foreign Language Annals*, 38(1), 100-107.
- Nunan, D. (1988). *The learner-centered curriculum*. Cambridge: University Press.
- O' Malley, J. & Chamot, A. (1990). *Learning strategies in second language acquisition*. Cambridge, England: Cambridge University Press.
- O'Mara, F., & Lett, J. (1990). *Foreign language learning and retention: Interim results of the language skill change report*. Advanced technology, Inc.: Reston, VA.
- Oxford, R. (2001, March). *Language learning strategies, proficiency, and autonomy: What they mean in the new millennium*. Symposium conducted at the meeting of the Desert Language and Linguistics Society, Provo, Utah.
- Oxford, R. (1993). Research on second language learning strategies. *Annual Review of Applied Linguistics*, 13, 175-187.
- Oxford, R. (1990). *Language learning strategies: What every teacher should know*. New York: Newbury House.
- Oxford, R. & Burry-Stock, J. (1995). Assessing the use of language learning strategies worldwide with the ESL/EFL version of the Strategy Inventory for Language Learning (SILL). *System*, 23(1), 1-23.
- Oxford, R. & Ehrman, M. (1995). Adults' language learning strategies in an intensive foreign language program in the United States. *System*, 23(3), 359-386.
- Oxford, R. & Nyikos, M. (1989). Variables affecting choice of language learning strategies: A synthesis of studies with implications for strategy training. *System*, 17, 235-247.

- Pajares, F. & Schunk, D. (2001). Self-efficacy beliefs and school success: Self-efficacy, self-concept, and school achievement. In: *Perception*, pp 239-266. London: Ablex Publishing.
- Ramirez, A. (1986).. Language learning strategies used by adolescents studying French in New York schools. *Foreign Language Annals*, 19(2), 131-141.
- Park, G. (1997). Language learning strategies and English proficiency in Korean university students. *Foreign Language Annals*, 30(2), 211-221.
- Rubin, J. (1987). Learner strategies: Theoretical assumptions, research history and typology. In A. Wenden & J. Rubin (Eds.), *Learner strategies in language learning*, pp. 15-30. New Jersey: Prentice/Hall International.
- Rubin, J. (1975). What the “good language learner” can teach us. *TESOL Quarterly*, 9, 41-51.
- Schmidt, R. & Watanabe, Y. (2001). Motivation, strategy use and pedagogical preferences. In Z. Dornyei & R. Schmidt (Eds.), *Motivation and second language acquisition*, pp. 313-352. Honolulu: University of Hawaii.
- Shmais, W. (2003). Language learning strategy use in Palestine. *TESL-EJ*, 7(2). Retrieved from <http://tesl-ej.org/ej26/a3.html>.
- Stern, H. (1975). What can we learn from the good language learner? *Canadian Modern Language Review*, 31, 304-318.
- Tyacke, M. (1991). Strategies for successes: Bringing out the best in a learner. *TESL Canada Journal* 8(2), 45-56.
- Wharton, G. (2000). Language learning strategy use of bilingual foreign language learners in Singapore. *Language Learning*, 50(2), 203-244.
- Wilkinson, J. (1987). *The Imamate Tradition of Oman*. Cambridge: Cambridge

- University Press. *Foreign Language Annals*, 38(1), 90-99.
- Woodrow, L. (2005). The challenge of measuring language learning strategies. *Foreign Language Annals*. Vol. 38, 1, 90-98.
- Yang, N. (1999). The relationship between EFL learners' beliefs and learning strategy use. *System*, 27, 515-535.

Notes

- i Many studies (see, e.g., Shmais, 2003) use GPA as a predictor of language proficiency. The assumption in such studies is that the higher the GPA, the more proficient the learner in the foreign language is.
- ii Hatch and Lazaraton (1991: 310) maintain that when the cell sizes are unbalanced, "the assumption of equal variance may be violated."
- iii It should be noted that ANOVA gives exactly the same results as the *t*-test.
- iv Despite the significant difference among the four groups ($F = 0.98$, $p = .042$), Schaffé post-hoc test showed tendency of difference, but it was not significant. As a result, LSD, which can locate differences when the mean differences are not big.
- v In step 1 of the regression analysis, the exponential B indicates that the odds of having a higher GPA increased by a factor of 2.626 with every point on a five-point Likert scale in the case of cognitive strategies. This ration jumped to 3.065 in the final step.

Appendices

Language Learning Strategies Questionnaire

Dear student,

This questionnaire is an attempt to explore the language learning strategies used by the students of the English Department at SQU. Your participation in this study and honest responses to the questionnaire items are highly appreciated. Please be assured that your information will remain strictly confidential and will be used solely for the purpose of this research project.

Part One:

1. Name: _____ ID: _____
2. Gender: Male Female
2. Year in the English Department: First Second
 Third Fourth
3. Program: Education Art Translation
4. Overall GPA: _____
5. GPA in English courses: _____
6. Language(s) spoken at home: _____
7. How do you rate your proficiency in English? Excellent Good Fair Poor

B. Directions:

This form of the Strategy Inventory for Language Learning (SILL) is for students of English as a foreign language. You will find statements about learning English. Please read each statement carefully. Then, next to each statement, make a check mark (✓) in the answer box that tells

how true of you the statement is.

1. **Never** or almost never true of me (لاينطبق علي أبدا)

2. **Seldom** true of me (نادرا ما ينطبق علي)
3. **Sometimes** true of me (أحيانا ينطبق علي)
4. **Often** true of me (عادة ينطبق علي)
5. **Always** or almost always true of me (دائما ينطبق علي)

		1	2	3	4	5
		Never	Seldom	Sometimes	Often	Always
1	I think of relationships between what I already know and new things I learn in English					
2	I use new English words in a sentence so I can remember them.					
3	I connect the sound of a new English word and an image or picture of the word to help me remember the word.					
4	I remember a new English word by making a mental picture (صورة ذهنية) of a situation in which the word might be used.					
5	I use rhymes (كلمات مسجعة) to remember new English words.					
6	I use flashcards (البطاقات) to remember new English words.					
7	I physically act out (أقوم بتمثيل) new English words.					
8	I review English lessons often.					
9	I remember new English words or phrases by remembering their location on the page, on the board, or					

	on a street sign.					
10	I say or write new English words several times.					
11	I try to talk to native English speakers.					
12	I practice the sounds of English.					
13	I use the English words I know in different ways.					
14	I start conversations in English.					
15	I watch English Language TV shows spoken in English or go to movies spoken in English					
16	I read for pleasure in English.					
17	I write notes, messages, letters, or reports in English.					
18	I first skim an English passage (read over the passage quickly) then go back and read carefully.					
19	I look for words in my own language that are similar to new words in English.					
20	I try to find patterns (أنماط متكررة) in English.					
21	I find the meaning of an English word by dividing it into parts that I understand.					
22	I try not to translate word-for-word.					

23	I make summaries of information that I hear or read in English.					
24	To understand unfamiliar words, I make guesses.					
25	When I can't think of a word during a conversation in English, I use gestures (اشارات) .					
26	I make up (أخترق) new words if I don't know the right ones in English.					
27	I read English without looking up every new word.					
28	I try to guess what the other person will say next in English.					
29	If I can't think of an English word, I use a word or a phrase that means the same thing.					
30	I try to find as many ways as I can to use my English.					
31	I notice my English mistakes and use that information to help me do better.					
32	I pay attention when someone is speaking English.					
33	I try to find out how to be a better learner of English.					
34	I plan my schedule so I will have enough time to study English.					
35	I look for people I can talk to in English.					
36	I look for opportunities to read as much as possible in English.					

37	I have clear goals for improving my English skills.					
38	I think about my progress in learning English.					
39	I try to relax whenever I feel afraid of using English.					
40	I encourage myself to speak in English even when I am afraid of making a mistake.					
41	I give myself a reward or treat when I do well in English.					
42	I notice if I am tense or nervous when I am studying or using English.					
43	I write down my feelings in a language learning diary.					
44	I talk to someone else about how I feel when I am learning English.					
45	If I don't understand something in English, I ask the other person to slow down or say it again.					
46	I ask English speakers to correct me when I talk.					
47	I practice English with other students.					
48	I ask for help from English speakers.					
49	I ask questions in English.					
50	I try to learn about the culture of English speakers.					

Table 4. Preference of language learning strategies by their means

Rank	Strategy No.	Strategy Category	Strategy Statement	Mean
<i>High usage (M = 3.5 -5.0)</i>				
1	32	MET	I pay attention when someone is speaking English	4.44
2	38	MET	I think about my progress in learning English	4.20
3	33	MET	I try to find out how to be a better learner of English.	4.19
4	29	COM	If I can't think of an English word, I use a word or a phrase that means the same thing.	4.17
5	31	MET	I notice my English mistakes and use that information to help me do better.	4.08
6	15	COG	I watch English Language TV shows spoken in English or go to movies spoken in English.	4.05
7	30	MET	I try to find as many ways as I can to use my English.	3.79
8	40	AFF	I encourage myself to speak in English even when I am afraid of making a mistake.	3.77
9	39	AFF	I try to relax whenever I feel afraid of using English.	3.73
10	37	MET	I have clear goals for improving my English skills.	3.71
11	24	COM	To understand unfamiliar words, I make guesses.	3.64
12	1	MEM	I think of relationships between what I already know and new things I learn in English.	3.62
13	27	COM	I read English without looking up every new word.	3.61
14	16	COG	I read for pleasure in English.	3.59
15	36	MET	I look for opportunities to read as much as possible in English.	3.55

16	12	COG	I practice the sounds of English.	3.53
<i>Medium usage (M = 2.5-3.4)</i>				
17	17	COG	I write notes, messages, letters, or reports in English.	3.49
18	45	SOC	If I don't understand something in English, I ask the other person to slow down or say it again.	3.49
19	9	MEM	I remember new English words or phrases by remembering their location on the page, on the board, or on a street sign.	3.48
20	21	COG	I find the meaning of an English word by dividing it into parts that I understand.	3.46
21	49	SOC	I ask questions in English.	3.45
22	42	AFF	I notice if I am tense or nervous when I am studying or using English.	3.44
23	10	COG	I say or write new English words several times.	3.41
24	18	COG	I first skim an English passage (read over the passage quickly) then go back and read carefully.	3.40
25	13	COG	I use the English words I know in different ways.	3.37
26	35	MET	I look for people I can talk to in English.	3.37
27	19	COG	I look for words in my own language that are similar to new words in English.	3.35
28	34	MET	I plan my schedule so I will have enough time to study English.	3.34
29	47	SOC	I practice English with other students.	3.31
30	28	COM	I try to guess what the other person will say next in English.	3.30
31	50	SOC	I try to learn about the culture of English speakers.	3.29

32	3	MEM	I connect the sound of a new English word and an image or picture of the word to help me remember the word.	3.29
33	2	MEM	I use new English words in a sentence so I can remember them.	3.26
34	8	MEM	I review English lessons often.	3.19
35	4	MEM	I remember a new English word by making a mental picture of a situation in which the word might be used.	3.15
36	22	COG	I try not to translate word-for-word.	3.10
37	25	COM	When I can't think of a word during a conversation in English, I use gestures.	3.09
38	14	COG	I start conversations in English.	3.08
39	44	AFF	I talk to someone else about how I feel when I am learning English.	3.05
40	11	COG	I try to talk to native English speakers.	3.03
41	20	COG	I try to find patterns in English.	3.02
42	48	SOC	I ask for help from English speakers.	3.00
43	46	SOC	I ask English speakers to correct me when I talk.	2.88
44	23	COG	I make summaries of information that I hear or read in English.	2.87
45	41	AFF	I give myself a reward or treat when I do well in English.	2.73
<i>Low usage (M = 2.4 or below)</i>				
46	26	COM	I make up new words if I don't know the right ones in English.	2.43
47	5	MEM	I use rhymes to remember new English words.	2.39

48	6	MEM	I use flashcards to remember new English words.	2.35
49	7	MEM	I physically act out new English words.	2.20
50	43	AFF	I write down my feelings in a language learning diary.	2.16

MET (Metacognitive strategies), COG (Cognitive strategies), MEM (Memory strategies), COM (Compensatory strategies), SOC (Social Strategies), AFF (Affective Strategies).

Footnotes

¹ Many studies (see, e.g., Shmais, 2003) use GPA as a predictor of language proficiency. The assumption in such studies is that the higher the GPA, the more proficient the learner in the foreign language is.

² Htch and Lazaraton (1991: 310) maintain that when the cell sizes are unbalanced, “the assumption of equal variance may be violated.”

³ It should be noted that ANOVA gives exactly the same results as the *t*-test.

⁴ Despite the significant difference among the four groups ($F = 0.98, p = .042$), Schaffé post-hoc test showed tendency of difference, but it was not significant. As a result, LSD, which can locate differences when the mean differences are not big.

⁵ In step 1 of the regression analysis, the exponential B indicates that the odds of having a higher GPA increased by a factor of 2.626 with every point on a five-point Likert scale in the case of cognitive strategies. This ration jumped to 3.065 in the final step.

**Writing Performance Relative to Writing Apprehension,
Self-Efficacy in Writing, and Attitudes towards Writing:
A Correlational Study in Turkish Tertiary-Level EFL**

Dilek Yavuz Erkan

Çukurova University, School for Foreign Languages Adana/Turkey

dilek_yavuz@yahoo.com

Ayten İflazoğlu Saban

Çukurova University, Education Faculty, Adana/Turkey

iaayten@cu.edu.tr

Bio Data:

Dilek YAVUZ ERKAN holds a BA ELT from Cukurova University, an MATEFL from Bilkent University and a PhD from Cukurova University. She has taught graduate and undergraduate tertiary EFL courses at Center for Foreign Languages of Cukurova University since 1994. Her particular interests lie in the areas of the computer and the Internet use in EFL and affective dimensions of ELT. She is particularly interested in self-efficacy, multiple intelligence, emotional intelligence and NLP.

Ayten IFLAZOGLU SABAN holds an MA and PhD from Cukurova University, Education Faculty, Curriculum and Instruction Department. She has taught graduate and undergraduate Curriculum Development and Instruction courses at Cukurova University Education Faculty Elementary Education Department since 1998. Her particular interests lie in the areas of the teaching and learning strategies in teacher education and cooperative learning, multiple intelligence, elementary education, curriculum

development and classroom management.

Abstract

The purpose of this study is to identify whether writing performance in students of English as a foreign language (EFL) is related to writing apprehension, self-efficacy in writing, and/or attitudes towards writing. The subjects were tertiary-level EFL188 students at Çukurova University School for Foreign Languages (YADIM) in Turkey. Three instruments were used to collect data: a writing apprehension test (WAT), a self-efficacy in writing scale (SWS), and a questionnaire on attitudes towards writing (WAQ). The study was carried out in two phases. The first phase involved adapting the WAT and WAQ so these would be tailored to our study population. In the second phase, all subjects were tested with all three instruments in a 2-hour period on the same day. WAT, SWS and WAQ were administered in the first hour, and students were then given 45 minutes to write a composition on a given topic. The compositions were graded and these marks were taken to indicate students' overall writing performance points. The results of the study suggest that, in these tertiary-level EFL students, writing apprehension and writing performance are negatively correlated, writing apprehension and writing self-efficacy are negatively correlated, and writing apprehension and attitude towards writing are positively correlated. Writing apprehension and writing attitude are supposed to be negatively correlated but although WAQ was devised to test attitude, it has subscales that measure apprehension, too. That is why the result regarding the relationship between writing apprehension and attitude towards writing is surprising.

Keywords: Attitudes towards writing, EFL context, self-efficacy in writing, writing, writing apprehension, writing performance.

Introduction

Writing is a language skill that is essential to academic success. Since it is an active, productive skill, students learning to write in a foreign language face multiple challenges. For this group, writing requires thinking strategies that allow the individual to express

him or herself competently in the other language, and is a complex activity that requires a certain level of linguistics knowledge, writing conventions, vocabulary and grammar. As noted by Celce-Mercia (1991), expressing one's ideas in written form in a second or foreign language, and doing so with reasonable accuracy and coherence, is a major achievement. The complexity of writing as a task tends to heighten anxiety levels in students who are taking writing courses. This anxiety can often demotivate the student or lead to discouragement, and thus may result in negative attitudes towards writing (Gere, 1987; Sharples, 1993). Most students, low and high achievers alike, find writing difficult and view it as something they just have to persevere through in order to pass certain exams (Yavuz & Genç, 1998). This may relate to affective elements such as student attitudes, writing apprehension and self-efficacy in writing. As Pajares (2003) and Pajares, & Valiante (1996) suggest, our experience has shown that if a student is unwilling to express him or herself in writing, lacks confidence in his or her ability to write, or feels apprehensive about writing, then the student is unlikely to be proficient at writing composition.

Thus, over the past 30 years, an enormous amount of research has been done on the writing-composition processes that student writers undertake (Daly, 1978; Daly & Miller, 1975a, 1975b; Daly & Wilson, 1983; Faigley, Daly & Witte, 1981; McCarthy, Meier, & Rinderer, 1985; Onwuegbuzie, 1999; Pajares & Johnson, 1993; Phinney, 1991; Shell, Colvin, & Bruning, 1995; Shell, Murphy, & Bruning, 1989). Approaches to writing instruction have gone through numerous pendulum swings over the decades. Investigators have sought interesting and practical methods to enhance student writing, but an action research study carried out at Cukurova University (YADİM) (Yavuz & Genç, 1998) revealed that regardless of the methods used, students tend to exhibit negative attitudes

towards writing. Even students who are proficient at other language skills, namely, reading, speaking, listening and grammar, share the same problems with writing: They are afraid to make writing errors; they lack self-efficacy in writing. In short, they feel that actually expressing themselves in written English is beyond their command of the language.

Most language learners at all levels believe that writing is one of the most difficult language skills to master (Kurt & Atay, 2007; Latif, 2007; MacIntyre, & Gardner, 1989, 1991; Sağlam, 1993; Shrewsbury, 1995). Students learning English as a foreign language (EFL) at the Center for Foreign Languages at YADIM are no exception. The results of a questionnaire completed by 150 graduate and 353 undergraduate EFL students at YADIM at the end of the 2001-2002 academic year revealed that students believe writing and speaking are the most difficult skills to learn (Yavuz-Erkan, 2004). Yavuz and Genç (1998) also concluded that students view writing composition as something they “have to make it through” in order to pass the YADIM final exam. At YADIM, teachers consistently find that when students come to the writing sections of exams, many leave the classroom without trying to write even a few sentences. This may suggest that the students are extremely apprehensive and lack self-efficacy in writing.

It is reasonable to speculate that success with writing in a foreign language may be related to attitudes towards writing, apprehension about writing (hereafter referred to as “writing apprehension”), and self-efficacy in writing. The assumption that these elements do affect writing performance in this subject group was the impetus for this research. To investigate these possible relationships in the Turkish tertiary-level EFL context, the following research questions were formulated:

- Is there a significant difference between writing apprehension and self-efficacy in writing, attitudes towards writing and writing performance?
- Is there a significant relationship among writing apprehension, self-efficacy in writing, attitudes towards writing and writing performance?

The investigation was largely based on the following assumptions from earlier research:

- Writing apprehension is a strong predictor of academic performance in first language (Pajares & Valiante, 2001).
- Academic self-efficacy beliefs are strongly predictive of academic performance (Lent, Brown, & Larkin, 1984). Language learners with high self-efficacy who believe they can learn a language are more likely to learn a language than learners who believe they cannot learn a language (Templin, Shiroku and Taira, 1999). Students who lack confidence in skills they possess are less likely to engage in tasks in which those skills are required, and they will more quickly give up in the face of difficulty.
- Writing attitude influences writing achievement in first language (Graham, Berninger, & Fan, 2007).

English Language Instruction in Turkey

In Turkey, English language competency is the main determinant for having a well-paid, prestigious job. Turks have considered English the most popular and prestigious foreign language to learn since World War I (Acar, 2004; Ministry of Education). Today, English classes are an indispensable part of Turkish education system and more and more importance is being placed on English language education. Turkey has a dual education

system: private and state. Eight years of primary education is mandatory (between ages 6 and 14 years), and 4 years of state-paid optional secondary education is available for those 14 to 18 years of age. English is the primary foreign language and is taught in private schools starting at kindergarten. In state schools, however, English classes start in fourth grade and continue throughout secondary school. Students are exposed to 2 hours of basic English instruction per week in state schools (mainly following a grammar-based curriculum), whereas private schools offer the same basic English-language curriculum but devote more time for classes. Depending on the school, this can be 8-10 hours of instruction per week based on a task-based and integrated-skills curriculum that encourages use of language. Regarding university-level English education, at so-called “English-medium” universities, first-year students receive intensive English-language teaching, usually 24 to 30 hours of instruction per week. In such programs, students are taught using task-based or integrated-skills syllabi or a combination of both. In the research context, a combination of both is used. At the universities that are not English-medium, students have two hours of English classes a week throughout their education.

Overview of Writing Issues in the Research Context

Writing is one of the most important skills for learners at English-medium universities like Çukurova University to master. In the research context, at Çukurova University Center for Foreign Languages (YADIM), acquiring good writing skills is considered to be of utmost importance. YADIM students are tertiary-level and are learning English for academic purposes. At their departments they are required to write projects and to sit written exams in order to complete their education. Reading and writing are the major language skills students must master to fulfill school requirements at their departments.

Reading of English is a receptive skill and therefore less problematic for a university student to learn. In contrast, writing is a productive skill and considered one of the most difficult to learn.

The process-oriented approach to writing has been adopted in the research context, YADIM. One of the activities that has emerged and gained popularity with the introduction of this approach is the “Writing Portfolio.” Paulson, Paulson, & Meyer (1991, p. 60) define writing portfolio as “a purposeful collection of student work that exhibits the student’s efforts, progress, and achievements in one or more areas.” Students of tertiary level at most universities in Turkey are required to keep a writing portfolio just like students studying at Çukurova University, YADIM. In a typical writing class, YADIM students write a rough draft and the work is then edited by a peer who comments on the content, provides feedback, checks the work for mechanical mistakes such as grammar, punctuation and capitalization, and sometimes makes corrections. Then a conference is held with the teacher, who also reads the work. Required changes are made in light of the peer’s and teacher’s suggestions and, finally, a finished copy is produced and placed in the student’s portfolio folder. What makes process writing valuable is that the same piece of writing is worked on until a satisfactory and effective final draft is achieved through peer and teacher editing. This stepwise method helps forge lasting improvements in student writing. In addition, editing peers’ work teaches learners how to critique, which may lead them to look at their own work critically and consequently improve their own writing. In some foreign language centers of Turkish universities, including YADIM, word processing is also used and makes process writing easier because of the ease with which drafts can be created.

Research Context

This study was conducted at the Center for Foreign Languages at Çukurova University (YADIM). This center offers an integrated-skills program to graduate and undergraduate students. The program consists of four levels (elementary, pre-intermediate, intermediate, upper-intermediate), each taught in an 8-week period. At the end of the academic year, each YADIM student takes a standard proficiency test before continuing on in their main area of study. This proficiency test is designed to measure the student's ability in English regardless of the language training they have had throughout the year. In addition, each student takes an achievement test every 2 months. Thus, they complete four achievement tests and one proficiency test before finishing the program. Each of these tests includes one section solely devoted to writing. At YADIM, students' writing skills are assessed based on a portfolio of revised essays the student produces over the course of the semester, and based on the timed-essay section of the achievement and proficiency tests.

Method

Subjects

Twelve randomly selected tertiary-level classes of 24 students (188 students in total) took part in this study. The subjects were all Turkish students learning English as a foreign language at Çukurova University, an "English-medium" university as mentioned above. At the time of the study, the students had been attending YADIM for 8 months and ranged from 18 to 22 years of age. All subjects had been enrolled in beginner English classes at the start of the academic year and were intended to have achieved intermediate-level academic proficiency in English after 8 months of the intensive EFL program.

Measurements

Three instruments were used for this research were: A writing apprehension test (WAT), a self-efficacy in writing scale (SWS), and an attitude-towards-writing questionnaire (WAQ).

Writing Apprehension Test (WAT)

For this study, we used Daly and Miller's (1975a) WAT to assess students' apprehension about writing. This is a 26-item questionnaire that features 13 items with positive polarity and 13 with negative polarity. Scoring is done on a 5-point Likert scale that asks the subject to state whether they agree or disagree with statements about writing. This inventory was specifically designed to measure self-reported writing apprehension.

To date, the WAT is the instrument most widely used to measure anxiety related to second-language writing. This test was originally developed with reference to first-language learners, particularly English native speakers, and therefore may not tap the essential aspects of foreign-language writing anxiety. Cheng (2004) stated that, although on the whole the Daly-Miller WAT has been shown to be an instrument of satisfactory internal consistency reliability with concurrent and predictive validity, it needs to be adapted in order to be used to study second-language writing. We adapted the WAT for use with our Turkish EFL students as follows:

Phase 1. First, the WAT was translated into Turkish by three EFL instructors and the two researchers. Second, translations were compared and necessary changes were made so that all items in the scale were clear to Turkish EFL students. Then the scale was read by three Turkish lecturers. It was revised based on their comments and then administered to 15 tertiary-level EFL students. After this, the final version of the scale was established

with feedback from the 15 students. Once in its final Turkish form, the scale was back-translated. The team then discussed any discrepancies between the two versions until all members reached a consensus. The final version turned out to be very similar to the original one. (We used the Turkish version of WAT, and when the Turkish version was translated into English to ensure construct validity, we saw that it was almost the same as the original version so since WAT appears on the Internet, we think there is no point in adding our translation (from Turkish to English) to the Appendix. We did not use it anyway, we used the Turkish version. Do you think we should have added the Turkish version?)

Phase 2. Next, the modified 26-item WAT was administered to 263 tertiary-level EFL students to assess its validity and reliability. Exploratory factor analysis with principal components estimation and varimax rotation was applied. Results indicated that five factors with Eigen values above 1.00 accounted for 62.13% of the variance. However, two factors had Eigen values above 2.00 (9.72 and 2.84). A scree plot of the factors showed that a solution with two factors was possible.

Phase 3. Factor loads were studied in this two-factor solution found through varimax rotation and items with factor loads of .30 or lower were dropped from the scale. Once this was done, the first subscale consisted of 12 items, the second subscale consisted of 11 items, and three items were omitted. The two factors accounted for 49.59% of the variance.

Phase 4. The results and the items loaded on two factors ended up to be very similar to the findings for the validity-reliability analysis done by Daly and Miller (1975). As those authors concluded, the two factors represented a separation of negatively and positively worded questions rather than two interpretable components of the overall measure.

Consequently, one factor was generated. For this solution, the Kaiser-Meyer-Olkin measure of sample adequacy was .92. Item analyses for both scales revealed that total items' correlations ranged from .36 to .77, the item mean values ranged from 2.61 to 3.84, and the standard deviations ranged from .93 to 1.19. The Cronbach alpha reliability coefficient for the scale was .92. A split-half technique indicated that the reliability of the instrument was .86.

Self-Efficacy in Writing Scale (SWS)

The self-efficacy scale developed by Yavuz-Erkan (2004) was used to assess the students' self-efficacy in writing. Based on the self-efficacy construct proposed by Bandura (1977), Yavuz-Erkan developed a 21-item writing self-efficacy scale to grade the strength of subjects' belief in their writing ability. The items of the scale were graded with the four-tier system Likert scale: Strongly Disagree, Disagree, Agree, or Strongly Agree. Each statement on the scale was preceded by the phrase "I can ..." (Appendix 1).

The reliability and validity were made by Yavuz-Erkan (2004). The scale was found to be cronbach alfa coefficient were .88 for the first factor (*Content*), .80 for the second factor (*Design*), .77 for the third factor (*Unity*), .74 for the fourth factor (*Accuracy*), and .50 for the fifth factor (*Punctuation*). According to the factor analysis results, the variance explained with five factors was found to be 66.16

In the current study, in order to make the Cronbach alpha reliability coefficient of punctuation subscale closer to the reliability point .70, the above-mentioned scale was administered to 189 YADİM students. Analyses yielded a solution with four factors. Four factors accounted for 66.16% of the variance. The Cronbach alpha reliability coefficients were .92 for the first factor (*Content*), .94 for the second factor (*Design-Unity*), .74 for

the third factor (*Accuracy*), and .72 for the fourth factor (*Punctuation*). In this application, the *Design* and *Unity* subscales blended to form a single subscale even though they appeared to be different subscales in Yavuz-Erkan's scale.

Based on its robust psychometric properties, the researchers deemed this new writing self-efficacy scale a reliable and valid tool for assessing self-efficacy in writing in foreign language. The scale was administered to the subjects in order to determine their self-beliefs linked to writing in English.

Attitude-Towards-Writing Questionnaire (WAQ)

The third measure used in this study was Rose's (1984) attitude-towards-writing questionnaire (WAQ), a tool for examining the link between attitudes about writing and student writers' actual writing performances. Attitude refers to the respondent's feelings about his or her writing ("I think my writing is good") and evaluation of that writing ("I think of my instructors reacting positively to my writing"). Rose's questionnaire has been shown to be valid for English and Spanish bilingual student populations (Betancourt & Phinney, 1988).

This questionnaire was adapted to Turkish in the following steps:

Phase 1. First, the WAQ was translated into Turkish by the researchers and three English instructors, the translation was compared and the necessary changes were made. Then the translation was edited by three Turkish teachers and necessary changes were made. Then the questionnaire was given to 15 students. More modifications were made and then the questionnaire was translated into English and compared with the original version. The two versions of the questionnaire were similar.

Phase 2. After the above-mentioned language validation study was completed, the 24-

item questionnaire (The explanation we made above for WAT applies to your suggestion here, too) was administered to 190 YADIM students. Principal components analysis was done to determine the factor structure of the scale. The results revealed that six factors with Eigen values above 1.00 accounted for 62.81% of the variance. There were also two factors (7.38 and 2.84) with Eigen values above 2.00. A scree plot of these factors showed that it is possible to reach a solution with two factors. However, no limitation was performed in order to reach a solution suitable for the structure of the original questionnaire.

Phase 3. Only the loads of the factors obtained through the analysis were examined. Items with factor loads of .30 or lower and items with the difference between factor loads that were attached to more than one item below .20 were dropped from the scale. In this case, the first subscale contained five items, the second subscale had four items, the third scale had three items and the fourth subscale had two items. Eight items were omitted from the scale. Results indicated that five factors accounted for 67.89% of the variance.

Phase 4. The results and the items loaded to factors were very similar to the validity-reliability study done by Rose (1984). Specifically, the five subscales found in the present study overlap those that Rose identified. For the present study, the first subscale found was named “Blocking”, the second “Complexity”, the third “Editing”, the fourth “Attitude”, and the fifth “Lateness.”

Phase 5. The adequacy of the sample was tested with the Kaiser-Meyer-Olkin test. For this solution, the Kaiser-Meyer-Olkin measure of sample adequacy was .82.

The items’ common variance values ranged from .65 to .91, and the total items’ subscale correlations ranged from .65 to .88. The item mean values ranged from 3.17 to 3.85, and the standard deviations ranged from .91 to 1.28. The Cronbach alpha reliability

coefficients were .83 for the first factor (*Blocking*), .75 for the second factor (*Complexity*), .67 for the third factor (*Editing*), .79 for the fourth factor (*Attitude*), and .76 for the fifth factor (*Lateness*).

Statistical Analysis

The data from the study were all analyzed using SPSS 11.5 program. To analyze the data Pearson correlation coefficient, independent groups t-test and one-way variance analysis (ANOVA) were used. The analyses were carried out at a significance level of $p = .05$.

Respondents were classified as high, moderate, or low in writing apprehension on the basis of their responses to WAT. Responses were summed for each person so that a high score always indicated high apprehension. Individuals scoring one standard deviation above or below the group apprehension score ($M = 66.68$, $SD = 6.66$) were operationally defined as high and low, respectively, in apprehension. Respondents whose scores fell within one standard deviation of the mean were classified as moderates. A one-way analysis of variance with three levels of apprehension was computed on the apprehension scores to insure that the group differences expected were present. A significant overall effect was observed, $F(2, 185) = 270.134$, $p < .0001$. Follow-up analyses using the conservative Scheffé procedure for multiple comparisons indicated that each group mean was significantly different ($p < .01$) from every other one. For the primary test of the hypotheses all groups were used.

Findings

One-way ANOVA was done for the three levels of writing apprehension (high, moderate, low) using the overall score on the achievement test as a dependent measure. The results

shows that, the higher the achievement test score the better the individual's performance, the group with high-level writing apprehension ($X = 6.07$, $SD = 4.01$) performed significantly more poorly [$F(1,185) = 6.020$, $p < .01$] on the achievement test than the groups with low and moderate writing apprehension ($X = 10.48$, $SD = 5.72$ and $X = 8.93$, $SD = 6.30$, respectively) One-way ANOVA was also done to assess for significant differences among the writing-apprehension groups' results on the SWS and WAQ. These results are shown in Table 1.

Table 1. Mean, standard deviations and one-way ANOVA results for writing apprehension levels and SWS and WAQ points

		Writing Apprehension Level								
		Low (N=88)		Moderate (N=73)		High (N=27)				
		Mean	SD	Mean	SD	Mean	SD	F	Sig.	Scheffé
	Writing Achievement	10.48	5.72	8.93	6.30	6.07	4.01	6.020	.003	H ≈ L H ≈ M
	Writing Self-Efficacy	78.23	13.47	70.57	15.36	62.24	11.70	15.191	.000	L ≈ M L ≈ H M ≈ H
	Design-Unity	38.52	7.24	33.87	8.00	29.02	6.64	19.053	.000	L ≈ M L ≈ H M ≈ H
Writing Self-Efficacy Sub-Scale	Content	17.54	4.20	15.83	4.47	14.07	3.60	7.901	.001	L ≈ M L ≈ H
	Accuracy	9.05	2.28	8.46	2.47	7.37	2.01	5.490	.005	L ≈ H
	Punctuation	7.82	1.78	7.25	1.97	6.89	2.14	3.265	.040	Not significant among groups
Writing Attitude Sub-Scale	Blocking	16.25	4.08	17.98	3.89	20.97	3.01	15.873	.000	H ≈ L H ≈ M
	Complex	12.76	3.34	13.78	3.19	16.16	2.21	12.144	.000	H ≈ L H ≈ M
	Editing	9.08	2.82	10.08	2.57	10.64	2.95	4.525	.012	H ≈ L
	Attitude	6.62	1.20	6.95	2.15	5.74	1.81	3.461	.033	M ≈ H
	Lateness	5.15	2.05	6.11	2.17	8.07	2.04	20.538	.000	H ≈ L H ≈ M M ≈ L

H= High, M= Medium, L = Low

The results for writing apprehension levels, total SWS scores, and scores for each of the

subscales studied reveal that the students with high self-efficacy in writing had low-level writing apprehension. One-way ANOVA revealed a significant difference among the means for SWS total score and subscale scores for the groups of students with moderate and high writing apprehension. [SWS total $F(2,185) = 15.191$, $p < .01$, Design-Unity $F(2,185) = 19.053$, $p < .01$, Content $F(2,185) = 7.901$, $p < .01$ Accuracy $F(2,185) = 5.490$, $p < .01$ Punctuation $F(2, 185) = 3.265$, $p < .05$]. The Scheffé-F test was applied to determine which pairs of groups had significantly different results. Scheffé's F-testing revealed that students with low-level writing apprehension had significantly better scores than students with high-level writing apprehension for the above parameters. "Punctuation" was the only subscale in which Scheffé's test revealed no significant difference between these two groups.

When writing apprehension levels and points of attitude towards writing were studied, it was seen that means of 'Blocking, Complex, Editing, Lateness' subscales increased when writing apprehension level increased, it was also observed that means of 'attitude' subscale differed among levels and students with high apprehension level had lower means than the other levels (moderate, low).

To find out if the difference between mean scores was significant one-way variance analysis was performed. The results showed that there was a significant difference between low, moderate and high apprehension levels in terms of all subscales of WAT [Blocking $F(2, 185) = 15.873$, $p < .01$, Complex $F(2, 185) = 12.144$, $p < .05$, Editing $F(2, 185) = 4.524$, $p < .05$, Lateness $F(2,185) = 20.538$, $p < .01$, Attitude $F(2, 185) = 3.461$, $p < .05$]. Follow up analyses using Scheffé's multiple comparison procedure revealed that the differences among the three means were statistical significant. However, the magnitude of the means differences varied. Table 1 summarizes the important statistical information.

To assess whether students' scores for the WAT, SWS and WAQ were related to actual writing achievement scores, Pearson correlation coefficients were calculated. Table 2 shows these results.

Table 2. Pearson correlation analysis for writing achievement scores relative to WAQ, SWS, and WAT .

		Writing Achievement	Writing Apprehension	Writing Self-Efficacy	Writing Attitude
	Writing Apprehension	-.23(**)	1	-.38(**)	.47(**)
	Writing Self-Efficacy	.27(**)	-.38(**)	1	-.46(**)
Writing Self-Efficacy Sub-Scale	Design-Unity	.30(**)	-.41(**)	.95(**)	-.43(**)
	Content	.14	-.34(**)	.81(**)	-.46(**)
	Accuracy	.26(**)	-.24(**)	.81(**)	-.36(**)
	Punctuation	.08	-.12	.50(**)	-.18(*)
Writing Attitude Sub-Scale	Blocking	-.20(**)	.38(**)	-.50(**)	.85(**)
	Complex	-.06	.32(**)	-.43(**)	.71(**)
	Editing	-.02	.21(**)	-.08	.61(**)
	Attitude	.17(*)	-.02	.42(**)	-.19(*)
	Lateness	-.22(**)	.43(**)	-.53(**)	.71(**)

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

The Pearson correlation coefficient for writing achievement and WAS total score was -.23, indicating a negative relationship between writing apprehension level and writing achievement. The correlation coefficient for WAS and SWS was -0.38 whereas total points between WAS and WAQ is 0.47. As the table indicates, both these relationships were statistically significant.

The Pearson correlation coefficient for writing achievement and SWS total score was 0.27, indicating a positive relationship between self-efficacy in writing and writing performance. When the values were calculated for the different SWS subscales relative to writing achievement, the coefficients indicated positive statistical relationships for

Design-Unity (0.30) and *Accuracy* (0.26). The other subscales were not statistically correlated with writing achievement.

When the correlation coefficients for WAQ subscales and writing achievement were calculated, the results showed significant negative relationships for *Blocking* and *Lateness* (*Blocking* -0.20, *Lateness* -0.22) and a significant *positive relationship* between *Attitude* and writing *performance* (*Attitude* .17).

Discussion and Conclusion

The term “writing apprehension” was coined by Daly and Miller (1975, p. 244), to describe “the tendency of a person to avoid the process of writing - particularly when it is to be evaluated in some way”. Learners experiencing this feeling often find it difficult to express themselves, avoid writing as much as possible and find such classes unfavorable. Theoretically, apprehension often impairs and hinders performance so those with low writing apprehension are supposed to perform better on tests of writing skills than apprehensive writers (Krause, 1994). In the present study, we tested the hypothesis that individuals with low-level writing apprehension would perform significantly better on a test of writing skills than those with high-level writing apprehension. The results confirmed that this theory holds true.

Previous research has demonstrated a clear, consistent and significant correlation between a student’s level of writing apprehension and writing performance in both first and second languages. Indeed, our results were in line with those of many other researchers (Daly, 1978; Daly & Miller, 1975a, 1975b; Daly & Wilson, 1983; Faigley, Daly & Witte, 1981; McCarthy, Meier, & Rinderer, 1985; Onwuegbuzie, 1999; Pajares & Johnson, 1993; Phinney, 1991; Shell, Colvin, & Bruning, 1995; Shell, Murphy, &

Bruning, 1989). Daly (1978) found that students with high levels of writing apprehension consider writing to be unrewarding, and that these individuals will avoid classes that involve writing assignments, if possible. His 1978 study of 3,602 undergraduate writing-composition students examined the differences in writing skills between groups with low- and high-level writing apprehension. Daly concluded that students with low writing apprehension tend to have stronger writing skills than those with high apprehension. This result is supported by Lee and Krashen (1997), who reported that native speakers of Chinese in Taiwan who had higher writing apprehension tended to receive lower evaluations on the composition section of the Senior High School Examination in Taiwan ($r = -0.21$). Powell's study (1984) also had a similar outcome, with a modest relationship ($\tau = -0.27$) between writing apprehension and writing performance of students in a variety of university-level writing courses.

Apart from Gungle and Taylor's (1989 cited in Phinney, 1991) adaptation of the Daly-Miller WAT and a study of bilingual students by Betancourt and Phinney (1988), very little research has been done on writing apprehension in second-language writers. Our study expands this effort and looks specifically at relationships between writing apprehension and writing performance in the Turkish tertiary-level EFL context (see page 5 above). If, as our research shows, writing apprehension is a predictor of student competence in writing, then teachers should do their best to minimize writing apprehension.

Bandura (1997) stated that academic self-beliefs are strongly predictive of academic performance. Beach (1989), Meier, McCarthy, & Schmeck, (1984), Shell, Murphy & Bruning (1989), McCarthy, Meier & Rinderer (1985), Zimmerman and Bandura (1994), and McCarthy, Meier, & Rinderer, (1985) have all investigated the predictive value of

self efficacy relative to writing performance. These researchers all concluded that the two constructs are related; that is, self-efficacy is a predictor of actual writing performance. If self-efficacy is a predictor of student competence, then teachers should pay attention to students' perceptions of their personal competence as well as their actual competence.

Pajares (2003) asserts that psychological states such as anxiety and stress provide information about efficacy beliefs. Our study also showed that students with high self-efficacy beliefs regarding writing in a foreign language had low-level writing apprehension. This finding supports the conclusions of Daly and Wilson (1983), Onwuegbuzie (1999) and Pajares & Johnson (1993) that writing apprehension is negatively correlated with self-efficacy. Similarly, it supports Pajares and Valiante's (1996) findings that elementary students' self-efficacy beliefs about writing capability predicted their writing performance, and that these beliefs also directly influenced the students' writing apprehension in their first language. In our study, we found no statistically significant difference among the three different writing-apprehension groups (low, moderate, high) with respect to *Punctuation* subscale scores in the SWS. This suggests that students at all apprehension levels have similar beliefs about their ability to use punctuation. In their 1997 study of self-efficacy in writing, Pajares & Valiante concluded that students with high self-efficacy scores had low-level writing apprehension. McCarthy, Meier and Rinderer (1985) also found that self-efficacy and writing anxiety were correlated with essay scores of undergraduate college students. Shell *et al.* (1989) investigated the writing self-efficacy of undergraduates and reported a significant correlation between students' confidence in their writing skills and their holistic score on a 20-minute essay (0.32).

In our study, we also hypothesized that writing performance is related to writing

apprehension, self-efficacy in writing, and attitudes towards writing. The results indicate negative correlations between writing apprehension and writing performance ($r = -0.23$) and between writing apprehension and writing self-efficacy ($r = -0.38$), and a positive correlation between writing apprehension and writing attitude ($r = 0.47$) (Table 2). This last result was surprising as one would expect that when one's attitude towards writing is good, one's apprehension about writing would be lower. As Phinney (1991) explains, while testing attitude, WAQ has such subscales that measure apprehension too, although it was devised to test attitude.

Implications and Suggestions for Further Study

Writing is not only a cognitive but also an emotional activity; thus, the affective components of writing strongly influence all phases of the writing process (McLeod, 1987). The results of our study underscore the need for nurturing the affective dimensions of EFL learning. The implication is that EFL practitioners should investigate students' beliefs about their writing capabilities, their attitudes towards writing, and their level of apprehension about writing. Our findings point to all these as important predictors of academic writing performance in tertiary-level EFL students. As Bandura (1986) suggests, self-efficacy beliefs are developed primarily through enactive attainment; that is, people's confidence grows as they attempt and complete tasks and the feeling of success increases confidence whereas failure decreases it. Bandura also acknowledges that verbal persuasion, that is, messages individuals receive from authority figures, can reinforce and increase self-efficacy.

As well as helping students with writing apprehension, self-efficacy and attitude, teachers should also make efforts to help students understand how their affective

processes can influence their EFL writing performance. In short, teachers should make every effort to help their students increase competence through confidence.

As noted, our investigation provides evidence for the roles and importance of writing apprehension, attitudes towards writing, and writing self-efficacy in relation to actual writing performance. This research could be considered a preliminary investigation on which follow-up work could be based. In a comparative future study, it would be interesting to assess whether student performance improves after training or coaching on self-efficacy, apprehension, and attitude related to writing.

References

- Acar, K. (2004). Globalization and language: English in Turkey. *Sosyal Bilimler* 2 (1), 1-10.
- Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review*. 84, 191-215.
- Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. Englewood Cliffs, NJ: Prentice Hall.
- Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York: W. H. Freeman and Company.
- Beach, R. (1989). Showing students how to assess: Demonstrating techniques for response in the writing conference. In C. M. Anson (Ed.), *Writing and Response*, 127-148. Urbana, IL: NCTE.
- Betancourt, F. & Phinney, M. (1988). Sources of writing block in bilingual writers. *Written Communication*, 5(4), 461-478.
- Celce-Murcia, M. (1991). *Language teaching methodology*. London: Prentice Hall.

- Cheng, Y (2004) A measure of second language writing anxiety: Scale development and preliminary validation. *Journal of Second Language Writing*. 13(4), 313-335.
- Cheng, Y.S., Horwitz, E.K., & Schallert, D. (1999). Language anxiety: Differentiating writing and speaking components, *Language Learning* 49, 417–446.
- Daly, J. A. & Wilson, D. A. (1983). Writing apprehension, self-esteem and personality. *Research in the Teaching of English*. 17, 327-341.
- Daly, J. A., & Miller, M. D. (1975). The empirical development of an instrument to measure writing apprehension. *Research in the Teaching of English*, 9, 242-249.
- Daly, J. and Miller, M. (1975a). The empirical development of an instrument to measure writing apprehension. *Research in the Teaching of English*, 9, 242-249.
- Daly, J. and Miller, M. (1975b). Further studies in writing apprehension: SAT scores, success expectations, willingness to take advanced courses, and sex differences. *Research in the Teaching of English*, 9. 250-256.
- Daly, J.A. (1978). Writing apprehension and writing intensity in business and industry. *Journal of Educational Research*, 72, 10-14.
- Faigley, L., Daly, J. A., & Witte, S.P. (1981). The role of writing apprehension in writing performance and competence. *Journal of Educational Research*, 75, 16-21.
- Gere, A. R. (1987). *Writing groups: History, theory and implications*. Carbondale: Southern Illinois University Press.
- Graham, S., Berninger, V. & Fan, W. (2007). The structural relationship between writing attitude and writing achievement in first and third grade students. *Contemporary Educational Psychology*, 32(3), 516-536
- Hadaway, N. L. (1987). *Writing apprehension among second language learners*. Unpublished Doctoral Dissertation, Texas A & M University, College Station.

- Krause, K.L.(1994). *The effect of goal-setting and planning on the writing competence of secondary school students*. Paper presented at the Annual Conference, Australian Association for Research in Education, Newcastle, NSW.
- Kurt, G. & Atay, D. (2007). The effects of peer feedback on the writing anxiety of prospective Turkish teachers of EFL. *Journal of Theory and Practice in Education* 3 (1), 12-23
- Latif, M. A. (2007). The factors accounting for the Egyptian EFL university students' negative writing affect. *Essex Graduate Student Papers in Language & Linguistics* 9, 57-82
- Lee, S.Y. and Krashen, S. (1997). Writing apprehension in Chinese as a first language. *ITL: Review of Applied Linguistics*, 115-116, 27-37.
- Lee, S.Y., (2001). The relationship of writing apprehension to the revision process and topic preference: A student perspective In: P.-H. Chen and Y.-N. Leung, Editors, *Selected papers from the tenth international symposium on English teaching*, Crane, Taipei, Taiwan, pp. 504–516.
- Lent, R. W., Brown, S. D. & Larkin, K.C. (1984). Relation of self-efficacy expectations to academic achievement and persistence. *Journal of Counseling Psychology*, 31, 356-362.
- MacIntyre, P. D. & Gardner, R. C. (1989). Anxiety and second language learning: Towards a theoretical clarification. *Language Learning*, 39, 251-275.
- MacIntyre, P. D. & Gardner, R. C. (1991). Language anxiety: Its relation to other anxieties and top-processing in native and second Languages. *Language Learning*, 41, 513-534.
- Masny, D., & Foxall, J. (1992). *Writing apprehension in L2*. (ERIC Document

Reproduction Service No. ED 020 882).

McCarthy, P., Meier, S., & Rinderer, R. (1985). Self-efficacy and writing. *College Composition and Communication*, 36, 465-471.

McLeod, S. (1987). Some thoughts about feelings: The affective domain and the writing process. *College Composition and Communication*, 38, 4, 426-435.

Meier, S., McCarthy, P. R., & Schmeck, R. R. (1984). Validity of self-efficacy as a predictor of writing performance. *Cognitive Therapy and Research*, 8, 107-120.

Ministry of Education -Turkey: <http://www.meb.gov.tr>

Onwuegbuzie, A. J. (1999). Writing apprehension among graduate students: Its relationship to self-perceptions. *Psychological Reports*, 84, 1034-1039.

Pajares, F & Valiante, G. (1997). Influence of self-efficacy on elementary students' writing. *Journal of Educational Research*, 90(6), 353-360.

Pajares, F. & Johnson, M. J. (1993). Confidence and competence in writing: The role of self-efficacy outcome expectancy, and apprehension. *Research in the Teaching of English*, 28, 313-31.

Pajares, F. & Valiante, G. (1996). *Predictive utility and causal influence of the writing self-efficacy beliefs of elementary students*. Paper presented at the Annual Meeting of the American Educational Association, New York, NY.

Pajares, F. (2003). Self-efficacy beliefs, motivation, and achievement in writing: a review of the literature. *Reading & Writing Quarterly*, 19: 139-158.

Pajares, F., & Johnson, M. J. (1994). Confidence and competence in writing: The role of writing self-efficacy, outcome expectancy, and apprehension. *Research in the Teaching of English*, 28, 313-331.

Pajares, F., & Valiante, G. (2001). Gender differences in writing motivation and

- achievement of middle school students: A function of gender orientation? *Contemporary Educational Psychology*, 24, 366-381.
- Paulson, F. L., Paulson, P. R., & Meyer, C. A. (1991). What makes a portfolio a portfolio: Eight thoughtful guidelines to help educators encourage self-directed learning. *Educational Leadership*, 48(5), 60-63.
- Phinney, M. (1991). Word processing and writing apprehension in first and second language. *Computer and Composition*, 11(1), 65-82.
- Powell, B. J. (1984). A comparison of students' attitudes and success in writing. *Journal of Negro Education*, 53, 114-123.
- Rose, M. (1984). *Writer's block: The cognitive dimension*. Carbondale, IL: Southern Illinois University Press.
- Sağlam, G. (1993). The attitudes of some Turkish students toward responses to their EFL writing. *Unpublished Masters Thesis*. The Institute of Economic and Social Science of Bilkent University. Ankara/Turkey.
- Sharples, M. (ed.) (1993). *Computer supported collaborative writing*. Springer-Verlag.
- Shell, D. F., Colvin, C., & Bruning, R. H. (1995). Self-efficacy, attributions and outcome expectancy mechanisms in reading and writing achievement: Grade-level and achievement level differences. *Journal of Educational Psychology*, 87, 386-398.
- Shell, D. F., Murphy, C. C. & Bruning, R. H. (1989). Self-efficacy and outcome expectancy mechanisms in reading and writing achievement. *Journal of Educational Psychology*, 81, 91-100.
- Shrewsbury, M. B. (1995). *The effects of collaborative learning on writing quality, writing apprehension, and writing attitude of collage students in a developmental English program*. Unpublished Dissertation Abstract. The College of Human

- Resources and Education of West Virginia University, Morgantown/West Virginia
- Song, M. (2007). A promising practice for improving writing quality and decreasing writing apprehension of Korean students. *Modern English Education*, 8(2), 113-135.
- Templin, S.A, Shiroku, M & Taira, K. (1999). Self efficacy syllabus. Retrieved from <http://www.jalt-publications.org/tlt/articles/1999/04/templin> [21.02.2009].
- Wu, Y. (1992). *First and second language writing relationship: Chinese and English*. Unpublished doctoral dissertation, Texas A & M University, College Station.
- Yavuz, D. & Genc, A.B. (1998). Flexibility of setting up a writing unit at YADIM, *Unpublished Action Research Study*. Çukurova University, Adana, Turkey.
- Yavuz-Erkan, D. (2004). *Efficacy of cross-cultural e-mail exchange for enhancing EFL writing: A perspective for tertiary-level Turkish EFL learners*. Unpublished Dissertation Abstract. Çukurova University, The Institute of Social Sciences English Language Teaching. Adana/Turkey
- Zimmerman, B. J. & Bandura, A. (1994). Impact of self-regulatory influences on writing course attainment. *American Educational Research Journal*, 31, 845-862.

Appendix 1

Writing Efficacy Scale

Read each statement below and then use the following scale to indicate various degrees of effectiveness. Of course, there are no right or wrong answers to such questions, so do not spend too much time on any one statement, but select the answer that best applies to you. Thank you for your cooperation.

1= I do it very well 2= I do it well 3= I do not do it well 4= I do not do it well at all

		1	2	3	4
1	I can write interesting and appropriate response to a given topic				
2	I can easily cover all the information that should be dealt within a given topic.				
3	I can use appropriate style to the task.				
4	I can easily match style with topic				
5	I can generate ideas to write about easily.				
6	I can think of ideas rapidly when given a topic to write about.				
7	I can write on an assigned topic without difficulty.				
8	I can easily find examples to support my ideas.				
9	I can justify my ideas in my compositions.				
10	I can write grammatically correct sentences in my compositions.				
11	I can use complex language in writing without difficulty.				
12	I can produce error free structures.				
13	I can spell very well.				
14	I can use the punctuation correctly.				
15	I can edit my compositions for mistakes such as punctuation, capitalization, paragraphing.				
16	I can easily use structures I have learned in my class accurately.				
17	I can link ideas together easily.				
18	I can use transition words correctly to make my composition a better one.				
19	I can use connectors correctly to make my composition a better one.				
20	I can use a wide range of vocabulary in my compositions.				
21	I can use synonyms in a composition rather than repeating the same words over and over again.				
22	I can write a brief and informative overview of a given topic.				
23	I can manage my time efficiently to meet a deadline on a piece of writing.				
24	I can rewrite my wordy or confusing sentences to make them				

	clearer.				
25	I can extend the topic to fit in a given word limit.				
26	I can choose and defend a point of view.				
27	I can make long and complex sentences.				
28	I can fulfill a writing task without difficulty within a given time limit.				

The Role of Portfolio Assessment and Reflection on Process Writing

Ahmad Sharifi

University of Guilan

Jaleh Hassaskhah

University of Guilan

Bio Data:

Ahmad Sharifi received his M.A. in TEFL from Guilan University. His main area of interest is issues in assessment and testing. He is currently teaching undergraduate courses in Shahid Sattari Air University, Iran.

Dr. Jaleh Hassaskhah is a lecturer at the university of Guilan. She has presented many papers in different conferences. She also teaches undergraduate and postgraduate courses in the university of Guilan, Iran.

Abstract

Language teaching and testing have always been highly interrelated in the sense that it's been impossible to work in either without taking the other into account. By the movement of language teaching in the direction of learner-centered approach, testing and assessment have begun to apply the same approach. However, it seems that applying a single test at the end of the course is still popular. Since any single measure seems incapable of estimating the diversity of skills, knowledge, processes, and strategies that combine to determine student progress, scholars have begun to incorporate alternative assessment techniques to yield more useful information about students' achievement and classroom instruction. One of these alternative assessment techniques is portfolio assessment. The

present study has devoted itself to investigate the effect of using portfolio assessment technique and reflection activities on students' writings and process writing. This study, which followed the quasi-experimental design, was conducted in a class of 20 students at Shahid Sattari Air University of Iran. During one semester, they took ten tests: five pretests, and five posttests. The portfolio-based teaching in the second half of the semester was introduced as the treatment. Also, a validated questionnaire was given to students to express their attitude to portfolio-based learning. A set of paired-sample *t* tests was run to compare the students' tests. Each pair of tests was compared to see how much progress they made over time. The level of significance in this study was .001 as a result of using repeated measurement and Bonferroni test. Based on the findings achieved in this study, the effectiveness of the treatment was confirmed. Further, the students' responses to questionnaire indicate that their attitude was positive to portfolio-based learning.

Keywords: Assessment, Portfolio, Portfolio Assessment, Writing Assessment, Formative Assessment, Summative Assessment, Reflection, Process Writing

Introduction

A study into the history of language teaching and pedagogy suggests that there has been a close interrelationship between language teaching and testing. This interrelationship has been so close that it has been impossible to work in either without taking the other into account (Farhady, Jafarpoor, & Birjandi, 1994). The conceptual framework guiding the development of curriculum and instruction practices in teaching of English as a second language (ESL) has undergone significant modification during the last fifteen years. This shift in pedagogical theory has resulted in the increasing use of student-centered communicative approaches in the classroom. These approaches include process writing, process reading, communicative competence, and whole language (Goodman, 1989; Heymsfeld, 1989) and are distinguished from prior practices by their focus on language function and meaning and the processes of learning.

Proponents of process-oriented curricula and instruction concur that traditional assessment techniques are often incongruent with current ESL classroom practices. Standardized testing is seen as particularly antithetical to process learning and has been attacked vigorously not only in ESL, but throughout the field of education (Rothman, 1990; Shepard, 1989). Because of the incompatibility of process learning and product assessment and the discrepancy between the information needed and the information derived through standardized testing, educators have begun to explore alternative forms of student assessment techniques, one of which is portfolio assessment. It is increasingly cited as a viable alternative to standardized testing (Wolf, 1989).

As language teaching has moved in the direction of learner-centered approach, testing and assessment have begun to incorporate the measures that reflect the type of tasks which are more learner-centered and authentic (Birenbaum, 1996). However, it seems that the current trend in some educational systems is that teachers provide students with a single test at the end of the course. Whereas standardized tests serve a purpose in education, they are neither infallible nor sufficient. Many educators like Flood and Lapp (1989) acknowledge that any "single score . . . almost always fails to accurately report student overall progress" (p. 509). A single measure seems incapable of estimating the diversity of skills, knowledge, processes, and strategies that combine to determine student progress. Therefore, the multifaceted nature of language proficiency makes it difficult for any single test to measure it (Genesee & Upshur, 1996).

To compensate for limitations associated with using standardized tests, professionals in ESL education often use a combination of formal and informal assessment techniques for monitoring student language development. One of the assessment techniques which represent the combined and integrated form of formal and informal assessment is

portfolio assessment. What distinguishes assessment from testing is that assessment integrates instruction into itself in the sense that assessment is at the service of learning and measuring, but the area of testing can be characterized by a separation of instruction and testing activities, and by measuring products solely in the form of a single test (Wolf, Bixby, & Glen, 1991 p. 53).

The last two decades have also seen substantial growth in the use of reflection. Reflection is seen by many as a form of meta-cognition or "thinking about thinking" (Swartzendruber-Putnam, 2000). English (1998) states, "Metacognitive ability is the determining factor that enables writers to adjust accordingly to varying task demands and contexts. It facilitates the selection, allocation of techniques, and strategies for successful task completion" (p.5). Although many scholars have made statements about the value of reflection (e.g. Swartzendruber-Putnam, 2000; English, 1998), Swain (2002) sums up the generally accepted benefits of reflection: "Reflection enables us to evaluate experience, learn from mistakes, repeat successes, revise, and plan" (p.12).

Research Questions and Hypotheses

The researcher intends to seek answers to the following questions:

1. Do reflective activities for the portfolio have any effect on the improvement of students' process writing?
2. Do reflective activities for the portfolio improve students' writing quality in terms of depth (It enables students to show quality work, which is done without pressure and time constraints, and with the help of resources, reference materials and collaboration with others.), breadth (A wide range of skills can be demonstrated.)

- and growth (It shows efforts to improve and develop, and demonstrates progress over time)?
3. Does portfolio assessment with its multiple perspectives on the student's performance contribute to the assessment process?
 4. Do reflective activities for the portfolio have any effect on the student's growth over time, including the student's abilities, knowledge, skills, and attitudes?

The researcher has put forward the following hypotheses:

1. Reflective activities for the portfolio do not have any significant effect on the improvement of students' process writing.
2. Reflective activities for the portfolio do not have any specific effect on students' writing quality.
3. Portfolio assessment does not contribute to the assessment process.
4. Reflective activities for the portfolio do not have any significant effect on the student's growth over time, including the student's abilities, knowledge, skills, and attitudes.

Literature Review

Portfolio Assessment

The concept of portfolio assessment is not new. Portfolios originated with artists' collections of their works and have long been used to demonstrate competencies. In response to the need for alternative and more authentic assessment practices, portfolios have become a common alternative to the traditional assessment methods (Mayer & Tusin, 1999). Based on the constructivist theories, which advocate that learning has to be constructed by the learners themselves, rather than being imparted by the teachers,

portfolio assessment requires students to provide selected evidence to show that learning relevant to the course objectives has taken place. They also have to justify the selected portfolio items with reference to the course objectives (Steffe & Gale, 1995).

Biggs (1996) holds that the preparation of an assessment portfolio is an active process involving collecting, synthesizing and organizing possible relevant items to provide the best evidence of achievement of the learning objectives; a process that demands ongoing assessment, reflection and justification. There is also the assumption that during the process of preparing an assessment portfolio, learning is enhanced as students are encouraged to reflect on their experience, identify learning needs and initiate further learning (Harris, Dolan, & Fairbairn, 2001). Such an assumption, however, should be supported with empirical evidence if the full potential of portfolio assessment is to be realized.

Genesee and Upshur (1996) define portfolio as follows:

A portfolio is purposeful collection of students' work that demonstrates to the students and others their efforts, progress, and achievements in given areas. Student portfolios have been inspired by professionals such as photographers and architects as a means of keeping a record of their accomplishments to show to others. Second language portfolios can have a very specific focus, such as writing, or broad focus that includes example of all aspects of language development. Students should have their own portfolios, which can be a conventional file folder, a small cardboard box, a section of a file drawer, or some other such receptacle (p. 99).

They maintain that the value of portfolios is in the assessment of student achievement. They are particularly useful in this respect because they provide a continuous record of

students' language development that can be shared with others. Genesee and Upshur clearly state that reviewing portfolio can increase the students' involvement in and ownership of their own learning. The positive effects of portfolios student learning arise from the opportunities they afford students to become actively involved in assessment and learning.

Theoretical Background of Portfolio Assessment

The underlying philosophy of this alternative approach to evaluation is that students are encouraged to become more autonomous and to take more responsibility for their work, including the evaluation of it. The following is what Yancey (1992) thinks of assessment:

Assessment is no longer seen as a process where one party submits his or her work to another with no influence on how the work is performed or interpreted because . . . all the parties are bona fide participants, and the person whose performance is being assessed is more than an object of someone else's perusal (p. 18).

Also Belanoff (1994) believes that portfolio assessment promotes participation and autonomy by allowing students to select the work on which they will be evaluated to reflect on their work to take control of revision and have the opportunity to produce substantive revision to be granted the time to grow as writers; to take risks with their writing, and to seek advice from peers. The result is that evaluation becomes a positive force to encourage growth, maturity, and independence, rather than a means of pointing out deficiencies.

Self-Assessment and Peer Assessment

A study into the literature will make it evident that self-assessment and portfolios are intertwined (Farr & Tone, 1994). It is impossible to find an article that mentions self assessment that does not do it in the context of discussing portfolio assessment. Portfolio assessment is the only methodology that responds directly to the goal of training students to assess their own success (ibid). It incorporates collecting and reviewing artifacts, understanding progress through record keeping, documenting interests and preferences, conferencing with teacher and peers. It also combines instruction with assessment that allows for self-reflection and self-evaluation.

Students can become better language learners when they engage in deliberate thought about what they are learning and how they are learning it. In this kind of reflection, students step back from the learning process to think about their language learning strategies and their progress as language learners. Such self assessment encourages students to become independent learners and can increase their motivation (McMullan, 2006).

Crooks (2001) also maintains that self-assessment provides students with the opportunity to understand the grading system. They can eliminate the controversy regarding subjective grading and gain ownership in their learning process. When students are involved with self-assessment, they are better able to work with other students, exchange ideas, get assistance when needed, and be more involved in cooperative and collaborative language-learning activities. As these students go about learning, they begin to construct meaning, revise their understanding, and share meanings with others.

The benefits of incorporating peer assessment into the regular assessment procedures have been discussed in a number of studies. Peer assessment is believed to enable learners to develop abilities and skills denied to them in a learning environment in which

the teacher alone assesses their work. In other words, it provides learners with the opportunity to take responsibility for analyzing, monitoring and evaluating aspects of both the learning process and product of their peers. Research studies examining this mode of assessment have revealed that it can work towards developing students' higher order reasoning and higher level cognitive thought , helping to nurture student-centered learning among undergraduate learners, encouraging active and flexible learning and facilitating a deep approach to learning rather than a surface approach (Gibbs, 1992). Peer assessment can act as a socializing force and enhances relevant skills and interpersonal relationships between learner groups.

Reflection and Portfolio

Despite the prevalence of a constructivist and "dialogic" pedagogy used by many writing teachers, little has been written explicitly on the role of reflection in the writing classroom. In the preface to "Dialogic Classroom", Galin and Latchow (1998) identify three things that must happen for the successful establishment of a "Dialogic Classroom", the third being reflection (cited in Crooks, 2001).

Gallagher (2001) also maintains that reflection is a major component of portfolios as it helps students to learn from experience and practice, thereby helping them to bridge the theory-practice gap. He says through the reflective process students are able to identify gaps in knowledge and/or skills and competence, but also to reconfirm and document strengths, skills and knowledge.

Studies on Portfolio Assessment and Reflection

Although there have been a good number of studies on the alternative

assessments, there seems to be little research related to the effect of it along with reflection, which was employed in the current study. It appears that due to practicality and time issues, this technique has been neglected in studies on alternative assessment. However, some research has been done on reflection. Two experimental studies conducted by Fontana and Fernandes (1994), and Frederikson and White (1997) have shown that students who have opportunities to reflect on their work and self-assess themselves show greater improvement than those who do not. Other studies done by Mccurdy and Shapiro (1992), Sawyer, Graham and Harris (1992) also show performance gains related to reflection and self-assessment.

Methodology

Participants

This study was conducted with 20 intermediate-level male students at Shahid Sattari Air University. They were in the age range of 22 to 28 with different cultural background. As far as the writing instruction is concerned, they had already had some informal lessons on paragraph writing. Actually, due to the practical problems, no random selection procedure was made. During the semester, the researcher administered 10 tests after teaching a particular writing lesson which served as five pretests and five posttests.

Instruments

According to the purpose of the study, a number of instruments for collecting the relevant data were used. The writing tests, the instructional material used in the treatment sessions, assessment and rating checklists. They are presented and explained below.

Writing Tests

As Raimes (1983) suggests, in the process approach, students are trained to generate ideas for writing, think of the purpose and audience, and write multiple drafts in order to present written products that communicate their own ideas. This approach was also selected since portfolio assessment and process writing are considered to be natural partners, and that both show effort and development very clearly. The topic of the test was chosen based on the topics of some IELTS books.

Writing Test Scoring Criteria

Of the common scoring method, analytic method is the most objective and valid one (Farhady, et al., 1994) in that the learners writing components, which are content, organization, accuracy, and complexity, are scored based on criteria, and the mean score is considered the final score. In this study, the Weigle's (2004) essay scoring criteria was used to score the students' essays (See Appendix A). In this scale, composition profile is used to score the students' performance on writing components. Each paper is rated on these components. Two qualified raters, who were trained at Iran Language Institute, scored the papers, and the results were analyzed to estimate the inter-rater reliability.

Instructional Material

As the focus of the study was on the written performance of the learners, the researcher tailored some materials from related books like *Academic Writing Course* (Jordan, 1999), and *The Longman Handbook for Writers and Readers* (Anson &

Schwegler, 1998). The researcher used these materials for the writing package to teach in the class. The researcher developed a writing package which consisted of eighteen writing lessons. These lesson, which were basically based on Anson and Schwegler (1998), had the following main sections: Part 1: writing and reading; Part 2: revising and shaping writing; and Part 3, editing and proofreading. Some model paragraphs were also extracted and incorporated into the package. Different grammatical points, basics of paragraph writing, essay writing, and so forth were among the writing tips dealt with in the package.

Assessment and Rating Checklists

In addition to the above-mentioned instruments, the following assessment tools were also employed by the researcher in this study:

- Self/peer assessment with rating scales
- Checklist with criteria (such as: clear presentation, relevant vocabulary, correct spelling), depending on the task
- Guided reflection on the task

These criteria were used to make the students' portfolio-based teaching more objective. Students needed to familiarize themselves with the types of analytical questions which they might need to answer in order to organize their portfolio. The questions were: What work am I most proud of? What are my goals? How are my language goals changing over time? When do I know I've done good work? What does my portfolio reveal about me and my learning style?

Questionnaire

An already validated questionnaire was given to students to express their attitude to portfolio-based teaching and reflective activities. This questionnaire, which was adopted from McMullan (2006), has three parts: questions about personal and professional development; general statements; and portfolio effectiveness.

Design of the Study

The design used for this study is both quantitative and qualitative in nature. It is quantitative in that a quasi-experimental design has been adopted. Also the students' performances on writing have been described in terms of organization, content, accuracy, and complexity. Their performances were checked using repeated measure statistical technique to see the amount of students' improvement over the period of time.

The design is also qualitative in that students' growth over time has been described through the portfolio they have prepared. There have been some criteria students had to meet in their portfolio, which will be discussed later. Also, a validated questionnaire was given to students to express their attitude to portfolio-based teaching and reflective activities.

Procedure

On the first day of the course, as the researcher talked through the syllabus, he introduced students to the portfolio project. He explained that they would talk about the portfolio shortly after midterm. The researcher told them that focusing on the writing tasks at hand would be more than adequately preparing them for the project and that it was crucial that they keep everything they write in a folder. Although it was important for students to

know about the portfolio project from the outset, the researcher deferred discussing the portfolio any further at that time because he didn't want the portfolio to become an end product towards which students were self-consciously aiming their efforts for a grade. Up to the half of the semester, the researcher administered five pretests before giving treatment. These pretests were accompanied by teaching essay writing without requiring students to keep portfolio and do reflective activities.

In the second half of the semester, he provided students with a handout that explained the portfolio and told them why he was asking them to do it. Then, he made the objectives of the course as explicit as possible. To prepare students for this project, the researcher asked them frequently throughout the term to answer questions about their writing, to reflect on each others' writing, and to characterize their own writing and writing processes in terms of what they saw others doing. In addition, he designed writing assignments that asked students to think about the issue. Finally, the students took the tests along with their treatment. Like the pretests, these tests were scored by two raters the mean of which was considered the final score. The reason students took these tests in the order mentioned is that the researcher intended to see the students' growth in process writing over the period of time.

Data Collection and Analysis Procedures

Portfolios are collections of relevant work that reflect students' individual efforts, development, and progress over a designated period of time and can provide students, teachers, parents, and administrators with a broad picture of each student's growth over time. To do so, the following were considered for inclusion of data:

1. Cover Letter “About the author” and “What my portfolio shows about my progress as a learner” (written at the end, but put at the beginning). The cover letter summarizes the evidence of a student’s learning and progress.
2. Table of Contents with numbered pages.
3. Entries - both core (items students have to include) and optional (items of student’s choice). The core elements will be required for each student and will provide a common base from which to make decisions on assessment. The optional items will allow the folder to represent the uniqueness of each student. Students can choose to include “best” pieces of work, but also a piece of work which gave trouble or one that was less successful, and give reasons why.
4. Dates on all entries, to facilitate proof of growth over time.
5. Drafts of written products and revised versions; that is, first drafts and corrected/revised versions.
6. Reflections appeared at different stages in the learning process (for formative and summative purposes or both).

Besides these procedures, which were used for the qualitative part of the study, the students' scores on pretests and posttests were used to be analyzed through the statistical procedure of paired t test. Further, a validated questionnaire was given to students to express their attitude to portfolio-based teaching and reflective activities.

Result

For the quantative part of the study, the researcher used the paired-sample t test and time series to see if the application of reflection through portfolio had any effect on the students' writing. Before that, the researcher administered five pre-tests before giving

treatment. These pretests, which were administered every week, were accompanied by teaching essay writing without requiring students to keep portfolio and do reflective activities. The students were given treatment along with an administration of five post-tests. To see the students' possible improvement, every two tests were compared. The level of significance set in this study is .001. The reason for this is that *t* test is used for two comparisons only, and because there were 10 comparisons in this study (i.e. Pretest 1 with Pretest 2, Pretest 2 with Pretest 3, etc), the new level of significance was set as a result of applying Bonferroni test, having the previous level of significance divided by the number of comparisons: $.01/10 = .001$

The following is the description of the data gained by paired sample *t* test.

Table 1 Paired-Sample *t* Test for the Students' Pretest 1 and Pretest 2

	Paired Differences					T	df	Sig. (2-tailed)
	Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
				Lower	Upper			
Pair 1 pretest1 - pretest2	-.04250	.10036	.02244	-.08947	.00447	-1.894	19	.074

Paired Samples Statistics

Std. Error Mean	Std. Deviation	N	Mean	
.21863	.97776	20	5.9875	pretest1 Pair 1
.21384	.95634	20	6.0300	pretest2

Paired Samples Correlations

Sig.	Correlation	N	
.000	.995	20	pretest1 & pretest2 Pair 1

As shown in the table, there is not any significant difference between students' scores on the first and second pretests. The students' mean score moved from 5.98 on pretest 1 to 6.03 on pretest 2.

Table 2 Paired-Sample *t* Test for the Students' Pretest 2 and Pretest 3

		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
					Lower	Upper			
Pair 1	pretest2 - pretest3	-.07500	.10822	.02420	-.12565	-.02435	-3.099	19	.006

Paired Samples Statistics

Std. Error Mean	Std. Deviation	N	Mean	
.21384	.95634	20	6.0300	pretest2 Pair 1
.20979	.93821	20	6.1050	pretest3

Paired Samples Correlations

Sig.	Correlation	N	
.000	.994	20	pretest2 & pretest3 Pair 1

Table 2 indicates that there isn't any significant difference between students' Pretest 2 and Pretest 3. There is an improvement in the students' scores, moving from the mean score of 6.03 on Pretest 2 to 6.10 on Pretest 3.

Table 3 Paired-Sample *t* Test for the Students' Pretest 3 and Pretest 4

		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
					Lower	Upper			
Pair 1	pretest3 - pretest4	-.01750	.11502	.02572	-.07133	.03633	-.680	19	.504

Paired Samples Statistics

Std. Error Mean	Std. Deviation	N	Mean	
.20979	.93821	20	6.1050	pretest3 Pair 1
.19946	.89199	20	6.1225	pretest4

Paired Samples Correlations

Sig.	Correlation	N	
.000	.993	20	pretest3 & pretest4 Pair 1

As seen in table 3, there isn't any significant difference between students' scores on Pretest 3 and Pretest 4. There was a slight improvement 9 (only .02) on students mean score, moving from 6.10 on Pretest 3 to 6.12 on Pretest 4.

Table 4 Paired-Sample *t* Test for the Students' Pretest 4 and Pretest 5

		Paired Differences					t	Df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
					Lower	Upper			
Pair 1	pretest4 pretest5	-.03250	.09904	.02215	-.07885	.01385	-1.468	19	.159

Paired Samples Statistics

Std. Error Mean	Std. Deviation	N	Mean	
.19946	.89199	20	6.1225	pretest4 Pair 1
.20297	.90770	20	6.1550	pretest5

Paired Samples Correlations

Sig.	Correlation	N	
.000	.994	20	pretest4 & pretest5 Pair 1

Table 4 suggests that there isn't any significant difference between students' scores on pretests 4 and 5. The improvement on these tests is also slight. The mean score just moved from 6.12 to 6.15 on Pretest 5.

Table 5 Paired sample *t* Test for the Students' Pretest 5 and Posttest 1

		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
					Lower	Upper			
Pair 1	pretest5 posttest1	.28250	.20537	.04592	-.37862	-.18638	6.152	19	.000

Paired Samples Statistics

Std. Error Mean	Std. Deviation	N	Mean	
.20297	.90770	20	6.1550	pretest5 Pair 1
.18031	.80637	20	6.4375	posttest1

Paired Samples Correlations

Sig.	Correlation	N	
.000	.978	20	pretest5 & posttest1 Pair 1

As seen above, there is a significant difference between Pretest 5 and Posttest 1. The students' improvement is from the mean score of 6.15 on Pretest 5 to 6.43 on Posttest 1. This result implies that the students' improvement might be due to the treatment given.

Table 6 Paired-Sample *t* test for the Students' Posttes1 and Posttest 2

		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
					Lower	Upper			
Pair 1	posttest1 posttest2	.28750	.34254	.07659	-.44781	-.12719	3.754	19	.001

Paired Samples Statistics

Std. Error Mean	Std. Deviation	N	Mean	
.18031	.80637	20	6.4375	posttest1 Pair 1
.18777	.83972	20	6.7250	posttest2

Paired Samples Correlations

Sig.	Correlation	N	
.000	.914	20	posttest1 & posttest2 Pair 1

As illustrated in table 6, there is a significant difference between the performance of students on Posttest 1 and 2. This result implies that the students' performance improves

as time goes on; in other words, the students' improvement growth over time is significantly different.

Table 7 Paired-Sample *t* Test for the Students' Posttest 2 and Posttest 3

		Paired Differences				t	df	Sig. (2-tailed)	
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
					Lower	Upper			
Pair 1	postets2 postest3	.12000	.37781	.08448	-.29682	.05682	1.420	19	.172

Paired Samples Statistics

Std. Error Mean	Std. Deviation	N	Mean	
.18777	.83972	20	6.7250	postets2 Pair 1
.21070	.94227	20	6.8450	postest3

Paired Samples Correlations

Sig.	Correlation	N	
.000	.916	20	postets2 & postest3 Pair 1

Unlike Posttest 2, Posttest 3 doesn't show any significantly different improvement. Though there is an evidence of improvement, it's not significantly different. The mean score has moved from 6.72 on Posttest 2 to 6.84 on Posttest 3.

Table 8 Paired-Sample *t* Test for the Students' Posttest 3 and Posttest 4

		Paired Differences					T	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
					Lower	Upper			
Pair 1	posttest3 - posttest4	-.19500	.29509	.06598	-.33311	-.05689	2.955	19	.008

Paired Samples Statistics

Std. Error Mean	Std. Deviation	N	Mean	
.21070	.94227	20	6.8450	posttest3 Pair 1
.22886	1.02349	20	7.0400	posttest4

Paired Samples Correlations

Sig.	Correlation	N	
.000	.958	20	posttest3 & posttest4 Pair 1

Table 8 supports improvement in the students' performance on Posttest 4 compared to Posttest 3, though not significantly at $p < 001$ value. This table also implies that the students' improvement in performance over time is significantly different.

Table 9 Paired-Sample t Test for the Students' Posttest 4 and Posttest 5

		Paired Differences				T	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference			
					Lower	Upper		
Pair 1	postets4 postest5	.17500	.21244	.04750	-.27443	-.07557	3.684	.001

Paired Samples Statistics

Std. Error Mean	Std. Deviation	N	Mean	
.22886	1.02349	20	7.0400	postets4 Pair 1
.24720	1.10550	20	7.2150	postest5

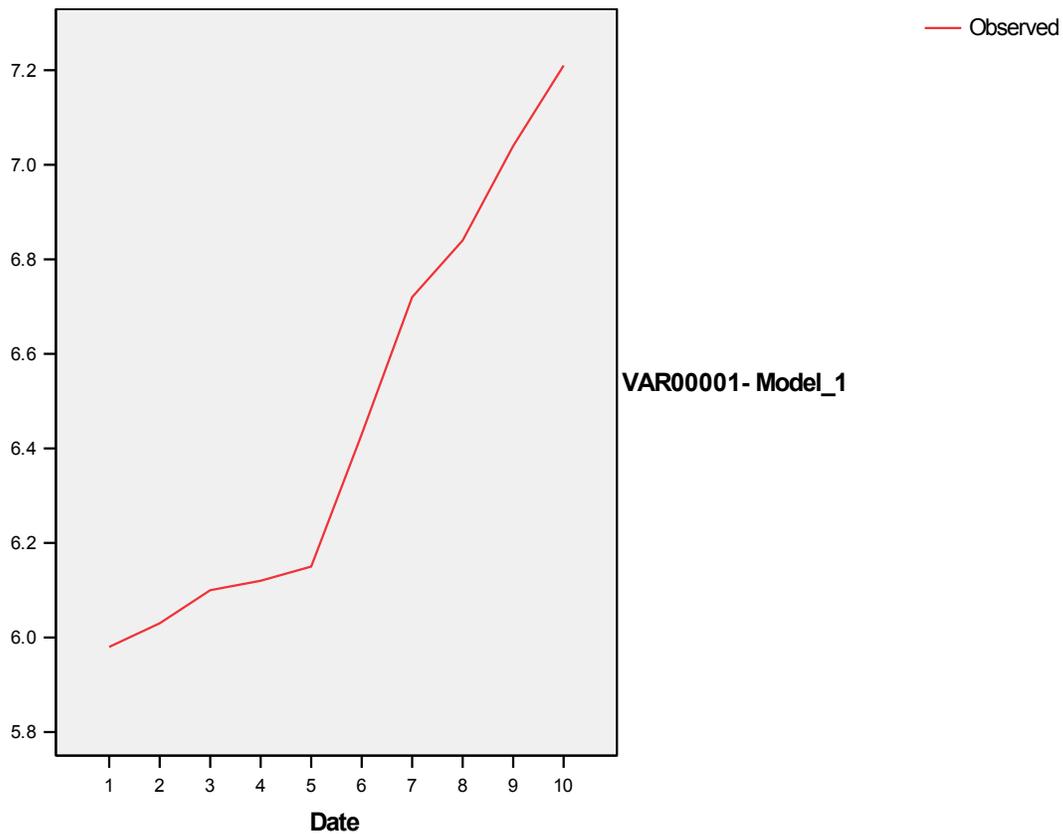
Paired Samples Correlations

Sig.	Correlation	N	
.000	.983	20	postets4 & postest5 Pair 1

And finally, table 9 also illustrates that Posttests 4 and 5 are also significantly different. This is another evidence of students' improvement over the period of experiment. The students' growth over time is also shown through the following diagram.

Graph 1 The Students' Growth Over Time

A



As shown in the diagram above, there was a gradual improvement from the mean score of 5.98 to 6.15 before they received the treatment. The improvement turned out to be significant by the time they received treatment, moving from the mean score of 6.43 on the first posttest to 7.21 on the last one.

As a result of the data analysis run above, it can be concluded that due to the treatment effect, the experimental group outperformed the control group. Therefore, the null hypotheses 1 and 2 are rejected.

Qualitative Studies

Students' Reflective Letter and Course Objectives

White (2005) put forth a new method of scoring called "Phase 2 Scoring ". This method requires that two documents be developed as part of assessment: first, a set of goals set by faculty or teacher for the particular course, program, or purpose for which the portfolio is submitted; the second one is a reflective letter to readers composed by the students showing that those goals have been met. This reflective letter allows the students to evaluate the course in terms of their own experience of it. Phase 2 scoring is highly dependent on the reflective letter. It is used as evidence that the goals have been met. It is now possible to give a reliable and reasonably quick reading to the portfolios in hand. In this part students' portfolios are discussed through reflective letters.

The course objectives we have provided students with were adopted from White (2005). The course objectives were divided into three main parts: goals of rhetorical knowledge; goals of process writing; and goals of conventions. The table below shows the percentage of achievement in each goal by students, which suggests students made a big achievement in portfolio program.

Table 10. The Percentage of Achievement in Each Goal by Students

Rhetorical knowledge	Process writing	Conventions
86%	94%	87%

Table 10 suggests that 94% of the students had writing portfolio with brainstorming, draft papers, revised papers, peer revision and self revision, and the final draft. The students' portfolio suggests the students had a great improvement in process writing. With the quantitative and qualitative data gathered, we can conclude that the treatment was effective. Therefore, null Hypotheses 1, 2, and 4 are rejected. Hypothesis 3 is also rejected since the students' portfolio helped the researcher have a greater view of students' improvement and assess their achievement more effectively.

Questionnaire

An already validated questionnaire was given to students to express their attitude to portfolio-based teaching and reflective activities. This questionnaire, which was adopted from McMullan (2006), has three parts: questions about personal and professional development, general statements, and portfolio effectiveness. Each is discussed below.

The analysis of the questionnaire shows that 89% of students agreed or strongly agreed portfolios helped them to take responsibility for their own professional development, and 78% agreed or strongly agreed that portfolios enhanced their reflective skills. Also 75% felt that portfolios helped them to become aware of their strengths and weaknesses, and that they helped them to become independent learners (77% agreed or strongly agreed). Even more students (80%) felt that portfolio helped to promote critical thinking, and more than this (81%) felt that portfolio helped them to improve their self-esteem. students felt that portfolios took a great deal of time to complete (73% agreed/strongly agreed) and caused them much anxiety (64% agreed/strongly

agreed). On the other hand, 75% of the students felt they had good reflective writing skills. Table 4.4.2.2 shows that students felt that portfolios took a great deal of time to complete (73% agreed/strongly agreed) and caused them much anxiety (64% agreed/strongly agreed). On the other hand, 75% of the students felt they had good reflective writing skills. 80% of respondents agreed or strongly agreed that they liked the portfolio as an assessment tool, and 81% agreed or strongly agreed that they liked the portfolio as a developmental learning tool.

When respondents were asked how effective they felt portfolios were, the majority said they were very effective. 72% felt that portfolios were very effective in helping them to learn from practice, and 59% felt that they helped them to be prepared for practice. When it came to the effectiveness of portfolios in assessing learning and competence, 66% felt that portfolios were very effective in assessing learning and 54% in assessing competence.

Discussion

As stated in the literature, portfolio encourages students to enhance their reflective skills (Grant & Dornan, 2001) and help them become aware of their strengths and weaknesses (Priest & Roberts, 1998). They help students to take responsibility for their own professional development and promote critical thinking (Wenzel et al., 1998). In addition, they help students to develop independent learning and increase their feelings of self-esteem and confidence (Harris et al., 2001). In this study also, they did help students to enhance their reflective skills and to develop a sense of responsibility for their own professional development. A possible reason why students in this study reported that portfolios helped them enhance

their reflective skills and their critical thinking skills could be due to the amount of support and guidance students felt they received with their portfolio use. If students do not receive ongoing support and guidance and regular feedback sessions on how to use the portfolio and on how to reflect, their reflective skills will remain shallow and the deeper critical thinking skills will not develop sufficiently.

Implication of the Study

In general, the results of this study have five main implications for syllabus designers, materials developers, and language teachers. First, this study may be helpful to syllabus designers because they can gain insights from the results of this study that in designing a syllabus, flexibility should be emphasized, because only in that case the students can be expected to actively participate in the teaching process. Portfolio assessment plays a significant role in incorporating the learner's suggestions and opinions into decision making and the instructors' opinions. This study has shown that such an approach is helpful. So syllabus designers should also consider and value learners' rights to formulate their own decisions, suggestions, and criticism while designing syllabuses. Therefore, the findings of this study suggest that language syllabuses should be flexible and consider the learner's participation in decision making.

Second, the study can also be useful for materials developers. That is, they should consider the learner's own performances and for the possibility of manipulating material if learners are more likely to make progress that way.

Third, language teachers can also benefit from the results of this study. This study may give language teachers the insight that incorporating formative assessment into the classes

helps both the teachers and the students identify their own strengths and weaknesses and provide modification where and when needed. Moreover, the implication that can be drawn is likely to provide language teachers with insightful guidelines to conduct portfolio assessment while teaching language skills and components in their own classes. In short, the results imply that helping learners to develop the habits of self-assessment and critical, reflective thinking and providing them with feedback in their learning are effective techniques to be incorporated into the writing classes.

Fourth, the findings of this study might be applicable to the assessment of all language skills including listening comprehension, speaking ability, and to the assessment of language components such as pronunciation, vocabulary, and grammar. However, many investigations are still needed to support this implication.

Fifth, this study may also have implications for assessing the progress of the students of translation. Mastery of different structures of both the foreign language and the native language and different semantic relationships of both languages in the process of converting one language into the other can be continuously checked through the use of portfolio assessment. Strengths and weaknesses in different parts and at different stages can be identified through the student's involvement in self-evaluation and reflection on their own work in the process of diary-keeping and portfolio assessment.

To this end, the findings of this study may be applied to all educational fields and in all educational settings where the aim is to raise student's consciousness and awareness of their own learning strategies and to help them develop some kind of criticality and reflectivity towards what and how they are taught and how they learn. Particularly, this study may pave the way for teachers to help their students proceed towards meaningful learning away from resorting to memorization which is an inadequate

technique at present.

Suggestions for Further Research

This study was conducted to university-level students. It can be fruitful if other research on the same issue is extended to pre-university level students such as high school students.

This study was conducted to investigate the effect of portfolio assessment and reflection on writing process. It can be a good idea for the interested researchers to investigate the effect of using portfolio and reflection on the improvement of other language skills such as listening, speaking, and sub skills such as grammar and vocabulary.

Because the students in this study were all male, a further area for research can be to investigate the relationship between gender and portfolio and self assessment to check the possible difference between male and female students' performances.

Research can also be conducted as to the effectiveness of faculty training and participation in the portfolio process and specific components of the mentoring process. The quality of the mentoring that faculty provide to their students should also be investigated.

References

Anson, Ch. M., & Schwegler, R. A. (1998). *The Longman handbook for writers and readers*.

New York: Longman.

Belanoff, P., & Dickson, M. (Eds.). (1991). *Portfolios: Process and product*. Portsmouth, NH:

- Boynton/Cook Heinemann.
- Biggs, J. B. (1996). The teaching context: The assessment portfolio as a tool for learning. In Biggs, J.B. (Ed.), *Testing: To educate or to select?* Education in Hong Kong at the Crossroads, (pp. 227-271), Hong Kong Educational Publishing Co.
- Birenbaum, M. (1996). Towards a plurastic approach to assessment. In M. Birenbaum & F. Dochy (Eds). *Alternatives in assessment of achievement. Learning, Processes and Prior Knowledge*, 3-31.
- Crooks, T. (2001). *The validity of formative assessment*. Paper presented to the British Educational Research Association Annual Conference, University of Leeds, September 2001.
- English, J. A. (1998). Moo-based metacognition: Incorporating online and offline reflection into the writing process. *Kairos*, 3, 1.
- Farhady, H.; Jafarpoor, A.; & Birjandi, P. (1994). *Testing language skills: From theory to practice*. Tehran: SAMT.
- Farr, R. & Tone, B. (1994). *Portfolio and performance assessment: Helping students evaluate their progress as readers and writers*. Orlando: Harcourt Brace & Co.
- Flood, J., & Lapp, D. (1989). Reporting reading progress: A comparison portfolio for parents. *The Reading Teacher*, 42, 508-514.
- Fontana, D., & Fernandes, M. (1994). Improvements in mathematics performance as a consequence of self-assessment in Portuguese primary school pupils. *British Journal of Educational Psychology*, 64 (3), 407-417.
- Frederiksen, J.R., & White, B.J. (1997). *Reflective assessment of students' research within an inquiry-based middle school science curriculum*. Paper presented at the annual meeting of the American Educational Research Association, Chicago, IL.
- Gallagher, P. (2001). An evaluation of a standard-based portfolio. *Nurse Education Today*, 21,

- 409-416.
- Galín, E. J., & Latchow, J. (1998). The dialogic classroom: Teachers integrating computer technology, pedagogy, and research. In K. B. Yancey (Ed.), *Portfolios in the writing classroom: An introduction*. Urbana, IL: NCTE.
- Genesee, F. & Upshur, J. (1996). *Classroom-based evaluation in second language education*. Cambridge: Cambridge University Press.
- Gibbs, G. (1992). Down with essays. *The New Academic*, 1 (2), 18-19.
- Goodman, K. S. (1989). Whole language is whole: A response of Heymsfeld. *Educational Leadership*, 46 (6), 69-70.
- Harris, S., Dolan, G. & Fairbairn, G. (2001). Reflecting on the use of student portfolios. *Nurse Education Today*, 21, 278–286.
- Heymsfeld, C.R. (1989). Filling the hole in whole language. *Educational Leadership*, 46 (6), 65-68.
- Jordan, R. R. (1990). *Academic writing course*. London: Collins E.L.T.
- McMullan, M. (2006). Students' perceptions on the use of portfolios in pre-registration nursing education: A questionnaire survey. *International Journal of Nursing Studies*, 43, 333-343.
- Mayer, D.K. & Tusin, L.F. (1999). Pre-service teachers' perceptions of portfolios: Process versus product. *Journal of Teacher Education*, 50 (2), 131–139.
- McCurdy, B.L., & Shapiro, E.S. (1992). A comparison of teacher monitoring, peer monitoring, and self-monitoring with curriculum-based measurement in reading among students with learning disabilities. *Journal of Special Education*, 26 (2), 162-180.
- Raimes, A. (1983). *Techniques in teaching writing*. Oxford: Oxford University Press.

Rothman, R. (1990). Ford study urges new test system to open the gates of opportunity.

Education Week, 2, 1-12.

Shepard, L. A. (1989). Why we need better assessments. *Educational Leadership*, 46 (7), 4-9.

Steffe, L. & Gale J. (1995). *Constructivism in Education*. Erlbaum: Hillsdale, NJ.

Swain, S. (2002). Studying teachers' transformations: Reflection as methodology. *The Clearing*

House, 72, 1, 40-52.

Swartzendruber-Putnam, D. (2000). Written reflection: Creating better thinkers, better writers.

English Journal, 1, 88-93.

Weigle . S.C. (2004). Integrating reading and writing in a competency test for non-native speakers

of English. *Assessing Writing*, 9, 27-55

White, E. M. (2005). The Scoring of Writing Portfolios: Phase 2. *College Composition and*

Communication, 56 (4), p. 584.

Wolf, D. P. (1989). Portfolio assessment: Sampling student work. *Educational Leadership*, 46

(7), 35-39.

Wolf, D., Bixby, J. & Glen, J. (1991). To use their minds well: Investigating new forms of

student assessment. *Review of Research in Education*, 17, 31-74.

Yancey, K. B. (1998). Reflection in the writing classroom. In K. B. Yancey (Ed.), *Portfolios in the*

writing classroom: An introduction. Urbana, IL: NCTE.

Appendix A: Essay scoring criteria, adapted from S.C Weigle (2004), *Assessing Writing*, 9 (27-55)

Rhetoric: Content	R h e t o r i c : Organ ization	Language: Accuracy	Language: Range and complexity
9-10	9-10	9-10	9-10
The treatment of the assignment completely fulfills the task expectations and the topic is addressed.	Clear and appropriate organization plan.	The essay is clearly written with few errors; errors do not interfere with comprehension.	The essay uses a variety of sentence types accurately.
Fully developed range of evidence for generalizations and supporting ideas is provided in a relevant and credible way.	Effective introduction and conclusion.	Includes consistently accurate word forms and verb tenses.	Uses a wide range of academic vocabulary.
Uses ideas from source text well to support thesis.	Connections between and within paragraphs are made through effective and varied use of transition and other cohesive devices.	Word choices are accurate and appropriate.	Source text language is used sparingly and accurately incorporated into writer's own words.
7-8	7-8	7-8	7-8
The treatment of the assignment fulfills the	Clear	The essay is clearly written but contains some errors which do not interfere	The essay uses a variety of sentence types.

task expectations completely and the topic is addressed clearly.	organization plan	with comprehension.	
Evidence for generalizations and supporting ideas is provided in a relevant and credible way.	Satisfactory introduction and conclusion	The essay may contain some errors in word choice, word form, verb tenses, and comprehension.	Good range of vocabulary used with at most a few lapses in register.
Ideas from source text used to support thesis.	Satisfactory connections between and within paragraphs using transitions and other cohesive devices.	----- -----	Some language from the source text may be present but is generally well incorporated into writer's own words.
5-6	5-6	5-6	5-6
The treatment of the assignment minimally fulfills the task expectations; some of the task may be slighted.	Adequate but simplistic organizational plan	Is generally comprehensible but contains some errors that distract the reader; at most a few errors interfere with comprehension.	Somewhat limited range of sentence types; may avoid complex structures.
Some evidence for generalizations and supporting ideas is provided	Introduction and conclusion present but may be brief.	The essay may contain some errors in word choice, word form, verb tenses, and comprehension.	Somewhat limited range of vocabulary
Ideas from source text are included but may not be explicitly acknowledged as such.	connections between and within paragraphs occasionally missing	----- ----- ----	May include extensive language from source texts with an attempt to incorporate text own language

3-4	3-4	3-4	3-4
The treatment of the assignment only partially fulfills the task expectations and the topic is not always addressed clearly.	Organization plan hard to follow	Contain many errors; some errors may interfere with comprehension.	Uses a limited number of sentence types.
Evidence for generalizations limited, and supporting ideas is insufficient and irrelevant.	Introduction and conclusion may be missing or inadequate.	Includes many errors in word choice, word form, verb tenses, and complementation.	----- ----- -----
May not include ideas from source text, or may consist primarily of ideas from source text without	connections between and within paragraphs frequently missing	----- -----	Extensive use of source text language with little integration with writer's words.
1-2	1-2	1-2	1-2
The treatment of the assignment fails to fulfill the task expectations and the paper lacks writing.	No apparent organization plan	Contain numerous errors that interfere with comprehension	Use simplistic and repetitive vocabulary that may not be appropriate for academic focus.
Evidence for generalizations and supporting ideas is insufficient and irrelevant.	Introduction and conclusion missing or clearly inappropriate.	Includes many errors in word choice, word form, verb tenses, and complementation.	Does not vary sentence types sufficiently.
----- ----- ---	Few connections between and within paragraphs	----- -----	May rely almost exclusively on source text language

Understanding the influence of L1 and lexical aspect in temporal acquisition: Quantitative and qualitative studies.

Sun Xiao-zhao & Du Juan

School of Foreign Languages, Dalian University of Technology

Bio Data:

Sun Xiao-zhao is an assistant professor in the School of Foreign Languages at Dalian University of Technology. He obtained his Ph.D. in English education from Hiroshima University. His current research interests include temporal acquisition, vocabulary acquisition and qualitative analysis of language learning process of EFL learners.

Du Juan is a lecturer in the School of English Studies at Dalian University of Foreign Languages. She has been teaching since 1996. Her research interests include translation studies, syllabus design and content-based instruction in classroom teaching.

This research was supported by a Scientific Research Start-up Fund for Returned Scholars, by the Ministry of Education in China.

Abstract

The two presented studies aim to make a comparative investigation on L1 influence and lexical aspect effect in temporal acquisition by Chinese and Japanese EFL learners. By using a mixed methods approach, two studies were conducted in order to present a more comprehensive and in-depth analysis of learners' performance in temporal marking. While Study One was a cross-sectional using a written error recognition and correction task to look at factors at work in EFL learners' temporal performance, Study Two attempted to examine the metacognitive process of their tense-aspect interpretation by

means of qualitative data obtained from retrospective interviews. As revealed in the results, L1 influence was found to be an active factor in Japanese learners' progressive marking performance. However, L1 influence was not found in Chinese learners. With regard to the lexical aspect effect, a strong progressive-activities association predicted by the Aspect Hypothesis was not found in the results. Through learners' verbal reports, Study Two provides qualitative evidence for the existence of L1 influence and learners' awareness of lexical aspect in tense-aspect performance. Detailed discussion is made on these findings as well as their pedagogical implications.

Keywords: temporal morphology; verbal semantics; L1 influence

Introduction

The study of tense-aspect morphology has always been the focus of language teaching and pedagogical materials. In response to the pedagogical necessity, studies on the acquisition of tense and aspect have greatly grown since 1980s. As noted by Bardovi-Harlig (2000), the development of research on the acquisition of systems of temporal expression reflects the development of research in second language acquisition in general.

In both L1 and L2 fields, the Aspect Hypothesis has long remained the major research interest in studies on tense-aspect acquisition. Focusing on the predictions of the Aspect Hypothesis, many studies have been conducted on L2 temporal acquisition in both naturalistic and instructed environments. As the most robust prediction of the Aspect Hypothesis, the attachment of perfective past with telic verbs gained extensive support in L2 studies (Bardovi-Harlig & Reynolds, 1995; Robison, 1995; Collins, 1999; Salaberry, 1998; Shirai & Kurono, 1998). However, counter evidence still exists. In his study, Rohde (1996, 1997) reported the distribution of inflected *versus* uninflected verb forms in past-time contexts. And he further argued that the use of past in his study is not determined by the verb's inherent aspect, which is clearly a challenge to the Aspect

Hypothesis. Housen (2002) also provided counterexamples to the Aspect Hypothesis. His longitudinal study of a child with L1 Dutch learning English showed a lack of the early association between achievements and simple past marking proposed by the hypothesis.

The association between progressive and activities is also well supported in L2 studies. A large number of cross-sectional studies of English confirmed this pattern (Bardovi-Harlig & Reynolds, 1995; Robison, 1995). However, while past-achievements association remains very robust in both L1 and L2 studies, more counter arguments are voiced with regard to progressive-activities association. Both Rohde (1996) and Robison (1990) reported non-targetlike overextensions of the progressive inflection –ing to state verbs, such as *liking, * loving and *smelling. Another example also comes from Rhode (1996) who found that his learners attached the progressive marker –ing more with achievements than with activities.

Besides, as Salaberry and Shirai (2002) point out, in tense-aspect acquisition studies, a simple form-meaning correlation is only part of the larger picture, as many other factors have to be considered. One of them is L1 influence, which also serves partially as a response to results that are incompatible with the Aspect Hypothesis. Much evidence has been found in support of L1 influence in tense-aspect acquisition (Robison, 1995; Andersen & Shirai ,1996; Cai, 2003), yet counter arguments can still be heard. (Bardovi-Harlig, 2000).

As mentioned above, although the hypothesis is claimed to be universal (Shirai & Kurono, 1998), research interest to test its universality has been maintained since its establishment. Besides, in the field of L2 acquisition of tense-aspect morphology, controversies still remain as to what factors influence the learner language of tense-aspect morphology, especially with regard to L1 influence. Although L1 influence has always

been a research focus in SLA, the number of studies exploring this issue in tense-aspect acquisition is still few. Even those researchers who have addressed this topic limited their studies on single language subjects (Salaberry, 1999; Collins, 2002; Cai, 2003). So it is of great importance to include two or more comparable language groups with their specific tense-aspect features taken into account. Just as Slabakova (2002) states in his review article, the effect of a learner's native language on his or her acquisition of aspectual properties in a second language has been curiously neglected so far. Much more precise research questions can be formulated if L1 transfer is taken into account and only properties that differ in the L1 and the L2 are investigated. In the present studies, by including two language groups, close examination was made on L1 influence based on more specified research questions.

Moreover, although some cross-sectional researches have obtained much solid evidence on learners' temporal acquisition, it is risky to base conclusions merely on the basis of quantitative data. This is particularly true with regard to Japanese EFL learners' acquisition and interpretation of the L2 English progressive aspect; although, both Garielle (2005) and Sun (2006) point out in their cross-sectional research that Japanese learners are subject to L1 influence in their L2 aspectual acquisition. Therefore, qualitative data reflecting learners' mental representation of the target forms are needed to provide validity to the previous arguments. Within the researcher's knowledge, very few studies have been done on the metacognitive process of learners' tense-aspect interpretation, irrespective of the one done recently by Liskin-Gasparro (2000) and Collins (2005). In her study of eight advanced learners of Spanish as a second language, Liskin-Gasparro (2000) found that they were successful at eliciting information from learners on the discourse, semantic, and instructional input factors that may have

influenced their performance on two oral narrative tasks. Collin's (2005) study reports on the findings from the retrospective interview protocols which encouraged learners to reflect on their performance on the controlled tasks. Two key factors emerged from the verbal reports: learners' awareness of semantic categories and their sensitivity to frequency effects in instructional input. Both appeared to constrain learners' access to new knowledge that would allow their interlanguage to develop to more advanced levels. Therefore, by using both a controlled written task and a retrospective interview, the present research is a relatively new attempt to explore the factors influencing the interlanguage development of tense and grammatical aspect.

Rationale

In order to get a better understanding of the present research focus, it is necessary to have some knowledge of the basic related concepts.

1. The Aspect Hypothesis

The Aspect Hypothesis is based on a theory of inherent lexical aspect. According to Vendler (1967), verbs can be classified into four types based on different semantic features in terms of inherent lexical aspect: states, activities, accomplishments and achievements, based on different semantic features, static / dynamic, durative / punctual and telic / atelic. This distinction is illustrated in Table 1. The static / dynamic distinction contrasts a static situation and a dynamic one. A dynamic situation can be either durative or punctual. A punctual situation can only persist for a very short time period. In contrast, a durative situation describes a relatively stable process. The telic / atelic distinction focuses on whether or not the situation has an end point. Thus, telic situations, once they

begin, progress step by step towards their climaxes when they finally stop. In the present study, states, activities and achievements will be the target verb types.

Table 1 Distinction of Inherent Lexical Aspect

	States	Activities	Accomplishments	Achievements	
punctual	–	–	–	+	
telic	–	–	+	+	
dynamic	–		+	+	+
examples	seem;want; feel; taste	talk; sing; walk;write	build a house; paint a picture	begin; finish; break; kill	

The Aspect Hypothesis claims that the learner’s selection of verbal morphology is related to the inherent lexical semantics of the verb phrase. Moreover, the hypothesis makes clear predictions about the distribution and direction of tense-aspect morphological markings (i.e. past with achievements, present with states, and progressive with activities).

2. Tense-Aspect System of English, Japanese and Chinese languages

In terms of tense, Chinese is well accepted as a tenseless language with no morphological inflections to mark time (Smith, 1991). Instead, it relies heavily on temporal adverbs and aspectual markers. In contrast, both English and Japanese are tense languages (Shirai, 1998). Thus, it is common to hear a Chinese student say such a sentence as, ‘I go to school yesterday.’ The following are examples of tense formation in these three languages. In both English and Japanese, the past tense is expressed by verb inflections.

However, in Chinese the perfective aspect marker is used to indicate the past.

English: Yesterday Tom **bought** a book.

Chinese: Zuo tian Tom **mai-le** yibenshu.

Yesterday Tom buy-ASP a book. ('le' is a perfective aspect marker)

Japanese: Kino Tom-wa hon-o katta.

Kino Tom-TOP book-ACC buy-PAST.

Note: ASP = aspect marker; TOP = topic marker; ACC = accusative case marker; PAST = past tense marker

Since tense marking is absent in Chinese, it is interesting to observe how the absence of the target grammatical item might influence Chinese learners' performance on past tense marking.

In terms of grammatical aspect, all three languages employ an aspect system to indicate time relation. With regard to the imperfective aspect, while all three languages employ aspect markers which function in some similar ways, Japanese shows an idiosyncratic feature which is absent in both English and Chinese. As such, there will be a brief discussion of the Japanese imperfective aspect marker *-te i* in comparison to English and Chinese.

As complex an aspect marker as *te iru* is, it should be noted here that the general meaning that *te iru* conveys is 'focus on the durative phase of a situation', and also that *te iru* can be generally characterized as 'durative imperfective'. This overarching meaning is always there, and depending on the different parameters (syntactic, semantic, and pragmatic), distinct senses such as progressive and resultative are instantiated (Shirai,

2000). When we make English imperfective aspect as a reference, it is worth mentioning that Japanese imperfective aspect marker *te iru* is not equal to the corresponding English *-ing* because *te iru* semantically allows more interpretations than *-ing* including progressive, resultative, perfect and habitual (Shirai, 1998). On the other hand, in terms of English progressive aspect marker *-ing*, *te iru* is the closest corresponding grammatical item in Japanese. The present paper will not discuss the semantics of progressive aspect in detail. Instead, it is to focus on the interaction between the imperfective aspect marker and verbal inherent aspect. In this respect, the Japanese imperfective aspect marker *-te i* behaves similarly to English and Chinese in that when combined with activity and accomplishment verbs, it mainly denotes a progressive meaning. However, Japanese differs from the other two languages in some important ways.

An important difference between Japanese and the other two languages is that Japanese can refer to a resultative state using the imperfective aspect marker *-te i* on achievement verbs, while no such meaning can be rendered in both English and Chinese. Since Japanese aspect marker *-te i* allows for more interpretations, some achievement verbs that are ungrammatical with progressive marking in both English and Chinese are compatible with *-te i*. Therefore, the achievement anomaly is missing from Japanese (Shirai, 1998).

For example:

Japanese: Tom-wa kare-ni kizui-*te iru*. (Japanese: resultative state)

Tom-TOP him-ACC notice-ASP-NONPAST.

English: *Tom is notice-*ing* him.

Chinese: *Tom zhu yi dao-*zhe* him.

Tom notice-ASP him.

In Japanese, there is only a very short list of state verbs compared with English and Chinese. Most English state verbs are often expressed in Japanese by the combination of achievement verbs and *-te i*, while in English and Chinese, states are described by state verbs that do not allow for any progressive marker in most cases. For example:

Japanese: Tom-wa kare-o shitte iru. (Japanese: resultative state)

Tom-TOP him-ACC know-ASP-NONPAST.

English: *Tom is knowing him.

Chinese: *Tom ren shi-zhe ta.

Tom know-ASP him.

From the above discussion, it is clear that Chinese and Japanese differ from English in some unique ways. On the one hand, Chinese shows an absence of tense inflection; on the other hand, Japanese allows for the imperfective aspect marker *-te i* to be combined with achievement verbs without the concept of achievement anomaly (as found in English and Chinese). In terms of EFL study, these special characteristics provide ground for an interesting research perspective, that is, Chinese learners have to acquire past tense marking which is new to them while Japanese learners have to narrow their L1 knowledge about combinations between the imperfective marker and verbal lexical aspect. Although we have to seriously consider the fact that, semantically speaking, progressive aspect is more complex than past tense when our focus is not on making a cross-item comparison but on comparing learners' performances on these two items separately across two nationality groups, it can still help further our understanding of how

the absence of L2 grammatical knowledge in L1 and partial fit between L1 and L2 grammatical knowledge might affect learners' perceptions of the target language.

Research Design and Research Questions

Since the researcher has conducted a series of experiments to investigate EFL learners' temporal marking from a comparative perspective, the present studies include two parts: Study One and Study Two, which aim to present a more comprehensive and in-depth analysis of learners' performance in temporal marking by using a mixed methods approach. Mixed model research refers to the use of both quantitative and qualitative strategies in several or all phases of the research study. The use of mixed methods is becoming increasingly popular, for it allows the researcher to answer both exploratory and confirmatory questions (Teddlie & Tashakkori, 2003). Study One was a cross-sectional study while Study Two tries to examine the metacognitive process of learners' tense-aspect interpretation by means of retrospective interviews. The interview was not given as a follow-up in Study One mainly because it had to be conducted immediately after the grammar test to avoid time lapse as much as possible. Therefore, individual grammar test and interview were considered more appropriate. As a replication of Study One, Study Two served as an attempt to obtain both exploratory evidence through interviews and more face validity through administering the test to another small group of subjects. As a whole, the following research questions were proposed, with the first two targeted in both studies and the last targeted in Study two.

1. Is there any L1 influence on Japanese learners' performance on progressive errors since in Japanese the imperfective aspect marker *te iru* allows wider

interpretation? And is there any L1 influence on Chinese learners' performance on base form errors since tense inflection is absent in their native language? (Study one and two)

2. Is there any lexical aspect effect found in the result? (Study one and two)
3. Are EFL learners aware of verbal lexical aspect? And is there any cognitive evidence of L1 influence found in learners' verbal reports, particularly with regard to Japanese learners' performance on ungrammatical progressive states and achievements? (Study two)

Next, results of the two studies will be presented and discussion will be made with regard to the research questions.

Study One

Subjects

The participants consisted of two groups of university students, 45 Japanese learners and 45 Chinese learners. Both groups of learners are 3rd-year students majoring in English at a national university. Besides, two groups share much in common in terms of study background. They are both in an acquisition-poor learning environment. Although reforms in English education have been carried out for some years, the mainstream English pedagogy still remains teacher-centered, textbook-centered and test-centered.

A language proficiency test based on the pre 1st and 2nd Grade English Language Proficiency Test of Japan was administered to both national groups, each of which was accordingly divided into two subgroups in terms of the language proficiency. An ANOVA analysis was made to ensure its validity. The results of language proficiency test

showed no difference between the two national groups: $F(1,86)= 0.417, p>.05$, while a significant difference was found between the two proficiency level groups: $F(1,86)=154.198, p<.001$.

Instrument and Scoring

In the present study, a 20-item error identification and correction test was designed with the test score as the dependent variable, and L1(2 levels), language proficiency (2 levels), lexical aspect(3 levels) and error category (2 levels) as the independent variables. Taking into account the above-mentioned linguistic difference, an equal number of verbs representing 3 lexical aspects (activities, achievements, and states) were included with 8 verbs for each lexical aspect which are equally distributed across two error categories: base form and progressive. Besides the 24 target verbs, 6 distractors were also included. Both the lexical aspects of the target verbs and the answers to the test were determined by a native linguist. All the target base-form verbs are ungrammatical in the test and correct answers are past tense forms. With regard to progressive errors, activity verbs in past tense are given in the test and verbs with progressive marking are the expected correct answers; while on state and achievement verbs progressive markings are ungrammatical and verbs in past tense are the correct forms. All the target verbs used are given in Appendix 1.

The learners were scored according to their performance on both identification and correction of the target error forms, each task accounting for one point. For example, if he or she could identify 'hearing' in 'he was hearing loud voices' as the erroneous item, one point would be given. In addition, if he or she succeeded in correcting 'hearing' as 'heard' which is the accurate grammatical form, another point would be attributed. Furthermore,

when the attempted form was not correct, but showed knowledge of past tense – for instance, ‘writed’ or ‘written’ for wrote; ‘talking’ for ‘was talking’ – half a point was deducted. Therefore, the full score for one verb is 2 points, which also serves as the basis for the calculation of mean scores in the following analysis. Sample items of the test are shown in Appendix 2.

Results and Discussion

Two ANOVAs were conducted on the test scores of base form category and progressive category respectively. Before we go into detailed discussion, results of the two ANOVAs will be reported.

In terms of progressive category, an ANOVA was made on the progressive category score with L1 (2 levels), language proficiency (2 levels), and lexical aspect (3 levels) as the independent variables. As the results revealed, the main effects of L1 and lexical aspect were found significant (L1: $F(1,86)=40.74, p<.001$; lexical aspect: $F(2,172)=7.56, p<.001$). Besides, interaction effect between L1 and lexical aspect was also significant ($F(2,172)=5.54, p<.005$).

In terms of base form category, another ANOVA was made on the base form category score with L1 (2 levels), language proficiency (2 levels), and lexical aspect (3 levels) as the independent variables. As the results revealed, the main effects of L1 and lexical aspect were found to be significant (L1: $F(1,86)=13.58, p<.001$; lexical aspect: $F(2,172)=10.45, p<.001$).

In the following section, a more in-depth discussion will be made on lexical aspect effect and possible L1 influence.

Lexical aspect effect

In the error task, past-tense activities were given in progressive context, while progressive states and achievements were given in past tense. Therefore, it is interesting to observe whether learners could be more successful in detecting errors on activity verbs and adding progressive markings to activity verbs in line with the progressive-activities association predicted by the Aspect Hypothesis.

As revealed by the analysis on the progressive category, lexical aspect effect was confirmed by both the main effect and interaction effect. With a close look at the main effect, it is interesting to note that learners performed best on achievement verbs, recognizing most errors in this category instead of the activity verb category. According to the Aspect Hypothesis, progressive markings are predicted to emerge with activities and spread through to achievements. It is also predicted that progressives will not be inappropriately used with statives (Andersen & Shirai, 1996; Shirai & Kurono, 1998). Following the hypothesis, the researcher expected learners to perform best on activities because of its strong tendency to be attached with progressive marking. Since in the present task errors on the activities category were given in progressive context, it should have been easier for learners to detect the unmarked errors on activities. However, contrary to the initial assumption, activities were not the most recognized in the present experiment.

Table 2 Basic statistics on progressive error category across Lexical Categories

	1 (States)	2 (Achievements)	3 (Activities)
N	90	90	90
Mean	1.43	1.60	1.44
SD	0.34	0.31	0.37

Table 3 Multiple comparison results on progressive error category across lexical categories

pair	r	nominal level	t	p
2 - 1	3	0.016	3.500	0.0005
2 - 3	2	0.033	3.303	0.0011
3 - 1	2	0.033	0.197	0.8437

MSe=0.111747, df=172, significance level=0.05

1=States; 2=Achievements; 3=Activities

One of the reasons might be that ungrammatical progressive markings on states and achievements were more conspicuous due to the sharp difference in their semantics. However, even if this was true, it is still surprising why activities failed to have a strong association with progressive markings as the Aspect Hypothesis predicted.

As the results revealed in the error task, the lexical aspect of activities curiously blocked the grammaticalization of progressive aspect instead of facilitating it. It seems that when progressive morphology was not provided and learners did not have to allocate their attention to the form in the sentence comprehension, they displayed a weakness in grammaticalizing activities in progressive marking. This was possibly caused by the heavy semantic overlapping between activities and progressive aspect. As we know, activities are set apart from the other dynamic verbs in that they describe actions or events that have inherent duration with no specific endpoint and display the same degree of homogeneity, such as run, play, dance, sleep, talk (Smith, 1983). Progressive aspect, on the other hand, is also defined essentially by continuousness (Comrie, 1976) or “action in progress” (Shirai & Andersen, 1995). Thus in the error task, the verbal semantics of activities seemed to have satisfied learners’ in-progress interpretation of the contextual meaning which should be grammatically expressed by progressive marking. According to some studies on language input and acquisition, it has been found that language learners

cannot process target language input for both meaning and form at the same time. When they are not pushed to notice the form, they process input for meaning only and do not attend to specific forms, and consequently fail to process and acquire them (Skehan, 1998). This could partially explain our present findings.

In addition, learners' failure to add progressive markings to activities in the error task also coincided with another well-accepted finding in temporal acquisition studies. According to many researches, at the beginning stages of acquisition, previous to the encoding of temporality with verbal morphology, temporal information is provided by lexical semantics, adverbs, calendric expressions, principles of discursive organization, and overall text structure (Von Stutterheim, 1991; Klein, 1994). Taking lexical aspect into consideration, we could argue that overlapping semantics between progressive and activities reinforced learners' common tendency in their language acquisition, such as process meaning only in the input and use of lexical devices in temporal marking. However, considering the fact that subjects in Study One were intermediate university learners, it is interesting to see that both lexical aspect effect and the other universal L2 acquisition constraints could be quite persistent.

These findings on activities remind us that caution is needed in interpreting the Aspect Hypothesis predictions, since many other factors might also be at work such as task effect, L1 influence, etc.

L1 influence

In terms of progressive errors, L1 influence was found in both the main effect and interaction effect between L1 and lexical aspect. Japanese learners accepted more error items than Chinese learners on states and achievements, while no significant difference

appeared on activity verbs (states: $F(1,258)=38.53$, $p<.001$; achievements: $F(1, 258)=18.46$, $p<.001$). This indicates a strong support of L1 influence on Japanese learners.

Table 4 Basic Statistics on progressive error category across lexical categories of both national groups

		States	Achievements	Activities
China	Mean	1.66	1.77	1.51
N=45	SD	0.37	0.31	0.38
Japan	Mean	1.19	1.44	1.37
N=45	SD	0.32	0.31	0.37

Table 5 Simple main effects of L1-lexical aspect interaction on progressive error category

effect	SS	df	MS	F
A(c1)	4.892	1	4.892	38.533
	0.000			
A(c2)	2.344	1	2.344	18.463
	0.000			
A(c3)	0.413	1	0.413	3.259
	0.072			
error		258	0.1269	
C(a1)	1.500	2	0.750	6.715
	0.001			
C(a2)	1.426	2	0.713	6.381
	0.002			
error		172	0.1117	

A: L1; C: lexical aspect; c1: states; c2: achievements; c3: activities
significance level=0.05

In contrast, in terms of the base form error category Chinese learners did not show any sign of negative L1 influence, outperforming Japanese learners in recognizing and correcting more base form errors across three lexical aspects although tense inflection is absent in their native language. In the following section an in-depth discussion will be made on why L1 influence was found in the Japanese group while being absent in Chinese group.

Table 6 Basic Statistics on base form error category across lexical categories of both national groups

		States	Achievements	Activities
China N=45	Mean	1.72	1.83	1.59
	SD	0.26	0.28	0.30
Japan N=45	Mean	1.57	1.65	1.49
	SD	0.30	0.24	0.38

As the results showed, Japanese learners underperformed on progressive marking items, especially on progressive achievements and states. Keeping in mind the linguistic background mentioned earlier, we might argue that their performance on progressive markings on achievements and states provided direct evidence of L1 influence in accordance with our research questions.

The research finding might be accounted for as follows: with a closer look at the transfer patterns of both groups, it is clear that the two groups differ most in the relationship between their respective L1 and L2. For Japanese learners, there is a partial overlap between L1 and L2 with regard to the interaction between imperfective markers

and lexical aspects while a much looser combination is allowed in L1 but not in L2; while for Chinese learners, tense inflection is completely absent in L1. White (1991) discussed the situations where a partial fit between L1 and L2 structure might pose problems to L2 learners in her research on argument structure. According to her, in the former case more efforts would be required for the learners to restrict the L1 to the L2 because the L1 permits properties not allowed in the L2. Thus it is reasonable to assume more acquisition difficulty and higher possibility of L1 influence. This reasoning has been generally supported in L2 argument structure studies (Juffs, 1996; Inagaki, 2001). Though dealing with different structures, this argument can still shed some insight on the present study. Since the L1 Japanese permits a much looser combination between the imperfective aspect marker and achievement verbs while the L2 English only permits a much restricted one, more L1-oriented difficulty might occur because it is clearly more challenging for learners to unlearn the L1-specific feature through not only positive evidence but also negative evidence in L2 input (White, 1991; Inagaki, 2002).

However, as found in the previous analysis, L1 influence was not found in Chinese learners' performance on base form errors. So where did Chinese learners' sensitivity to base form errors derive from? As to this question, some interesting clues were provided by the study of Tokowicz and MacWhinny (2005) on implicit and explicit measures of sensitivity to grammatical violations in second language acquisition. Using event-related brain potentials (ERPs), they investigated the contributions of explicit and implicit processes during second language (L2) sentence comprehension. As shown in different brain responses to grammatical and ungrammatical sentences, a strong grammaticality effect was found in the ERP data for the construction that was unique to the L2, suggesting that the learners were highly sensitive to these violations. This finding offered

a robust cognitive support to the present study. Under the past tense context, base form is the most direct grammatical violation to past tense which is absent in the L1 but unique to the L2. Similar to the finding in Tokowicz and MacWhinny's study, Chinese learners also exhibited a much higher sensitivity to this type of error.

Another question still deserves further consideration. Even though Chinese learners were equipped with explicit knowledge about past tense marking, generally speaking they could only be exposed to positive input, that is, grammatical past tense forms instead of ungrammatical base forms. How could we explain their success in disallowing these ungrammatical forms? One of the possible explanations was provided by White (1991) who argues that since the L2 properties not allowed in the L1 exist in the input, L2 learners may be able to notice them and arrive at the L2 grammar on the basis of positive evidence. Therefore, their task to reject base forms was made easier because only positive evidence in L2 input is required for Chinese learners to become aware of the tense marking of L2 which is not allowed in the L1. In other words, with the absence of tense marking in the L1, there is no competing similar grammatical form in Chinese learners' interlanguage which might cause some learning difficulty. There is reason to believe that this might have contributed to Chinese learners' better performance on base forms.

In summary, Study One provided answers to research question 1 and 2. First, with regard to L1 influence, it was assumed that partial overlap between the L1 Japanese progressive aspect and that of the L2 English may trigger L1 interference evidenced in higher acceptance of incorrect progressive marking on achievement and state verbs on the side of Japanese learners; on the other hand, no hypothesized L1 influence was found on Chinese learners. Besides, differing from the Aspect Hypothesis, activity verbs did not display a stronger association with progressive marking, which might be attributed to the

influence of lexical aspect and task effect.

Study Two

Study Two serves as a replication of Study One, while the focus is on the introspective interview conducted after the error identification and correction test was taken. Although Study One partially confirmed the influence of lexical aspect and L1 on learners' temporal marking, two questions still remain: Are learners aware of verbal semantics? Is there any cognitive evidence of L1 influence found in learners' verbal reports, particularly with regard to Japanese learners' performance on ungrammatical progressive states and achievements? Study Two tries to provide answers to these questions.

Subjects

The participants were 12 Japanese and 12 Chinese college students. Chinese learners were included as a comparison with Japanese learners in order to highlight any possible L1 influence in terms of progressive marking and base form. Participants in Study two were from the same universities as those in Study One, and they were all in their 1st or 2nd year. A language proficiency test based on the pre1st and 2nd Grade English Language Proficiency Test of Japan was administered to the participants to ensure that both groups were of the same proficiency level. The result of the language proficiency test showed no difference between the two groups ($F(1,14)=0.036, p>.05$).

Procedure

The same error recognition and correction test was used, since metacognitive awareness is often measured through learners' grammaticality judgments; particularly those which

require error correction and justification because they require learners to access and elaborate upon their linguistic knowledge, which is a reflection of metalinguistic awareness (Bialystok & Ryan, 1985).

The test was given to individual participants who were told to finish it in 30 minutes. Right after the paper test, a half-structured retrospective interview was conducted. The majority of the interview questions covered both correct and incorrect learner responses. For example, “Why do you think this item is wrong?”, “Would you please describe your thinking process to me?”, and “If this item is correct, could you explain its meaning to me?” The learner could view both the test and his or her responses while trying to reflect on his or her own performance. Besides these basic questions, the researcher also asked some probing questions in order to clarify and confirm what the learner reported. After Japanese learners were tested and interviewed, the researcher flew to China where the data were obtained following the same procedure. The interviews were conducted in Japanese with the Japanese group, while the Chinese group was interviewed in Chinese. Both interviews were recorded with the permission of the participants. The interviews were later transcribed and analyzed in terms of 1) verbal reports related to awareness of lexical aspect; 2) verbal reports related to L1 influence; 3) verbal reports related to types of interpretations of targeted temporal markings.

Results and Discussion

A 2 (L1) × 3 (lexical aspect) ANOVA was conducted on learners’ performance on progressive errors and base form errors respectively with the test score as the dependent variable, and L1 (2 levels) & lexical aspect (3 levels) as the independent variables. The results in Study Two largely echoed those in Study One. As for the progressive aspect,

achievement and state verbs were more challenging for Japanese students (on states: $F(1,66) = 16.00, p < .001$; on achievements: $F(1,66) = 5.40, p < .05$). Regarding the base form, Chinese learners outperformed their Japanese counterparts, showing a stronger awareness of it ($F(1,22) = 9.18, p < .01$). Some basic statistics are provided in Table 7 and 8. Therefore, it seems that Japanese learners have more difficulty than Chinese learners in recognizing ungrammatical interactions between achievement and state verbs and progressive marking, which indicates a possible L1 influence. By contrast, Chinese learners successfully rejected more base forms which are default verb forms in their L1.

The following part will make a discussion of learners' verbal reports. Since it was found that Japanese learners exhibited more difficulty in this task, the following analysis will be mainly devoted to Japanese learners' verbal reports relating to verbal semantics and L1 influence in accordance with the research questions.

Table 7 Basic Statistics of the Performance on Both Error Categories by the Two Groups

L1	Error 1		Error 2	
	Mean	SD	Mean	SD
China N=12	1.28	0.38	1.63	0.24
Japan N=12	0.98	0.35	1.42	0.32

Note: error1=progressive error; error2=base form error

Table 8 Basic Statistics of the Performance on Progressive Errors by Japanese and Chinese Learners across Three Lexical Aspects

	states	achievements	activities
Mean	1.33	1.54	0.97
China SD	0.41	0.33	0.41
Japan Mean	0.68	1.16	1.12
SD	0.45	0.32	0.29

Learners' Awareness of Verbal Semantics

As the learners' verbal reports showed, they displayed some awareness of verbal lexical aspect and some were even found to possess a certain degree of explicit knowledge of lexical aspect. As observed in their verbal reports, some learners were capable of interpreting temporal morphology by referring to the type of lexical aspect. The following table shows some typical samples of learners' awareness of verbal semantics. In the following tables, Japanese and Chinese learners were abbreviated as JL and CL, respectively.

Table 9 Sample Verbal Reports on Verbal Semantics of State and Achievement

Verbs	
Test Items	Learners' Reports (Translated)
They met a Spanish sailor standing at the top of the stairs. He <u>had</u> no time to shout to the others. Tom took the man's knife and <u>kill</u> him. Next, they were breaking open a door. Inside the room, they <u>found</u> a large number of knives.	I think progressive aspect should refer to some action which can last for a while. But 'breaking open', this action is so fast, and can take place just in a blink. So I think here -ing is strange. (JL3)
The young man <u>lie</u> in darkness in the bottom of the ship. The ship <u>was moving</u> up and down in the winds. Above him he was hearing loud voices—the men who <u>were trying</u> to sail the ship through the seas.	Well, it is just my guess. I feel that here 'heard' is more appropriate than 'hearing'. I feel that hear, this verb is instant. But 'hearing', the progressive use requires a certain time span. Well, this is my guess.(CL5)

<p>Mom <u>told</u> me that in the past the stores over there <u>were usually opened</u> till 10 p.m. On her way home from work she often <u>went</u> shopping there. She thought the food there <u>was tasting</u> wonderful.</p>	<p>This verb is the same as what I met just now such as want and know. This verb, ‘taste’, as I remember, is a state verb. So it can not go with –ing. (JL11)</p>
<p>Next morning I <u>talked</u> with my secretary when Holmes came into my room...</p>	<p>I think ‘talked’ is strange here because there is ‘when’ in the sentence. Besides, when you talk with someone, you often talk for some time. So I think... -ing is better. (CL7)</p>

Although learners tend to be affected by the lexical aspect in temporal processing, it is far from a major strategy in their interpretation repertoire. As shown in this study, learners mostly refer to the local context such as time adverbs, tense consistency, translation as well as rules of thumb to facilitate their comprehension. Taking all these factors into account, it could be rightly argued that lexical aspect is not the decisive factor in learners’ temporal marking. What’s more, it is worth noting that learners do not possess a systematic awareness of verbal semantics in their interlanguage and they tend to rely on simplified grammar rules as their main resource. This kind of item learning, in contrast with holistic learning, might be able to explain the learners’ unstable performance with tense-aspect tasks.

As we know, learners’ verbal reports reveal the interpretation/comprehension process when they deal with temporal expression. Thus, it is interesting to note that the devices they used for temporal comprehension largely overlapped with those for temporal production. For example, Dittmar (1981) mentioned that learners tended to use calendar expressions or adverbs such as “yesterday” with an infinitive to mark past or future time. Sato (1984) also indicates that second language learners rely on context and implicit inferences to mark temporality. According to her, devices used by L2 learners to express

temporality include temporal adverbials, locative adverbials, clause sequencing, calendar expression, interlocutor scaffolding, and implicit references. This overlapping is akin to the interactive nature of learners' comprehension and production. Some samples that relate to the above discussion are given in Table 10. In Table 11, the raw number of learners' verbal reports on major interpretation types across both groups is provided, which shows the frequency of a certain interpretation type appearing in the reports.

Table 10 Samples of Learners' Reports on their Temporal Interpretation

(translated)

Contextual meaning	From the previous part, I know that it is about the past time. Then I translated it into Japanese. And I think 'he was hearing a loud voice' is very strange. Yes, I judged it according to the context. (CL9)
Time adverb	Well, since there is the adverb 'while', I think it is better to use past progressive. (JL6)
Tense consistency	Because the latter parts were all in past tense, it must be wrong for this part to be written in past tense. (JL2)
Grammar rules	We learned this rule in our high school days, that is, progressive marking can not be used on the verbs such as want, know, etc.
Translation	I translated this part '...was looking handsome...' into Japanese, and I found no problem with it. (JL4)
Guess	I also felt confused about this part. So I just made a guess and chose it for my luck. (CL10)
Frequency effect	Well, because I have never seen this kind of usage before. (JL7)
Classroom training effect (by Chinese learners only)	This is very easy to me. When I saw the word 'write', I knew it must be wrong. Any word without s or ed is strongly suspicious. We did a lot of exercises on it in class. So I didn't give it any thinking when I wrote the answer. (CL2)
Verbal semantics	I think if you break something, it can be finished in an instant. (JL5)

Table 11 Number of Verbal Reports on Major Interpretation Types by both Groups across Two Error Categories

	Japan		China	
	Prog error	Base error	Prog error	Base error
Contextual meaning	45	15	37	4
Time adverb	29	19	31	3
Time consistency	26	28	20	6
Grammar rules	15	1	32	15
Translation	43	0	27	0
Guess	24	4	5	0
Frequency effect	12	0	19	0
Classroom training effect	0	0	0	25
Verbal semantics	8	0	9	0

In addition, learners' verbal reports also show that their sensitivity to the lexical aspect varies with verb types. It is clear that learners gave more reports on achievement verbs such as break and drop, while very few were made on state and activity verbs, especially on activity verbs. One of the reasons for this, in the researcher's assumption, is that because the task is given in a past context, ungrammatical progressive achievements are easier to identify due to the sharp difference between the action-ing-progressive marker and punctual/completed verbal semantics. In contrast, a progressive marker on state and activity verbs will not be as noticeable because of the proximity between the verbal semantics and aspect marker. This is especially true for activity verbs. On the other hand, there is also a possibility that different verbal classes carry different degrees of saliency for EFL learners. From our results, it seems that activities are far less salient than

achievements (See Table 12). Is the semantic saliency cognitively universal or is it bound within contextual factors? This seems to be an interesting new finding from the interview task and one which deserves further investigation in future studies.

Table 12 Awareness Frequency of Verbal Semantics across Three Verb Types

		Verb Types		
		States	Activities	Achievements
Verbal Semantics Mentioned	Verbs	2/8	1/8	4/8
	Learners	5/24	1/24	9/24

Note: Verbs: number of verbs with semantics mentioned/total number of verbs

Learners: number of learners mentioning semantics/total number of learners

L1 Influence

In this error correction task, Chinese learners were found to be more sensitive to the base form errors, while Japanese learners accepted more ungrammatical progressive marking on state and achievement verbs (especially state verbs). Since Japanese learners' higher acceptance of progressive states and achievements presents evidence for possible L1 influence, it is interesting to observe the distribution of temporal markings in the progressive error category in the grammar test within each national group. Table 13 provides the distribution of temporal markings in the progressive error category on the three lexical aspects across both national groups. And the raw number of verbs which carried the morpheme in question was given for each learner.

Table 13 Distribution of temporal markings in progressive error category on three lexical aspects across two groups

Learners	states	achievements	activities
	a b c d e f	a b c d e f	a b c d e f

1 C	3 1 0 0 0 0	1 1 0 2 0 0	3 0 0 0 1 0
J	1 2 0 1 0 0	0 3 0 1 0 0	2 2 0 0 0 0
2 C	1 1 1 1 0 0	4 0 0 0 0 0	2 2 0 0 0 0
J	0 4 0 0 0 0	0 3 0 0 0 1	4 0 0 0 0 0
3 C	1 1 0 2 0 0	4 0 0 0 0 0	4 0 0 0 0 0
J	1 1 1 0 0 1	2 2 0 0 0 0	1 1 1 0 0 1
4 C	1 1 2 0 0 0	3 0 0 1 0 0	2 2 0 0 0 0
J	0 3 1 0 0 0	0 4 0 0 0 0	1 1 2 0 0 0
5 C	3 1 0 0 0 0	3 0 0 1 0 0	0 4 0 0 0 0
J	0 3 0 1 0 0	3 1 0 0 0 0	2 2 0 0 0 0
6 C	2 0 1 0 1 0	4 0 0 0 0 0	3 1 0 0 0 0
J	0 4 0 0 0 0	2 2 0 0 0 0	1 3 0 0 0 0
7 C	1 1 1 1 0 0	1 1 0 2 0 0	0 4 0 0 0 0
J	2 2 0 0 0 0	2 0 0 1 0 1	4 0 0 0 0 0
8 C	4 0 0 0 0 0	3 0 0 0 0 1	3 0 0 0 1 0
J	0 4 0 0 0 0	2 2 0 0 0 0	0 4 0 0 0 0
9 C	2 2 0 0 0 0	3 0 0 0 1 0	2 0 1 0 0 1
J	3 1 0 0 0 0	4 0 0 0 0 0	1 1 2 0 0 0
10 C	4 0 0 0 0 0	4 0 0 0 0 0	2 2 0 0 0 0
J	0 4 0 0 0 0	1 1 1 0 0 1	1 3 0 0 0 0
11 C	4 0 0 0 0 0	3 1 0 0 0 0	0 2 1 0 1 0
J	3 1 0 0 0 0	4 0 0 0 0 0	0 4 0 0 0 0
12 C	1 1 1 1 0 0	4 0 0 0 0 0	0 2 2 0 0 0
J	1 1 1 0 0 1	4 0 0 0 0 0	0 2 1 0 0 1
total C	27 9 6 5 1 0	37 3 0 6 1 1	21 19 4 0 3 1
J	11 30 3 2 0 2	24 18 1 2 0 3	17 23 6 0 0 2

Note: C: Chinese learners; J: Japanese learners;

a: past tense marking; b: progressive marking; c: present marking; d: perfect marking; e: future marking; f: base form

Our interest is to observe the distribution of non-targetlike forms in both groups. In terms of Japanese learners, progressive markings on states and achievements occupy the major proportion of all the non-targetlike forms followed by present marking, base forms and perfect marking. In contrast, this was not found in the Chinese group, with present forms being the leading non-targetlike forms. Therefore, the distribution of temporal markings across the two national groups echoes the findings on L1 influence in the previous discussion. Next, discussion will be made with some more evidence from the

learners' verbal reports.

First, since Japanese learners were found to accept more ungrammatical progressive markings on achievement and state verbs, we were interested in how Japanese learners interpreted these progressive markings. In the interviews, students were asked to make verbal reports on both recognized and unrecognized error items. It was found that Japanese learners were able to recognize some error items by noticing the mismatch between progressive marking and achievement & state verbs, showing a target-like understanding of L2 progressive aspect. Sample reports are given in Table 14.

However, Japanese learners' understanding of English progressive aspect was not consistent, which can be seen from their weaker performance on their recognition of ungrammatical progressive achievement and stative verbs. In the interview, learners were asked to explain the meaning of the sentence containing the target item if that error item was not identified. And the learners usually produced a translation which they took as natural and semantically correct. In contrast, Chinese EFL learners could identify the ungrammaticality of state and achievement progressives more easily with the help of translation. For example, one learner just told the researcher that, 'It sounds very strange if I translate it into Chinese. So I think it is wrong.' Therefore, with regard to Japanese learners, the translations provide some clues to the semantic motivation for their L2 progressive interpretation. In the verbal reports, two patterns were representative of Japanese learners' translations of erroneous progressive markings on achievements and states. One of them is that learners tended to render English *-ing* into Japanese *-te i*, showing that they took *te i* as the equivalent of *-ing*. Meanwhile, learners also provided translations without resorting to *te i* structure, seemingly transferring the L1 resultative state meaning onto the L2 states and achievements progressives. Some examples are

given in Table 15.

Table 14 Sample Reports on Ungrammatical Progressive Markings on States & Achievements

Test Items	Learners' Report
I <u>was walking</u> with my friend on the street, and we <u>saw</u> some money lying on the ground. But nobody was around. Then my friend <u>said</u> to me, "somebody <u>was dropping</u> money on the street."	Well, I feel it is very strange because if you say dropping it means that the action is continuous and the thing you drop is still on the way. (JL12)
Something <u>went</u> wrong with my friend's computer. After he <u>checked</u> it he <u>said</u> to me in a loud voice, ' <u>are you noticing</u> the problem?'	I think it is strange to say that someone continues the action of noticing something. When you notice something, you just notice it. (JL6)

Table 15 Sample Translations on Semantic Interpretation of Progressive States & Achievements by Japanese Learners

Test items	Translations
...Tom turned to the others. 'The two men who think that we <u>are belonging</u> to them,' he said, 'find them'.	Kore wa watashi tachi wa karera ni <u>shozoku shiteiru</u> . (<u>shozoku shiteiru</u> = belonging) Note: here 'beloinging' was translated into the corresponding Japanese achievement verb plus teiru which indicates the resultative state. (JL1)
After he checked it he said to me in a loud voice, ' <u>are you noticing</u> the problem?'	Nani ka mondayi ni <u>kizuyitemasuka</u> ? Note: here the Japanese imperfective marker <i>teiru</i> was used to go with the English couterpart <i>be noticing</i> . The Japanese translation sounds perfectly natural although it denotes the meaning 'have you noticed the problem'.(JL8)

The evidence of L1 influence is not straightforward, but the learners' reflective reports did help us gain more insights into this complex issue. With both statistical data and interview data from Study one and two, a more convincing argument could be made on L1 influence in EFL temporal acquisition.

Learners' verbal reports also reminded us of another issue in SLA studies: optionality.

“Optionality” is a characteristic of developing L2 grammars, as discussed in White (1991). From the verbal reports, it can be seen that Japanese learners were entertaining two concepts of the lexical aspect-progressive interaction: on the one hand, they followed the basic target-like concept and could interpret the L2 progressive as action-in-progress; on the other hand, they also held onto their L1 concept of interpreting the states & achievements as resultative state meaning. Although some of their translations only indicated a possible resultative meaning due to the semantic complexity of *te i*, some other examples offer us much stronger evidence of the L1 semantic influence. For example, since in Japanese ‘shozoku shiteiru’ can only indicate a resultative state meaning, when Japanese learners interpreted ‘...we are belonging to them...’ as ‘*kore wa, watashitachi wa karera ni shozoku shiteiru*’, it is reasonable to assume that they transferred the L1 semantics into the L2. Among the studies on L2 tense and aspect acquisition, very few have been done on L1 influence on Japanese EFL learners’ progressive interpretation. Within the researcher’s knowledge, only Gabriele (2005) investigated this issue in her study on the role of transfer in the L2 acquisition of aspect. With a focus on the acquisition of the semantics of the English progressive marker, she employed an interpretation task to examine whether Japanese learners have more difficulty preempting the resultative state interpretation that is not available in the L2, but is an option in the L1. According to the results, learners showed a great deal of difficulty with past progressive on achievements in the interpretation task. Therefore, it was argued that Japanese learners did hold on to their L1 aspectual semantics of *te i* and tend to transfer it into the L2 semantic interpretation. Gabriele’s research provides support to the results of the present research although more efforts are needed to explore the complicated issue of L2 aspectual semantic interpretation.

As to the base form errors, which are not as semantically complex, it is no surprise that learners made much less report. What is interesting to the researcher is the different ways Japanese and Chinese learners describe this error item. In the interview, Japanese learners always called the base error ‘ganzaikei’, meaning present tense, and it seems that they mainly employed grammatical knowledge for the interpretation. But Chinese learners always referred to it as ‘yuanxing’, meaning uninflected verb form. Moreover, Chinese learners mostly referred to the classroom training which made this type of error detection easy to them. It seems that Chinese learners used the base form itself as a reminder in their interpretation, especially when they might be able to notice the pattern of the test along with their verbal reports.

To sum up, Study two supported the statistical results in Study one which partially answered research question 1 & 2. Moreover, it provided qualitative evidence in answer to research question 3, confirming the existence of L1 influence and learners’ awareness of lexical aspect in tense-aspect performance.

Pedagogical Implications

In terms of tense-aspect pedagogy, the most important task is to help learners construct a target-like form-form and form-function mapping. As the present studies showed, learner’s temporal acquisition is subject to a number of factors such as influence of lexical aspect and L1. On the basis of the current findings, the present studies highlighted the important role of input in tense-aspect acquisition as well as classroom instruction.

First, as discussed before, although progressive-activities association predicted by the Aspect Hypothesis was not found in the present studies, it could still be argued that learners’ tense-aspect interlanguage is not fully grammaticalized and their temporal

marking tend to be affected by verbal semantics. According to the distribution bias hypothesis, properties of the input promote the incorporation of an inappropriate form-meaning relationship in the interlanguage and learners may misperceive the meaning and distribution of a particular form that they discover in the input (Andersen, 1990). And in the researcher's assumption, insufficient input might also contribute to learners' difficulty in temporal acquisition as observed in learners' failure to add target-like progressive markings to activities in the present study. Therefore, richer and less biased language input is very important for learners to acquire non-prototypical temporal markings.

A number of instruction methods have been proposed as a way to help learners overcome the misleading effect of lexical aspect. One of them is so-called input flood, that is, providing students with positive evidence of the target structure through meaning-focused activities. In her study, Bardovi-Harlig (1995) reports on the effect of an ESL instructional unit in which she provided an input flood of contextualized activity verbs in the past perfective to EFL learners. The results show that the input flood technique did benefit learners in their temporal markings. However, as Sharwood-Smith (1981) argued, it might not lead to acquisition if learners just process the input for meaning. Therefore, in our teaching practice, input can be enhanced to make it more salient to the learners by using such techniques as underlining, italicizing, bolding, coloring, etc. But as J. White (1991) pointed out, even though input enhancement could be a valuable technique to draw learners' attention to the target form in the input, it may not be effective when the target structure involves L1-L2 contrasts. Thus more negative evidence is needed to make the target structure more salient in the input.

The importance of negative evidence was also noted in the present studies. When confronted with unmarked activity verbs in progressive context, learners tended to ignore

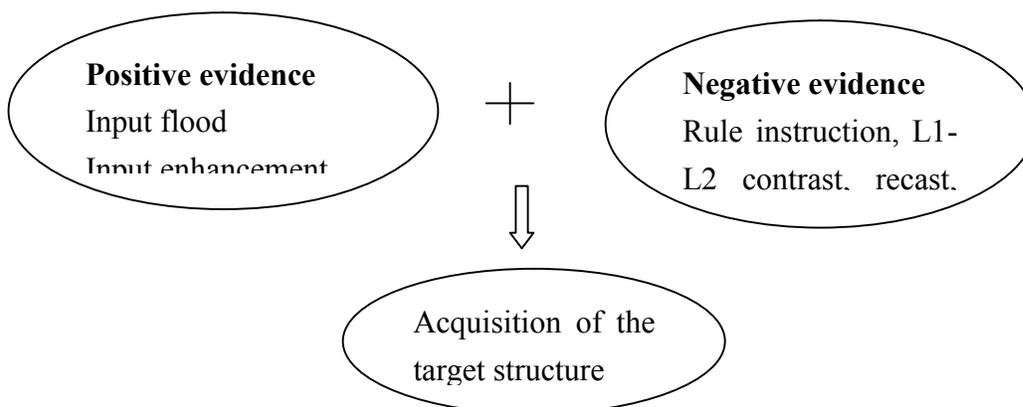
the grammatical marker due to heavy overlapping between verb semantics of activities and progressive aspect. And this kind of ungrammatical usage implying what is not possible cannot be found in positive evidence. Therefore negative evidence is highly necessary for a more target-like temporal marking.

The necessity of negative evidence is further accentuated by the finding of L1 influence in the present studies. As White (1990, 1991) argued, learners may require negative evidence when their interlanguage contains rules more general than those of the TL because negative evidence allows the L2 learners to know what the TL disallows. Her argument found new support in the present studies. When dealing with ungrammatical progressive markings on states and achievements, Japanese participants gave a much weaker performance, which indicated a possible L1 influence, that is, they have to learn to restrict their L1 concept and know what is not possible in L2.

While many teachers are eager to embrace the current pedagogical view that communicative-based, meaning-driven instruction is the most beneficial to L2 learners, they are also frustrated to see the unsatisfactory level of accuracy in learners. The reason seems quite evident: learners are often left unaware of their errors since their non-target-like production tends to be overlooked by teachers as long as there is no communication breakdown. In order to solve this problem, the present researcher hopes to remind language instructors of the importance of negative evidence, especially explicit negative evidence which often appears in the forms of metalinguistic rule instruction, evaluative judgements and L1-L2 contrast with overt reference to a target form. Take the progressive marking for example. According to Tomioka (1994), many Japanese textbooks just briefly explain the use of “be V-ing” under a section of “progressive aspect” and simply provide a list of verbs which do not take the progressive aspect. When

teachers teach progressive aspect in this way, students are easily misled to have the impression that English “*V-ing*” is the equivalent of Japanese “*V-te i*” and that there are only a limited number of exceptions. Therefore, when teachers provide little explicit rule instruction and negative evidence, progressive marking would be more challenging for Japanese learners who tend to turn to their L1 concept because this kind of instruction can not guarantee a well internalized L2 concept. In this case, it is highly necessary for teachers to provide explicit negative evidence by pointing out the source of error and offering further explanation in order to raise learners’ awareness of it.

In summary, based on the current findings and above discussions, the researcher attempts to suggest the following approach to the practical tense-aspect instruction:



Conclusion

In summary, the present studies provide answers to the three research questions. First of all, with statistical results from Study One and Two, L1 influence was found to be an active factor in Japanese learners’ progressive marking performance. And it was assumed that partial overlap between the L1 Japanese progressive aspect and that of the L2 English may trigger L1 interference evidenced in higher acceptance of incorrect

progressive marking on achievement and state verbs on the side of Japanese learners. In contrast, the absence of tense markers in L1 Chinese seems to have facilitated Chinese learners' detection of base form errors in the present studies, and no hypothesized L1 influence was found on Chinese learners. Besides, differing from the Aspect Hypothesis, a strong progressive-activities association was not found in the results, which might be attributed to the heavy semantic overlapping between activities and progressive aspect. In addition to the statistical results, Study Two provides qualitative evidence for the existence of L1 influence and learners' awareness of lexical aspect in tense-aspect performance. Moreover, findings in the present studies highlight the importance of negative evidence in classroom teaching, especially with regard to the challenge Japanese learners face in learning to disallow what is not possible in TL.

The present studies are limited in the following aspects: first, although with a mixed research design more knowledge has been obtained in terms of learners' interpretation of temporal marking, there is still much room for more in-depth investigation. In the future research, longitudinal data will be included. Secondly, caution should be taken in interpreting the results from Study Two. Due to the small sample size, it can only be viewed as a tentative attempt to tap into learners' thinking process in their temporal performance. What is more, as Ellis (2004) argued, there is always a distinction between possessing the knowledge itself and the ability to verbalize it, regardless of whether the learner possesses the metalanguage. In the future, a larger sample will be used for more qualitative evidence.

References

Andersen, R.W. (1990). Models, processes, principles and strategies: Second language

- acquisition inside and outside the classroom. In B. BanPatten & J. F. Lee (Eds.), *Second language acquisition – Foreign language learning* (pp. 45-78). Clevedon, UK: Multilingual Matters.
- Andersen, R.W. & Y. Shirai. (1996). Primacy of aspect in first and second language acquisition: The pidgin/Creole connection. In W.C. Ritchie & T.K. Bhatia (Eds.) *Handbook of second language acquisition* (pp. 527-570). San Diego, CA: Academic Press.
- Bardovi-Harlig, K. (1995). The interaction of pedagogy and natural sequences in the acquisition of tense and aspect. F. R. Eckman, D. Highland, P. W. Lee, J. Mileham and R. Rutkowski Weber (Eds.). *Second Language Acquisition Theory and Pedagogy*. Mahwah, New Jersey: Lawrence Erlbaum Associates.
- Bardovi-Harlig, K. (2000). *Tense and aspect in second language acquisition: Form, meaning and use*. Oxford: Blackwell.
- Bardovi-Harlig, K. & Reynolds, D. (1995). The role of lexical aspect in the acquisition of tense and aspect. *TESOL Quarterly* 29, 107-131.
- Bialystok, E., & Ryan, E. B. (1985). Towards a definition of metalinguistic skill. *Merrill-Palmer Quarterly*, 31, 229–251.
- Cai, J. T. (2003). *The Effects of Multiple Linguistic Factors on the Simple Past Use in English Interlanguage*. Beijing: Foreign Language Teaching and Research Press.
- Chalker, S. (1994). Pedagogical grammar: Principles and problems. In M. Bygate, (Ed), *Grammar and the language teacher* (pp.31-45). London: Prentice Hall.
- Collins, L. (2002). The roles of L1 influence and lexical aspect in the acquisition of temporal morphology. *Language Learning* 52, 43-94.
- Collins, L. (2005). Assessing second language learners' understanding of temporal

- morphology. *Language Awareness*, 14, 207-214.
- Comrie, B. (1976). *Aspect*. Cambridge: Cambridge University Press.
- Dittmar, N. (1981). On the verbal organization of L2 tense marking in an elicited translation task by Spanish immigrants in Germany. *Studies in Second Language Acquisition* 3, 136-164.
- Dulay, H. & M. Burt. (1974). Natural sequences in child second language acquisition. *Language Learning* 24, 37-53.
- Ellis, R. (2004). The definition and measurement of L2 explicit knowledge. *Language Learning* 54, 227-75.
- Gabriele, A., & Martohardjono, G. (2005). Investigating the role of transfer in the L2 acquisition of aspect. *Proceedings of Generative Approaches to Second Language Acquisition* 7. Somerville: Cascadilla Press Web Proceedings Project.
- Housen, A. (2002). The development of tense-aspect in English as a second language and the variable influence of inherent aspect. In M. Salaberry & Y. Shirai (Eds.), *Tense-aspect morphology in L2 acquisition*. Amsterdam: John Benjamins.
- Inagaki, S. (2001) Motion verbs with goal PPs in L2 acquisition of English and Japanese. *Studies in Second Language Acquisition*, 23, 153-70.
- Inagaki, S. (2002). Japanese learners' acquisition of English manner of motion verbs with locational/directional PPs. *Second Language Research*, 18, 3-27.
- Juffs, A. (1996). *Learnability and the lexicon: theories and second language acquisition research*. Amsterdam: John Benjamins.
- Klein, W. (1994). *Time in language*. London: Routledge.
- Liskin-Gasparro, J. (2000) The use of tense-aspect morphology in Spanish oral narratives: Exploring the perceptions of advanced learners. *Hispania* 83, 830-44.

- Robison, R. (1990). The primacy of aspect: Aspectual marking in English interlanguage. *Studies in Second Language Acquisition* 12, 315-330.
- Robison, R. (1995). The aspect hypothesis revisited: a cross-sectional study of tense and aspect marking in interlanguage. *Applied Linguistics*. 16, 344-370.
- Rohde, A. (1996). The aspect hypothesis and the emergence of tense distinctions in naturalistic L2 acquisition. *Linguistics* 34, 1115-1137.
- Salaberry, M. R. (1998). The development of aspectual distinctions in academic L2 French. *Canadian Modern Language Review* 54, 508-542.
- Salaberry, M. R. (1999). The acquisition of English past tense in an instructional setting. *System* 28, 135-152.
- Salaberry, R. & Y. Shirai (Eds.). (2002). *Tense-aspect morphology in L2 acquisition*. Amsterdam: John Benjamins.
- Sato, C. (1984). The syntax of conversation in interlanguage development. Unpublished doctoral dissertation, University of California, Los Angeles.
- Sharwood-Smith, M. (1981). Consciousness-raising, and the second language learner. *Applied Linguistics*, 2, 159-69.
- Shirai, Y. (1998). Where the progressive and the resultative meet: Imperfective aspect in Japanese, Korean, Chinese and English. *Studies in Language* 22, 661-692.
- Shirai, Y. (2000). The semantics of the Japanese imperfective –teiru: An integrative approach. *Journal of Pragmatics* 32, 327-61.
- Shirai, Y. & Andersen, R.W. (1995). The acquisition of tense-aspect morphology: A prototype account. *Language* 71, 743-62.
- Shirai, Y. & Kurono, A. (1998). The acquisition of tense-aspect marking in Japanese as a second language. *Language Learning* 48, 245-279.

- Skehan, P. (1998). *A cognitive approach to language learning*. Oxford: Oxford University Press.
- Slabakova, R. (2002). Recent research on the acquisition of aspect: an embarrassment of riches? *Second Language Research*, 18, 172-188.
- Smith, C. S. (1983). A theory of aspectual choice. *Language*, 59, 479–501.
- Thomason, S. G., & Kaufman, T. (1988). *Language contact, creolization, and genetic linguistics*. Berkeley and Los Angeles: University of California Press.
- Smith, C. S. (1991). *The parameter of aspect*. Dordrecht: Kluwer.
- Sun, xiaozhao. (2006). A Comparative Study on L1 Influence and Lexical Aspect in Past Tense and Progressive Marking by Japanese and Chinese EFL Learners. *International Journal of Curriculum Development and Practice* 8, 13-26.
- Teddle, C., & Tashakkori, A. (2003). Major issues and controversies in the use of mixed methods in the social and behavioral sciences. in A.Tashakkori & C. Teddle (Eds.), *Handbook of mixed methods in social & behavioral sciences* (pp.3-50). Thousand Oaks: Sage.
- Tokowicz, N. & MacWhinny, B. (2005). Implicit and explicit measures of sensitivity to violations in second language grammar. *Studies in second language acquisition* 17, 173-204.
- Tomioka, Taeko. (1994). A study of ‘V-te iru’ in Japanese. Online Resource. Retrieved November 21, 2006 from <http://homepage3.nifty.com/park/aspect.htm>.
- Vendler, Z. (1967). Verbs and times. In A. Vendler (ED.), *Linguistics and philosophy* (pp.97-121). New York: Cornell University Press.
- Von Stutterheim, C. (1991). Narrative and description: Temporal reference in second language acquisition. In T. Huebner & C. A. Ferguson (Eds.), *Crosscurrents in*

second language acquisition and linguistic theories (pp. 385–403). Philadelphia:

Benjamins.

White, L. (1990/91). The verb-movement parameter in second language acquisition.

Language Acquisition, 1, 337-60.

White, L. (1991). Argument structure in second language acquisition. *French Language*

Studies 1, 189-207.

Yang, Suying & Huang, Yueyuan. (2004). The impact of the absence of grammatical

tense in L1 on the acquisition of the tense-aspect system in L2. *International*

Review of Applied Linguistics in Language Teaching, 42, 49-79.

Appendix 1

Table 1 List of all the Target Verbs across Two Error Categories

	State Verbs		Activity Verbs		Achievement Verbs	
	Error	Correct	Error	Correct	Error	Correct
Base Form	feel →	felt	write →	wrote	kill →	
	appear →	appeared	lie →	lay	killed	
	belong	→	sail →	sailed	miss →	
	belonged		live →	lived	missed	
	hate →	hated			find →	
					found	
					jump →	
					jumped	
Progressive Form	was looking →	looked	looked at →	was	was hearing →	
	was wanting →	wanted	looking at		heard	
	were belonging →		played →	were	was dropping →	
	belonged		playing		dropped	
	was tasting →	tasted	talked →	was talking	were you noticing →	
			worked →	was	did you notice/have	
			working		you noticed	
					were breaking →	
					broke	

Appendix 2 Sample items of the test

1. Tom told us to come inside. There we met Horace Harker, the man who was A
B
looking quite handsome. He worked for a newspaper, and today he write a good
C D
story. He was very happy for it.
2. The young man lie in darkness in the bottom of the ship. The ship was moving up
A B
and down in the winds. Above him he was hearing loud voices---the men who were
C D
trying to sail the ship through the seas.
3. They met a Spanish sailor standing at the top of some stairs. He had no time to
A
shout to the others. Tom took the man's knife and kill him. Next, they were breaking
B C
open a door. Inside the room, they found a large number of knives.
D

The Effect of Metacognitive Strategy Instruction on EFL Learners' Reading Comprehension Performance and Metacognitive Awareness

Fatemeh Takallou
Payame Noor University, Iran

Bio Data:

Fatemeh Takallou has an M.A. in TEFL (Teaching English as a Foreign Language) from the Teacher Training University of Tehran in Iran. She is currently lecturing and researching in Payame Noor University, Kermanshah, Iran where she has taught courses in EFL, ESP, and translation to graduate and undergraduate students.

Abstract

As learners have an important role in new teaching methodologies, raising their awareness of learning strategies and helping them utilize these strategies is a crucial aim of teachers. One type of these learning strategies is metacognitive strategies including planning, self-monitoring and self-evaluation. The present study aimed at examining the effect of metacognitive (planning & self-monitoring) strategy instruction on EFL learners' reading comprehension performance (on authentic and inauthentic texts) and their metacognitive awareness. To this end, two tests (TOEFL and a reading comprehension test) and Strategy Inventory for Language Learning (SILL) were administered to 93 male and female EFL learners in four phases of this study. At the first phase, TOEFL was administered to all the students both to homogenize students regarding language proficiency and to validate the reading comprehension test. At the second phase, SILL was administered to two experimental and one control groups before strategy instruction. SILL assesses the frequency with which the subjects use a variety of techniques for foreign language learning. At the third phase, two experimental groups received five sessions of instruction on metacognitive strategies, one on planning and the other on self-monitoring strategy based on the Cognitive Academic Language Learning Approach (CALLA). Both experimental and control groups worked on authentic and inauthentic

texts (some articles from *Readers' Digest* and *Reading Skillfully III*). At the fourth phase, after completion of instruction, the reading comprehension test and SILL questionnaire were administered to all groups. Data analysis revealed that two experimental groups which received instruction on 'planning' and 'self-monitoring' outperformed the control group on the reading comprehension test. Moreover, text type played an important role in the subjects' reading comprehension. The subjects performed better on authentic texts. In addition, the results showed that experimental groups' awareness to metacognitive strategies significantly increased after instruction. The findings of the present study have implications for learners, teachers, and textbook writers in the realm of TEFL in particular and education in general.

Introduction

As learners have an important role in new teaching methodologies, raising their awareness of learning strategies and helping them utilize these strategies is a crucial aim of teachers. One type of these learning strategies is metacognitive strategies including planning, self-monitoring and self-evaluation. The present study aimed at examining the effect of metacognitive (planning & self-monitoring) strategy instruction on EFL learners' reading comprehension performance (on authentic and inauthentic texts) and their metacognitive awareness. To this end, two tests (TOEFL and a reading comprehension test) and Strategy Inventory for Language Learning (SILL) were administered to 93 male and female EFL learners in four phases of this study. At the first phase, TOEFL was administered to all the students both to homogenize students regarding language proficiency and to validate the reading comprehension test. At the second phase, SILL was administered to two experimental and one control groups before strategy instruction. SILL assesses the frequency with which the subjects use a variety of techniques for foreign language learning. At the third phase, two experimental groups received five sessions of instruction on metacognitive strategies, one on planning and the other on self-

monitoring strategy based on the Cognitive Academic Language Learning Approach (CALLA). Both experimental and control groups worked on authentic and inauthentic texts (some articles from *Readers' Digest* and *Reading Skillfully III*). At the fourth phase, after completion of instruction, the reading comprehension test and SILL questionnaire were administered to all groups. Data analysis revealed that two experimental groups which received instruction on 'planning' and 'self-monitoring' outperformed the control group on the reading comprehension test. Moreover, text type played an important role in the subjects' reading comprehension. The subjects performed better on authentic texts. In addition, the results showed that experimental groups' awareness to metacognitive strategies significantly increased after instruction. The findings of the present study have implications for learners, teachers, and textbook writers in the realm of TEFL in particular and education in general.

The current research in second language reading has focused on readers' strategies. Grabe (2002) reinforced the importance of efficient reading strategies. Reading strategies are of interest for what they reveal about the way readers manage their interactions with written text, and how these strategies are related to reading comprehension. Several empirical investigations have been conducted into reading strategies and their relationships to second language reading comprehension. More recent research has begun to focus on metacognition, i.e., cognition of cognition. These studies investigate the relationships among metacognitive awareness, strategy use, and reading comprehension. There seems to be enough evidence to be confident that strategy instruction can, indeed, be effective at helping students learn more successfully (Muñiz-Swicegood, 1994; Chamot, Barnhardt, El-Dinary, & Robbins, 1996; Oxford & Leaver, 1996; Cohen, Weaver, & Li, 1998).

Interactive reading processes involve the presence of both reader and text (Celce-Murcia & Olshtain, 2000). "Psychological and psycholinguistic research strongly indicate that the quality of reading as psychological processing of the target language input depends on whether the reader finds the text personally significant" (Nilsson, 2003, p. 4), that is, if the text relates to the reader's background knowledge and experiences, interests and information need.

According to Little, Devitt & Singleton (1989, p. 6) "this can be accomplished by using carefully chosen authentic texts". Lazaraton and Skuder (1997) found that even the most recent texts fell short on authenticity criteria used. For this reason, teachers need to become critical consumers of published materials. "Bring authentic data into the classroom can assist learners ... to make communicative meanings ... nonauthentic language, in some respects, actually makes the task for the language learner more difficult" (Nunan, 1999, pp. 80-81). On the other hand, it is said that authentic materials might be too culturally biased or they might be difficult for lower level learners to decode (Nunan, 1999).

As Grabe (1991) mentions reading is probably the most important skill for second language learners in academic context but according to Van Wyk (2001), many students enter higher education underprepared for reading demands which this is often due to their low level of reading strategy knowledge and lack of metacognitive control. Research on learning strategies focus on diverse issues of the reading comprehension process (Li & Munby, 1996; Mason & Krashen, 1997; Schoonen et al., 1998; Sheorey & Mokhtari, 2001).

Furthermore, Lazaraton and Skuder (1997) found that even the most recent texts fell short on authenticity criteria used. For this reason, teachers need to become critical

consumers of published materials. “Bring authentic data into the classroom can assist learners ... to make communicative meanings ... nonauthentic language, in some respects, actually makes the task for the language learner more difficult” (Nunan, 1999, pp.80-81). Some Studies are conducted on the relation between language learning strategies specifically metacognitive strategies, reading comprehension and text authenticity for example Li and Munby (1996), Schoonen and co-researchers (1998), Sheorey and Mokhtari (2001), and Smith (2003).

Although similar studies related to this research have been conducted in EAP context, the effect of metacognitive strategies instruction on learner’s reading comprehension performance regarding authentic and inauthentic texts has not been previously reported in EFL context. The present study focuses on the effect of instruction on planning and self-monitoring strategies on the EFL learners' reading comprehension performance. Moreover, it intends to explore the effect of text (authentic or inauthentic) on their reading comprehension performance.

This article reports a metacognitive strategy instruction study of reading in English as a foreign language. It is designed to address the following research questions:

1. Does instruction on planning strategy have any significant effect on EFL learners’ reading comprehension performance on authentic vs. inauthentic texts?
2. Does instruction on self-monitoring strategy have any significant effect on EFL learners’ reading comprehension performance on authentic vs. inauthentic texts?
3. Does instruction on metacognitive strategies have any significant effect on EFL learners’ metacognitive awareness?

Language Learning Strategy Instruction

Conscious development of reading skills is important because we are trying to equip students for the future (Nuttall, 2000). "It is impossible to familiarize them [students] with every text they will ever want to read; but what we can do is give them techniques for approaching texts of various kinds, to be used for various purposes, that is the essence of teaching reading" (Nuttall, 2000, p. 38). Strategy instruction is effective in promoting learner autonomy, or helping learners take control of their own learning (Stewner-Manzanares, Chamot, O'Malley, Küpper & Russo, 1985; Wenden & Rubin, 1987; Oxford, Talbott & Halleck, 1990). In addition, strategy training can help teachers become more aware of their students' needs and improve the relationship of their instruction to students' styles and strategies (Oxford, et al., 1990; Nyikos, 1996).

How is good strategy instruction carried out? This section explores what experts say in answer to this question, including general characteristics of optimal strategy instruction and a few instructional frameworks for strategy training.

Characteristics of Optimal Strategy Instruction

Researchers have come up with a number of characteristics of optimal strategy instruction (Iverson, 2005). As the following sections show, strategy instruction should be explicit, integrated, task-based, and individualized. It should also deal with affective factors and promote learner autonomy.

Explicit strategy instruction raises learners' consciousness both of their own strategy use and of the existence of other strategies (Oxford, 1994; Nyikos, 1996; Oxford & Leaver, 1996; Cohen, 1998).

Strategy instruction is much more effective when it is integrated into regular classroom

activities, rather than treated separately, and when numerous strategies are taught over a longer period of time (Oxford, et al., 1990; O'Malley & Chamot, 1990; Oxford, 1994; Oxford & Burry-Stock, 1995; Nyikos, 1996; Oxford & Leaver, 1996; Cohen, 1998).

Instructional Frameworks for Strategy Instruction

In addition to general guidelines about how strategy instruction should be carried out, several frameworks for strategy instruction exist. The following section focuses primarily on three different instructional frameworks.

The first model of strategy instruction, utilized in this study, is the Cognitive Academic Language Learning Approach, or CALLA (O'Malley & Chamot, 1990; Chamot & O'Malley, 1994; Chamot, Barnhardt, El-Dinary & Robbins, 1999). This approach focuses on the integration of three aspects of learning: content area instruction, academic language development, and explicit instruction in learning strategies. It is particularly targeted toward students who have at least an advanced-beginning or intermediate level of English proficiency. This model is presented very simply through five basic steps: preparation, presentation, practice, evaluation, and expansion:

In the Preparation phase, teachers provide advance organizers about the lesson, and students identify what they already know about a topic, using elaboration as a strategy. In the Presentation phase, teachers provide new information to students, using techniques which make their input comprehensible. Teachers can use advance organizers and encourage the use of selective attention, self-monitoring, inferencing, summarizing, and transfer. In the Practice phase, students engage in activities in which they apply learning strategies, often in cooperative small-group sessions. During this phase, the teacher should encourage the use of strategies such as grouping, imagery, organizational planning,

deduction, inferencing, and questioning for clarification. In the Evaluation phase, students reflect on their individual learning and plan to remedy any deficiencies they may have identified. Finally, in the Follow-Up Expansion phase, students are provided with opportunities to relate and apply the new information to their own lives, call on the expertise of their parents and other family members, and compare what they have learned in school with their own cultural experiences. (Chamot & O'Malley, 1987, p. 245)

Another model for strategy instruction is termed "Completely Informed Training," or sometimes "Strategy-Plus-Control Training." It was introduced by Oxford, et al. in 1990. It is actually presented as one part of an overall strategy training model, which focuses on steps that teachers should take when implementing strategy training, including "preparation on behalf of both students and teachers, issues that teachers should consider for instruction, the 'Completely Informed Training' method for instruction, and how teachers should follow up instruction with evaluation and revision" (Iverson, 2005, pp. 44-45).

The third instructional framework was presented by Oxford and Leaver in 1996. "It is a little different in that it is not so much a list of steps for the instructor to go through, but a division of types of instruction that can be carried out, according to the level of consciousness that each type of instruction promotes" (Iverson, 2005, p. 46).

Text Authenticity

The debate between Chomsky (1965) and Hymes (1972) led to a realization that communicative competence involved much more than knowledge of language structures and contextualized communication began to take precedence over form. This culminated in Communicative Language Teaching and paved the way for the reintroduction of

authentic texts which were valued for the ideas they were communicating rather than the linguistic forms they illustrated (Gilmore, in press). However, despite appeals for greater authenticity in language learning going back at least 30 years (O'Neill & Scott, 1974; Crystal & Davy, 1975; Schmidt & Richards, 1980; Morrow, 1981), movements in this direction have been slow. According to Gilmore (in press):

The debate over the role of authenticity, as well as what it means to be authentic, has become increasingly sophisticated and complex over the years and now embraces research from a wide variety of fields including discourse and conversational analysis, pragmatics, cross-cultural studies, sociolinguistics, ethnology, second language acquisition, cognitive and social psychology, learner autonomy, information and communication technology (ICT), motivation research and materials development ... it is important to attempt to bridge these divides and consolidate what we now know so that sensible decisions can be made in terms of the role that authenticity should have in foreign language learning in the future.

Advantages and Disadvantages of Authentic Materials

Guariento and Morely (2001) claim that we can take a lot of advantage of the use of authentic materials, but we are also aware of the reverse effects of it on learners.

The importance of using authentic material in language learning has been demonstrated conclusively (Schow, 1998). Wilkins (1976) points out authentic materials usually include incidental or improper English, which are a part of every day conversation, but can not be found in textbooks. This claim itself provides teachers with an excellent reason to integrate such materials in EFL classes.

Nunan (1999, p. 26) argues “that learners should be fed as rich a diet of authentic data

as possible because, ultimately, if they only encounter contrived ... texts, their task will be made more difficult". Authentic materials "will assist learners because they will experience the language item in interaction with other closely related grammatical and discourse elements" (Nunan, 1999, p. 27).

One of the other advantages of authentic materials is motivating the students. These materials usually promote a sense of achievement in the students. Moreover, changes are usually included in these kinds of materials, making it easy for the teachers and the learners to keep abreast of such changes (Guariento & Morley, 2001).

Authentic materials provide students with variety of materials which are not usually available in conventional teaching texts. These types of reading materials usually encourage reading for pleasure. Students are usually reluctant to read from their text books.

One of the main problems for learners may occur when students learn languages for the purpose of attaining success on an examination. In these situations, teachers have the responsibility to ensure that learners maximize their chances of success. If students are presented with authentic text they may not be given the necessary exposure to rules, patterns, or structures which they will need to achieve success on the examination. Moreover, authentic text may be a distraction to learners because it may be so apparent or introduce or focus on more language learning possibilities than a learner is able to comprehend or consolidate into examination related features (Murdoch, 1999).

It should be mentioned that the ability level of the student is an important factor in the choice of materials to be used. When teaching lower level learners, the instructor has to spend more time preparing authentic materials. Thus, without sufficient support materials, using authentic materials can add time constraints for the instructor.

It is said that authentic materials might be too culturally biased or they might be difficult for lower level learners to decode (Nunan, 1999). Often a good knowledge of cultural background is required when reading, as well as too many structures being mixed; causing lower levels problems when decoding the texts (Martinez, 2002).

Richards (2001) notes that authentic materials often contain difficult language, unneeded vocabulary items and complex language structures, which can often create problems for the teacher too. They can also become very dated, very quickly but unlike textbooks can be updated or replaced much easier and more cost effectively. The biggest problem with authentic materials is that if the wrong type of text is chosen, the vocabulary may not be relevant to the learner's needs and too many structures can create difficulty.

Also, it is mentioned that texts drawn from recognizably authentic sources may be seen as not pedagogically serious enough (Wallace, 1992). However, a number of scholars such as Widdowson (1990), Harmer (1991), and Nunan (2001) believe that the disadvantages of these materials are in minority, and that they should be included in EFL classes.

Research on learning strategies focus on diverse issues of the reading comprehension process (Li & Munby, 1996; Mason & Krashen, 1997; Schoonen et al., 1998; Sheorey & Mokhtari, 2001).

Some Studies are conducted on the relation between language learning strategies specifically metacognitive strategies, reading comprehension and text authenticity in EAP context for example Li and Munby (1996), Schoonen and co-researchers (1998), Sheorey and Mokhtari (2001), and Smith (2003).

As I know, similar studies related to this research have been conducted in EAP context,

but it has not been previously reported in EFL context. The present study focuses on the effect of instruction on planning and self-monitoring strategies on the EFL learners' reading comprehension performance. Moreover, it intends to explore the effect of text (authentic or inauthentic) on their reading comprehension performance.

Methodology

Participants

The participants in this study were 93 university students, 55 females and 38 males, majoring in Teaching English as a Foreign Language at Kermanshah Azad University. The subjects were assigned to groups by the university. Three intact classes were selected for the purpose of this study. Two classes were randomly assigned as experimental groups and the last one as the control group. Their age ranged from 20 to 25. They were Persian native speakers. Their homogeneity in terms of language proficiency was established through TOEFL (1995 version). In order to establish the homogeneity of the three groups in terms of general language proficiency, a one-way ANOVA was conducted to examine the probable difference among the performance of the three groups (EG1, EG2, and CG) before the experiment. The results indicated that there was not any significant difference between the mean scores of the subjects in the two experimental groups and the control group. Furthermore, they were homogeneous regarding eight years of English education in school and L1 background.

Instrumentation

To probe the research questions posed by the researcher, three instruments were used:

1. The Test of English as a Foreign Language (TOEFL, 1995 version)

This test was administered as a standardized measure to check the homogeneity of subjects in terms of language and also it was used as a criterion to validate teacher-made reading comprehension test. This TOEFL test consists of 100 items of three sections of structure and written expressions (40 items), vocabulary (30 items) and reading comprehension (30 items). Due to administrative limitations listening comprehension section was not included. The reliability of the test, as estimated against KR - 21 measure of internal consistency, turned out as .081 in the pilot study.

2. SILL Questionnaire, version 7.0 (Oxford, 1990)

In the present study, Oxford's (1990) Strategy Inventory for Language Learning (SILL) was chosen for this study because it is "perhaps the most comprehensive classification of learning strategies to date" (Ellis, 1994, p. 539) and has been widely used. Its Cronbach alpha reliability coefficients range from 0.89 to 0.98 in various studies (Oxford, 1986; Oxford & Nyikos, 1989; Ehrman & Oxford, 1989, 1990; Wildner-Bassett, 1992; Nyikos & Oxford, 1993; Bedell, 1993; Oxford & Burry, 1993; Oxford & Ehrman, 1995, cited in Oxford, 1996).

Reliability of the SILL is high across many cultural groups. Its validity rests on its predictive and correlative link with language performance as well as its confirmed relationship to sensory preferences (Oxford, 1996). This questionnaire was given before and after strategy instruction to ask the students about the frequency with which they used these two metacognitive strategies (planning and self-monitoring) SILL is composed of 50 items in six categories, in which part D is related to metacognitive strategies. SILL consists of series of statements such as "I try to find out how to be a better learner of English" to which students are asked to respond to a five-point Likert scale ranging from

1 (never or almost never) to 5 (always or almost always).

3. A reading comprehension test

This teacher-made reading comprehension test includes two sections of authentic and inauthentic texts. Four authentic reading texts were selected from *Reader's Digest* (September and February issues, 2005) which has interesting, popular, universal, and reader-friendly topics (Hwang, 2005). The magazine was chosen because according to Porter Ladousse (1999), magazines are example of authentic materials. Furthermore, they include different types of text with various illustrations which help students in implementing language learning strategies. Inauthentic reading comprehension texts were selected from chapters two, seven and nine of their course book, *Reading Skillfully III* (Mirhassani, 2003). It was taught during the course of Reading Comprehension III. On the whole, this reading comprehension test included 40 multiple choice questions (see Appendix).

Pilot Study

In order to assess the reliability of the reading comprehension test, the reading comprehension test was pretested with a sample group of 35 students having characteristics similar to the target group. Then after interpreting the collected data, five weak, malfunctioning and non-functioning items were removed from the whole test, and some were modified. After modification, estimated internal consistency measures (KR-21) revealed that the adequate reliability of 0.81 was attained. To establish the empirical validity of the reading comprehension test, the Pearson product-moment correlation between the reading comprehension tests (authentic and inauthentic) and TOEFL were calculated respectively as 0.67 and 0.60 which are significant at 0.05 level of significance.

Procedure

At the first phase, the TOEFL (1995 version) was administered to all the students. As it was mentioned above, it was used both to homogenize students regarding language proficiency and as a criterion to validate the reading comprehension test.

Table 1: Levene Test of Homogeneity of Variances for the Three Groups

Levene Test of Homogeneity of Variances			
Levene Statistic	df1	df2	Sig.
1.306	2	90	.276

At the second phase (next session), version 7.0 of the SILL that is a self-report instrument was administered to all groups. It assesses the frequency with which the subjects use a variety of techniques for foreign language learning. It was given before strategy instruction to ask the students about the frequency with which they use these two metacognitive strategies. On average, students completed the SILL within 25 minutes.

At the third phase, experimental groups received instruction on planning and self-monitoring based on CALLA model. They received instruction on metacognitive strategies for five sessions of ninety minutes during the semester. Both experimental and control groups worked on authentic and inauthentic texts (some articles from both *Readers' Digest* and *Reading Skillfully III*).

The five steps of the CALLA model (Preparation, Presentation, Practice, Evaluation, and Expansion) which are used in this study for instruction of planning and self-monitoring strategies.

1. Preparation. Effective strategy training requires a certain amount of preparation

which involves both input and output from the students. First of all, the teacher needs to elicit a certain amount of information from the students, in order to be better informed about the students' needs and make appropriate decisions about which strategies to teach and how to teach them (O'Malley & Chamot, 1990; Oxford, 1990; Oxford, et al., 1990; Chamot & O'Malley, 1994; Oxford & Leaver, 1996; Cohen, 1998; Chamot, et al., 1999).

2. Presentation. In the second stage of strategy instruction, learners are presented with a specific strategy or set of strategies to be taught. It is helpful for getting students to think about the strategy explicitly, discuss it, and remember it. The teacher describes how the strategy is used, why it is important and how it applies to the specific task at hand, and models it for the students with several examples (Oxford, et al., 1990).

3. Practice. In the third stage of strategy instruction, learners are given the opportunity to practice the strategy or set of strategies that are being targeted. One of the important elements of this phase is that it is integrated into the regular class work, so the students can make a solid connection between the new strategy and authentic tasks that they must accomplish. It is also important that the tasks are challenging enough to require the use of the new strategy, but not so difficult that they are overwhelming (Chamot, et al., 1999).

4. Evaluation. In the fourth stage of strategy instruction, learners reflect on their use of a specific strategy or strategies and evaluate its usefulness. This helps promote learner autonomy, and enable the instruction to be more individualized – both ingredients of good strategy instruction (O'Malley & Chamot, 1990; Oxford, 1990; Oxford, et al., 1990; Chamot & O'Malley, 1994; Cohen, 1998; Chamot, et al., 1999).

5. Expansion. In the final stage of strategy instruction (according to the CALLA model), learners are shown how to transfer the new strategy to different situations or tasks, and given opportunities to practice it. (O'Malley & Chamot, 1990; Oxford, 1990; Oxford, et

al., 1990; Chamot & O'Malley, 1994; Cohen, 1998; Chamot, et al., 1999).

At the fourth phase, SILL questionnaire and reading comprehension test were administered to all groups.

Data Analysis

In order to establish the homogeneity of the three groups in terms of general language proficiency, a one-way ANOVA was conducted to examine the probable difference among the performance of the three groups (EG1, EG2, and CG) before the experiment.

In order to probe the first and the second null hypotheses, and examine the effectiveness of explicit instruction of metacognitive strategies and compare their improvement with their counterparts in the control group, all the three groups took part in the same reading comprehension tests after completing the instruction. In order to test these two null hypotheses, the statistical technique of multivariate ANOVA was run. In this study planning and self-monitoring were independent variables and authentic and inauthentic texts in reading comprehension test were dependent variables.

In order to probe the last null hypothesis and examine the effect of metacognitive instruction on EFL learners' metacognitive awareness three paired samples t-tests were conducted.

Results

In order to establish the homogeneity of the three groups in terms of general language proficiency, a one-way ANOVA was conducted to examine the probable difference among the performance of the three groups (EG1, EG2, and CG) before the experiment.

The results indicated that there was not any significant difference between the mean

scores of the subjects in the two experimental groups and the control group.

Table 2: Descriptive Statistics for the Three Groups Performance on TOEFL

	<i>N</i>	<i>Mean</i>	<i>Std. Deviation</i>	<i>Std. Error</i>	<i>95% Confidence Interval for Mean</i>		<i>Minimum</i>	<i>Maximum</i>
					<i>Lower Bound</i>	<i>Upper Bound</i>		
<i>CONTROL</i>	31	57.29	15.674	2.815	51.54	63.04	29	99
<i>PLANNING</i>	31	65.16	16.988	3.051	58.93	71.39	30	99
<i>MONITORING</i>	31	57.45	13.825	2.483	52.38	62.52	38	99
<i>Total</i>	93	59.97	15.817	1.640	56.71	63.23	29	99

In order to probe the first and the second null hypotheses, and examine the effectiveness of explicit instruction of metacognitive strategies and compare their improvement with their counterparts in the control group, all the three groups took part in the same reading comprehension tests after completing the instruction. In order to test these two null hypotheses, the statistical technique of multivariate ANOVA was run.

As Table 4-3 shows the three groups have significant difference in their performance on reading comprehension tests but the difference in the two experimental groups is not significant.

Table 3: Pairwise Comparisons of the Three Groups on the Reading Comprehension Test

Measure: MEASURE_1

(I) GROUP	(J) GROUP	Mean Difference (I-J)	Std. Error	Sig. ^a	95% Confidence Interval for Difference ^a	
					Lower Bound	Upper Bound
CG	EG1	-1.871 *	.479	.000	-2.823	-.919
	EG2	-1.468 *	.479	.003	-2.419	-.516
EG1	CG	1.871 *	.479	.000	.919	2.823
	EG2	.403	.479	.402	-.548	1.355
EG2	CG	1.468 *	.479	.003	.516	2.419
	EG1	-.403	.479	.402	-1.355	.548

Based on estimated marginal means

*. The mean difference is significant at the .05 level.

a. Adjustment for multiple comparisons: Least Significant Difference (equivalent to no adjustments).

The results of the Tests of Between-Subjects Effects (Table 4) compare the mean scores of the three groups on the authentic and inauthentic sections of reading comprehension test. The F-values for the authentic and inauthentic sections are 5458.1 and 8.4. At 2 and 90 degrees of freedom, these F-values are greater than the critical value of 3.09 indicating that the three groups performed differently on the two tests.

Table 4: Test of Between-subjects Effect

Measure: MEASURE_1
 Transformed Variable: Average

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Intercept	39019.548	1	39019.548	5485.189	.000
GROUP	120.226	2	60.113	8.450	.000
Error	640.226	90	7.114		

Table 5 shows that mean difference in two authentic and inauthentic tests is significant.

Table 5: Pairwise Comparisons of the Authentic and Inauthentic Tests

Measure: MEASURE_1

(I) TEST	(J) TEST	Mean Difference (I-J)	Std. Error	Sig. ^a	95% Confidence Interval for Difference ^a	
					Lower Bound	Upper Bound
1	2	.946 *	.217	.000	.516	1.376
2	1	-.946 *	.217	.000	-1.376	-.516

Based on estimated marginal means

- *. The mean difference is significant at the .05 level.
- a. Adjustment for multiple comparisons: Least Significant Difference (equivalent to no adjustments).

In order to probe the last null hypothesis and examine the effect of metacognitive instruction on EFL learners' metacognitive awareness three paired samples t-tests were conducted (Table 6). As the level of significance shows, it is smaller than .05 in experimental groups. Therefore, metacognitive strategy instructions have positive effect on experimental groups' metacognitive awareness. As the last row of Table 4 shows, level of significance is bigger than .05, in other words, there is no significant difference in the

control group awareness. It can be concluded that the third null hypothesis is rejected.

Table 6: Results of the Paired-samples T- tests in SILL Questionnaire in EG1, EG2, & CG

	Paired Differences					t	df	Sig. (2-tailed)
	Mean	SD	Std. Error of Mean	95% confidence Interval of the Difference				
				Lower	Upper			
Pair 1 SILL1- SILL2 (EG1)	-258	3.225	.579	-7.441	-5.075	-10.806	30	.000
Pair 1 SILL1 - SILL2 (EG2)	-.097	1.758	.316	-5.742	-4.452	-16.143	30	.000
Pair 1 SILL1 - SILL2 (CG)	.0967	.30054	.05398	-.01346	.20701	1.793	30	.83

Discussion and Conclusion

The present study examined the effect of metacognitive strategies (planning & self-monitoring) instruction on reading comprehension performance. It was also intended to investigate probable effect of text type (authentic and inauthentic texts) on the EFL learners' performance in reading comprehension tests. Moreover, this research explored the effect of metacognitive strategy instruction on EFL learners' metacognitive awareness.

The first null hypothesis predicting that planning strategy instruction has no significant effect on the learners reading comprehension performance was rejected. As the results showed the first experimental group (EG1) outperformed the control group (CG).

Furthermore, it showed that the subjects performed better in the authentic reading comprehension test than the inauthentic one.

The second null hypothesis predicting that self-monitoring strategy instruction has no significant effect on the learners reading comprehension performance was also rejected. As the results showed the second experimental group (EG2) outperformed the control group (CG). Furthermore, it showed that the subjects' performance in the authentic reading comprehension test was better than the inauthentic one like the EG1.

Considering the third null hypothesis, i.e., metacognitive strategy instruction has no significant effect on EFL learners' metacognitive awareness, three paired sample tests were run to investigate the degree of EFL learners' awareness. It showed that students' awareness significantly increased after metacognitive strategy instruction.

The major concern of the present study was to explore the effectiveness of metacognitive strategies instruction on the reading comprehension performance of the EFL students and their awareness to metacognitive strategies. As it was shown, the experimental groups outperformed the control group on the reading comprehension performance. Thus, the metacognitive strategy instruction seems to have contributed to the improvement of students' reading comprehension performance. In other words, the explicit instruction and practice the experimental groups received about how to plan and how to monitor their reading, contributed to this improvement. In addition, the findings of this study indicate that metacognitive strategies instruction increases the experimental groups' metacognitive awareness. Meanwhile, it should be mentioned that both experimental and control groups outperformed in the authentic section of the reading comprehension test.

The findings of this study support other empirical studies on the effect of strategy

instruction on reading comprehension performance. People like Paris and Oka, 1986, O'Malley, 1987, Cross and Paris, 1988, Barnett, 1988, Mokhtari and Reichard, 2002, Trenchs Parera, 2006, Philip and Kim Hua, 2006, just to name a few, who worked on this issue. Moreover, it can be asserted that the model (CALLA) used to teach metacognitive strategies was a practical and useful one.

The findings are also in line with the positive effect of authentic materials on learners' reading comprehension performance and support the following researchers' findings such as Vigil, 1987, Allen, Bernhardt, Berry and Demel, 1988, Shrum and Glisan, 1994, Hauptman, 2000, and Pritchard and Nasr, 2004.

In addition, the findings of this study support the following researchers' findings on the correlation among metacognitive strategies, authenticity and reading comprehension in EAP context like Ems-Wilson, 2000, Smith, 2003, and Cain, 2004. The findings of the present study have implications for learners, teachers, and textbook writers in the realm of TEFL in particular and education in general.

This study is limited in terms of having two metacognitive strategies for instruction and also not considering gender as a variable. Therefore, this research needs other enough replications and further studies are needed to shed more light on the issue.

References

- Allen, E., Bernhardt, E. B., Berry, M. T. & Demel, M. (1988). Comprehension and text genre: An analysis of secondary foreign language readers. *Modern Language Journal*, 72(2), 163–172.
- Barnett, M. A. (1988). *Teaching reading in a foreign language*. ERIC Digest.
- Cain, S. D. (2004). *Using comprehension strategies with authentic text in a college*

- chemistry course*. Unpublished dissertation. University of Maryland, College Park.
- Celce-Murcia, M. & Olshtain, E. (2000). *Discourse and content in language teaching: A guide for language teachers*. Cambridge: Cambridge University Press.
- Chamot, A. U., & O'Malley, J. (1987). The cognitive academic language learning approach: A bridge to the mainstream. *TESOL Quarterly*, 21(2), 227-249.
- Chamot, A. U., & O'Malley, J. (1994). *The CALLA handbook: Implementing the cognitive academic language learning approach*. Reading, MA: Addison- Wesley.
- Chamot, A. U., Barnhardt, S., El-Dinary, P., & Robbins, J. (1996). Methods for teaching learning strategies in the foreign language classroom. In Rebecca L. Oxford (Ed.), *Language learning strategies around the world: Cross-cultural perspectives*. (Technical Report #13) (pp. 175-187). Honolulu: University of Hawaii, Second Language Teaching & Curriculum Center.
- Chamot, A. U., Barnhardt, S., El-Dinary, P., & Robbins, J. (1999). *The learning strategies handbook*. New York: Addison Wesley Longman, Inc.
- Chomsky, N. (1965). *Aspects of the theory of syntax*. Boston: MIT Press.
- Cohen, A. D. (1998). *Strategies in learning and using a second language*. London: Longman.
- Cohen, A. D., Weaver, S. J., & Li, T. Y. (1998). *The impact of strategies-based instruction on speaking a foreign language*. (CARLA Working Paper Series No. 4.) Minneapolis: Center for Advanced Research on Language Acquisition, University of Minnesota.
- Cross, D. R., & Paris, S. (1988). Developmental and instructional analyses of children's metacognition and reading comprehension. *Journal of Educational Psychology*, 80(2), 131.

- Crystal, D. & Davy, D. (1975). *Advanced conversational English*. Harlow: Longman.
- Ellis, R. (1994). *The study of second language acquisition*. Oxford: Oxford University Press.
- Ems-Wilson, J. (2000). Using elaborative interrogation to induce characteristics of polar and nonpolar solvents from animations of their molecular structures. *Dissertation Abstracts International*, 60(12), 4369A. (UMI No. 9955543)
- Get wise on serving size. (2005, February). *Reader's Digest*. 200-202.
- Gilmore, A. (In Press). *Authentic materials & authenticity in foreign language learning*. Nottingham University, UK.
- Grabe, W. (2002). *Reading in a second language*. Oxford: Oxford University Press.
- Guariento, W. & Morely, J. (2001). Text and task authenticity in the EFL classroom. *ELT Journal*, 55(4), 347-353.
- Harmer, J. (1991). *The practice of English language teaching* (2nd ed.). London: Longman.
- Hauptman, P. C. (2000). Some hypotheses on the nature of difficulty and ease in second language reading: An application of schema theory. *Foreign Language Annuals*, 33(6), 622-630.
- Hwang, C. C. (2005). *Effective EFL education through popular authentic materials*. Retrieved March 2005, from http://www.asian-efl-journal.com/March_05_ch.pdf
- Hymes, D. (1972). On communicative competence. In J. B. Pride & J. Holmes (Eds.), *Sociolinguistics* (pp. 269-93). Harmondsworth: Penguin.
- Iverson, K. J. E. (2005). *The instruction of language learning strategies for low-proficiency ESL learners*. Unpublished dissertation. Minnesota, Hamline University.

- Lazaraton, A. & Skuder, P. F. (1997). *Evaluating dialogue authenticity in ESL speaking texts*. Paper presented at the 31st Annual TESOL Convention, Orlando, FL.
- Little, D., Devitt, S. & Singleton, D. (1989). *Learning foreign languages from authentic texts: Theory and practice*. Dublin: Authentik in association with CILT.
- Martinez, A.G. (2002). *Authentic materials: An overview on Karen's linguistic issues*. Retrieved from <http://www3.telus.net/linguisticsissues/authenticmaterials.html>
- Mirhassani, A. (2003). *Reading skillfully III*. Tehran: Zabankadeh.
- Mokhtari, K. & Reichard, C. (2002). Investigating the strategic reading processes of first and second language readers in two different cultural contexts. *System*, 32, 379-394.
- Morrow, K. (1981). Principles of communicative methodology. In K. Johnson & K. Morrow (Eds.). *Communication in the classroom* (pp. 59–69). Harlow: Longman.
- Muñiz-Swicegood, M. (1994). The effects of metacognitive reading strategy training on the reading performance and student reading analysis strategies of third grade bilingual students. *Bilingual Research Journal*, 18, 83-97.
- Murdoch, Y. (1999). *Using authentic text in language classroom*. Retrieved from <http://www.cels.bham.ac.uk/resources/essays/yvette2.pdf>.
- Nilsson, K. (2003). *A meta-search approach to locating and classifying reading material for learners of Nordic languages*. Unpublished thesis, Uppsala University.
- Nunan, D. (1999). *Second language teaching and learning*. USA: Newbury House.
- Nunan, D. (2001). *Task-based language teaching*. Cambridge: Cambridge University Press.
- Nuttall, C. (2000). *Teaching reading skills in a foreign language*. UK: Macmillan Heinemann.

- Nyikos, M. (1996). The conceptual shift to learner-centered classrooms: Increasing teacher and student strategic awareness. In R. L. Oxford (Ed.), *Language learning strategies around the world: Cross-cultural perspectives* (pp. 109-117). Honolulu: Second Language Teaching and Curriculum Center.
- O'Malley, J. M. (1987). The effects of training on the use of learning strategies on learning English as a second language. In A. Wenden & J. Rubin (Eds.), *Learning strategies in language learning* (pp. 133-144). Cambridge: Prentice Hall International.
- O'Malley, J. M. & Chamot, A. U. (1990). *Learning strategies in second language acquisition*. Cambridge: Cambridge University Press.
- O'Neill, R., & Scott, R. (1974). *Viewpoints*. London: Longman
- Oxford, R. L. (1990). *Language learning strategies: What every teacher should know*. New York: Newbury House Publishers.
- Oxford, R. L. (1994). *Language learning strategies*. (An updated). Eric Clearinghouse on Languages and Linguistics, (online).
- Oxford, R. L. (1996). Employing a questionnaire to assess the use of language learning strategies. *Applied Language Learning*, 7(1&2), 25-45.
- Oxford, R. L., & Burry-Stock, J. A. (1995). Assessing the use of language learning strategies worldwide with the ESL/EFL version of the strategy inventory for language learning (SILL), *System*, 23(1), 1-23.
- Oxford, R. L., & Leaver, B. L. (1996). A synthesis of strategy instruction for language learners. In R. L. Oxford (Ed.), *Language learning strategies around the world: Cross-cultural perspectives* (pp. 227-246). Honolulu: Second Language Teaching & Curriculum Center, University of Hawaii.

- Oxford, R. L., Talbott, V., & Halleck, G. (1990). *Language learning strategies, attitudes, motivation, and self-image of students in a university intensive ESL program*. Paper presented at the 24th Annual TESOL Convention, San Francisco, CA.
- Paris, S. G. & Oka, E. (1986). Children's reading strategies, metacognition, and motivation. *Developmental Review, 6*, 25-26.
- Philip, B. & Kim Hua, T. (2006). Metacognitive strategy instruction for reading: correlation of cognition. *Journal e-Bangi, 1*, 1-27.
- Porter Ladousse, G. (1999). Review of exploring spoken English. *ELT Journal, 53*(2), 139–141.
- Pritchard, M. O. & Nasr, A. (2004). Improving reading performance among Egyptian engineering students: principles and practice, *English for Specific Purposes, 23*(4), 425-445.
- Richards, J. C. (2001). *Curriculum development in language teaching*. Cambridge: Cambridge University Press.
- Schmidt, R. W. & Richards, J. C. (1980). *Speech acts and second language learning*. *Applied Linguistics, 1*(2), 342-356.
- Schow, S. (1998). *The higher level of inferred learning: Second language vocabulary acquisition through authentic texts*. Unpublished MA thesis. Brigham Young University (BYU).
- Shrum, J. & Glisan, E. (1994). *Teacher's handbook: Contextualized language instruction*. Boston: Heinle & Heinle.
- Smith, B. L. (2003). The use of why-questions to enhance college science students' comprehension of their textbooks. *Dissertation Abstracts International, 64*(1), 62A. (UMI No. 3078197).

- Stewner-Manzanares, G., Chamot, A. U., O'Malley, J. M., Küpper, L., & Russo, R. P. (1985). *Learning strategies in English as a second language instruction: A teacher's guide*. Washington, DC: National Clearinghouse for Bilingual Education. (ERIC Document Reproduction Service No. ED338107).
- Test of English as a Foreign Language. (1995). *TOEFL test and score manual*. Princeton, NJ: Educational Testing Service.
- The latest advances in research, treatment and technology: Keep your brain young. (2005, September). *Reader's Digest*, 59-60.
- Trenchs Parera, M. (2006). *Reading strategies and strategy awareness in three EFL educated readers of English literary texts*. Retrieved from [www.atlantisjournal.org /Papers/28_2/MTrenchs.pdf](http://www.atlantisjournal.org/Papers/28_2/MTrenchs.pdf).
- Vigil, V. (1987). *Authentic texts in the college-level Spanish I class as the primary vehicle of instruction*. Unpublished doctoral dissertation, University of Texas at Austin.
- Wallace, C. (1992). *Reading*. Oxford: Oxford University Press.
- Wenden, A. & Rubin, J. (1987). *Learner strategies in language learning*. New Jersey: Prentice Hall.
- Widdowson, H. G. (1990). *Aspects of language teaching*. Oxford: Oxford University Press.
- Wilkins, D. A. (1976). *Notional syllabus*. Oxford: Oxford University Press.

The Ability of Taiwanese College Freshmen to Read and Interpret Graphics in English

Tzung-yu Cheng
China Medical University, Taiwan
marlinytc@hotmail.com

Bio Data:

Tzung-yu Cheng is an associate professor at China Medical University in Taiwan where he teaches courses in critical reading and academic reading. His current research interest includes content area reading, reading assessment, and sociology of language learning and teaching.

Abstract

This study is intended to assess the ability of Taiwanese college freshmen to read and understand English graphics, and at the same time evaluate the quality of English text-usage training dispensed in the high schools in Taiwan. The subjects, 211 freshmen, were drawn from a medical university in central Taiwan. They represented a group of above average students among the Taiwanese college freshmen majoring in sciences and engineering. The research instrument was the criterion-referenced graphics test, *GASS*. The data show that the pass rate is 13.27%. The survey of reading difficulty and the *t*-test, $t(177) = 5.933, p < .001$, suggest that in reading graphics, the ability to read English itself is a major cause of difficulty. The study concludes with suggestions for further studies and various recommendations for improving the teaching and learning of graphics in both high school and freshman college English classes.

Keywords: EFL Graphic Literacy; Graphic Literacy Assessment; Content Area Reading; Text Aids

Introduction

In Taiwan and in many southeastern Asian countries, academic excellence in English is pursued for higher status and as a means to keep abreast of globalization. Moreover, it is a popular practice in the academic world in many Asian countries to acquire English reading skills to study scholarly works in English, and to write texts and publish papers in English. One special feature of these academic texts is their extensive use of text aids to facilitate the comprehension and learning of their contents. Such learning and comprehension aids include a broad variety of graphs as shown below (Hunter, Crismore & Pearson, 1987):

1. Sequential graphs such as flowcharts, time lines and process charts;
2. Quantitative graphs such as bar graphs, line graphs, pie charts, bar charts;
3. Pictographs such as maps for physical, political and special purposes;
4. Tables and charts such as schedules, time tables and comparison charts.

These visual-spatial text adjuncts help communicate information and support thinking or learning processes (Schnotz, 2002). However, in the English native setting, text-usage skills instruction is often neglected. Equally, there are no studies exploring the current status of graphic literacy instruction in high school and college freshman English classes in Taiwan. Given the dearth of such related research, the objective of this paper is to determine the extent to which high school graduates in Taiwan are adequately prepared for reading and interpreting graphics in academic English.

In this study, the broad rationale is first presented along with references to past research as regards the functions of text aids. Following is the case study of this question, carried out among a group of Taiwanese college students. It is hoped that this preliminary

research will lead the way to more large-group studies. The findings of this study may be of interest to those who are involved in EFL/ESP programs in Asia.

Literature Review

“A good graphic representation can force us to discover things without knowing in advance what we are looking for” (Wainer, 1994, p.14). Graphs present concepts in a concise manner, distill a great deal of descriptive writing into a small amount of space, restate the information, and illustrate the author’s ideas (Askov & Kamm, 1982; Weintraub, 1967). Fry (1981) states that graphs are an important communication tool and often communicate better than words. They are concrete and direct, but truncated, and they require interpretation. Schnotz (2002, p.107) points out, “...verbal information and pictorial information can be kept simultaneously in working memory and, accordingly, it is easier for the learner to make cross-connections between the two different codes and later retrieval information.” As a result, he continues, if the subject matter is complex and/or if learners have low prior knowledge, then graphic representations increase comprehension.

The idea of using text adjuncts dates back to the seventeenth century, when Comenius in his ‘Didacta Magna’ suggested that envisioning information is extremely important for effective learning. However, it is only since the 1970s that research on comprehension of visual displays has been investigated systematically (Schnotz, 2002). Empirical results have documented that graphics facilitate comprehension and learning from the texts (Arnold & Dewyer, 1975; Bodemer, Ploetzner, Bruchmuller & Hacker, 2005; Booher, 1975). Booher (1975) found that the mixture of pictorial information and verbal instructions facilitates perceptual-motor tasks. Arnold and Dewyer (1975) reported that

the pictorial supplementation of verbal information, such as the use of diagrams, resulted in better performance and comprehension. Bodemer, Ploetzner, Bruchmuller and Hacker (2005) investigated the benefit of an instructional support method during learning with multiple representations, simulations, and animations. Three types of information integration were compared in two consecutive experiments: (1) presentation of non-integrated information, (2) presentation of pre-integrated information, and (3) active integration of information. They found that the integration of static representations before processing dynamic visualizations yielded better performance, and can provide a basis for a more systematic and goal-oriented behavior during simulation-based discovery learning.

Other studies have also documented that among readers of different ability levels, visual displays can have a supporting function for understanding and learning difficult materials (DeStefano & LeFevre, 2007; Montali & Lewandowski, 1996; Schnotz, 2002). In Montali and Lewandowski's study (1996), less skilled readers performed at a level commensurate with average readers when science and social studies passages were highlighted on computer screen while being voiced. DeStefano and LeFevre (2007), reviewing the cognitive load in hypertext reading, found that readers with low working memory and low prior knowledge were usually disadvantaged in hypertext. However, if the hypertext structure was hierarchical and consistent with that of the knowledge domain, learners with low prior knowledge were benefited. Schnotz (2002), in a review of the supportive function of visual-spatial text adjuncts, concluded that visual-spatial adjuncts seems to be especially evident with learners of low prior knowledge and low verbal skills: The more difficult a learning content is, the higher the learner's frequency of looking at adjunct visual displays.

Despite the strengths of graphic displays on reading comprehension, some experts in

graphic literacy warn that the very strength of graphs can create a problem (Lowe, 1993; Schnotz, 2002; Weintraub, 1967). Weintraub (1967) points out that graphic displays condense information; they are frequently difficult to interpret. In addition, Schnotz (2002) suggests that visual displays do not support communication, thinking, and learning automatically. The learner requires specific graphic schemata in order to be able to read off information from the visual-spatial configuration (Lowe, 1993). Nevertheless, learners often underestimate the informational content of pictures and believe that a glimpse would be enough for understanding and for extracting the relevant information (Mokros & Tinker, 1987; Schnotz, 2002; Weidenmann, 1989). Consequently, the informational content of pictures may be overlooked by students who do not have experience in reading them (Askov & Kamm, 1982).

Effective learning with visual-spatial text adjuncts can be fostered through instructional design by the teacher or instructional material and through adequate processing strategies by the learner (Schnotz, 2002). In an English-as-mother-tongue setting, students can learn graphic reading skills from their courses in mathematics, social studies, science, and reading from elementary school through to high school. However, Heilman, Blair and Rupley (1981) point out that teachers are not teaching text-usage skills and that these skills are often neglected. Studies have also found that secondary students reading at or below grade level experience confusion about the purpose and function of textbook aids (Mateja & Wood, 1982), and students tend to make inadequate use of these text aids while using textbooks (Dillner & Olson, 1982). Thomas and Moorman (1983) traced the causes and concluded that elementary teachers may skip lessons on study skills in the hope that secondary school teachers will compensate. Meanwhile, junior and senior high instructors assume students in their classrooms have already been taught how to utilize

the graphic representations in textbooks.

As an antidote to existing deficiencies in graphic literacy and to ensure that high school graduates are able to use graphic and text-usage skills, the Georgia Curriculum Guide for Social Studies outlined its own objectives for graphic instruction. In the core courses and college preparatory courses such as World History, Citizenship, and Economics, the objectives include enabling students to make a table of contents; to develop charts, tables, graphs, and grids to appropriately convey information; to use features of books for information; to interpret keys or legends for map reading; and to infer from data on a map or combination of maps (Gillespie, 1988).

With the swift emergence of new technologies, the use of graphic information in learning and instruction has become a specific challenge for educators (Schnotz, 2002). The effective supportive function of graphics on comprehension and learning has been well evidenced in the United States (Arnold & Dewyer, 1975; Bodemer, Ploetzner, Bruchmuller & Hacker, 2005; Booher, 1975; DeStefano & LeFevre, 2007; Montali & Lewandowski, 1996; Gillespie, 1988; Schnotz, 2002; Weintraub, 1967). To the contrary, studies concerning EFL learners' ability to read and interpret graphics in English do not exist in Taiwan. An analysis of the Mandarin content textbooks used in elementary schools and junior and senior high schools shows that instruction in graphic literacy also takes place in courses such as mathematics, science, and social studies. Do the teachers, like their counterparts in the United States, skip the instruction and assume that their students have mastered the skills in the lower grades? Furthermore, no empirical studies have documented that the learning of Mandarin graphic reading skills could be transferred to the reading of English graphics. Equally, there are no studies exploring the current status of graphic literacy instruction in high school and college freshman English

classes in Taiwan.

As graphics facilitate comprehension and the learning of information, the possession of graphic reading skills may compensate for the deficiencies in students' comprehension of the college discipline-specific texts. Accordingly, this study proposes to assess the ability of college freshmen in Taiwan to read and interpret graphics in English-language texts, especially these static text adjuncts found in college content area textbooks. This study was designed to answer the following four questions:

1. Are college freshmen ready to use the text aids and graphics used in English-language text books to facilitate comprehension?
2. Which of these graphics--graphs, tables of contents, tables and charts, index, time lines, and maps--are college freshmen most likely to interpret correctly?
3. Did college freshmen have any experience reading the graphics in English at high school?
4. Is an insufficient grasp of English a cause of difficulties in reading and interpreting graphics?

Method

Participants

The subjects consisted of 211 freshmen students at a medical university in central Taiwan. These participants came from five departments. They were selected to represent a group of above-average students regarding their overall school achievement among Taiwanese college freshmen majoring in sciences and engineering. The criteria used to define the ability of the participants are based on the results of the July Joint College

Entrance Examination held annually in Taiwan (CEEC, 2008).

Taiwanese high school students, at the end of Grade 10, are required to choose a track of study between humanities and sciences which include engineering and life sciences. Usually students proficient in mathematics choose the science track. In the examination, students in the science track are required to take five or six subjects: Mandarin Chinese, English, mathematics for science, biology, chemistry, and physics. Students in humanities are required to take five subjects: Mandarin Chinese, English, mathematics for humanities, history and geography. Both tracks take the same Mandarin and English tests as they read and use the same learning materials. The results of the tests are weighted on a scale of zero to 100. In 2007, 100,117 high school graduates registered for the examination. Among them, 49,788 students took the examination for the science track (CEEC, 2008).

In estimating the standing of the participants' overall school achievements among the nation's high school graduates, the following steps were followed:

1. The researcher listed the respective raw "minimum college department admittance score" (MCDAS) of the five departments in 2007: 320.97, 318.81, 280.70, 275.73, and 251.95. Then, the researcher averaged the MCDAS of the departments by the six required test subjects, which yielded the following mean MCDAS: 53.495, 53.135, 46.783, 45.955, and 41.991.
2. Next, the researcher computed the national mean MCDAS for the examinees in the science track by following this formula (sum of the national mean of each test subject \times the examinees taking each subject) \div (sum of the examinees in the six subjects): [Mandarin (54.46 \times 100,070) + English (31.12 \times 100,059) +

mathematics for science $(36.09 \times 49,788) +$ biology $(56.94 \times 30,002) +$ chemistry $(43.55 \times 46,236) +$ physics $(32.81 \times 46,092)] \div (100,070 + 100,059 + 49,788 + 30,002 + 46,236 + 46,092)$. The computations yielded a national mean MCDAS of 41.893, which is lower than any of the five departments by a range from 11.602 to 0.098 points.

Research Instrument

Graphics Assessment for Social Studies (GASS) was the measure for this study (Gillespie, 1988). GASS is a criterion-referenced test. It was designed to measure the ability of high school students to read and interpret graphics commonly found in social studies textbooks. The test is an untimed test and consists of 72 questions. The questions are grouped into six sections with twelve questions each. Each section measures the ability to read and interpret one of six graphic representations or images: graphs, tables and charts, time lines, a table of contents, an index, and maps.

Content validity evidence was obtained from four social studies teachers who assessed the item-objective congruence. The consistency was 97.06%. Predictive validity was estimated by comparing the test scores to the students' semester grades. The coefficient for test and grades was .64, and for retest and grades was .70. The internal consistency index for the first test was .95 and was .91 for the retest, using the Kuder-Richardson Formula 21 (Gillespie, 1988).

The standard of performance (cut-off criterion) was set at 85%. As Tyler (1973, p. 105) points out, "A child has demonstrated mastery ... when he has performed correctly 85% of the time. Some small allowance, like 15%, is needed for lapses common to all people." Gronlund (1982) also suggests that the mastery level on a multiple-choice test be set at

85%. Four high school social studies teachers, one test and assessment specialist, and six reading education specialists were invited to rate each item consistently with this criterion. The question “Should every high school graduate be able to answer this item correctly?” was used as a guide to rate the items (Jaeger, 1979). The mean of the responses made by the panel members was 85% and the median was also 85% (Gillespie, 1988).

This test was selected for the study because very few tests that exist now are designed to assess the ability to read and interpret graphics. Frequently-used tests, such as the *California Achievement Test* or the *Test of Academic Progress*, contain very few subsections measuring graphic skills, leaving judgments about students’ abilities tenuous at best. For five years prior to this study, the researcher administered parts of the instrument as remediation measures for Taiwanese college students attending his English classes. This test, though developed for social studies, was also deemed appropriate for the participants in sciences, as no differentiation of high school EFL instruction is made for students in either sciences or humanities in Taiwan.

Questionnaire

Two questions were formulated to explore whether the participants had been taught how to read graphics in high school EFL classrooms, and to determine whether or not English was a cause of difficulties in understanding the questions on the test. These are the two questions:

1. In your high school EFL classrooms, did your teachers teach you to interpret graphics similar to those in the test? If your answer is yes, write down the page number on which the same graphics appear.

2. Say whether the English in the items A) was difficult; B) was a little bit difficult; C) was not difficult to read and understand.

Test Procedure

The test was administered in four of the researcher's freshman English classes. In two consecutive periods of 50 minutes, the participants responded to the questions. Participants were told about the purposes of the test. They were encouraged to try their best to answer the questions on the test. They were also told that they were free to discontinue the test at any time. No dictionary was allowed.

Grading

Two college English teachers were invited to hand-grade the measure in a two-week period. The first teacher graded all the questions, added up the total scores, and tallied the number of responses to the questionnaire. The second teacher graded the answer sheets again for accuracy. The raw score was then converted to percent by following this formula: $(\text{total correct answers} \div 72) \times 100$.

Data Analysis

The SPSS version 11.5 was used to organize and analyze the data collected in the study. First reported were the descriptive statistics of the converted scores, and the frequency and percentages using 5% as an interval to depict the passing rates and the general performances. A graph was appended to illustrate the result. Error ratios were then calculated to rank the participants' performances of these graphics: graphs, charts and tables, time lines, table of contents, maps, and indices. Finally, the participants' responses to the questionnaires were reported in percentages. Also reported were the results of a *t*-

test comparing the participants who indicated that the English in the items was or was not difficult to read and understand. Tables were also formulated to visually delineate the results.

Results

Pass Rate and the Distribution of Performances

Table 1 presents the pass rate and the distribution by frequency and percentage of the performances of the 211 participants. The results show that 28 or 13.27% of the participants scored above the 85% of the cut-off criterion. Twenty nine or 13.744% of the participants scored near the passing criterion. The table also shows that 88.625% of the participants were able to answer more than 50% of the questions, with only 11.375% of the subjects out of the range.

Table1: Passing Rate and Distribution by Frequency and Percentage

Exact Range by 5%	Frequency	Percent	Cumulative Percent
95.00~100.0	00	00	00
90.27~90.27	05	02.370	02.370
86.11~88.88	23	10.900	13.270
80.55~83.33	29	13.744	27.014
75.00~79.16	37	17.536	44.550
70.83~73.61	30	14.218	58.768
65.27~69.44	19	09.004	67.772
61.11~63.88	18	08.531	76.303
55.55~59.72	15	07.109	83.412
50.00~54.16	11	05.213	88.625
45.83~48.61	07	03.318	91.943
40.27~44.44	09	04.265	96.208
36.11~38.88	05	02.370	98.578
31.94~34.72	03	01.422	01.422
Total	211	100.0	100.0

Table 2 lists the mean, median, mode, range, minimum, maximum and percentiles of the results. The range is 58.33 from the highest score, 90.27, to the lowest score,

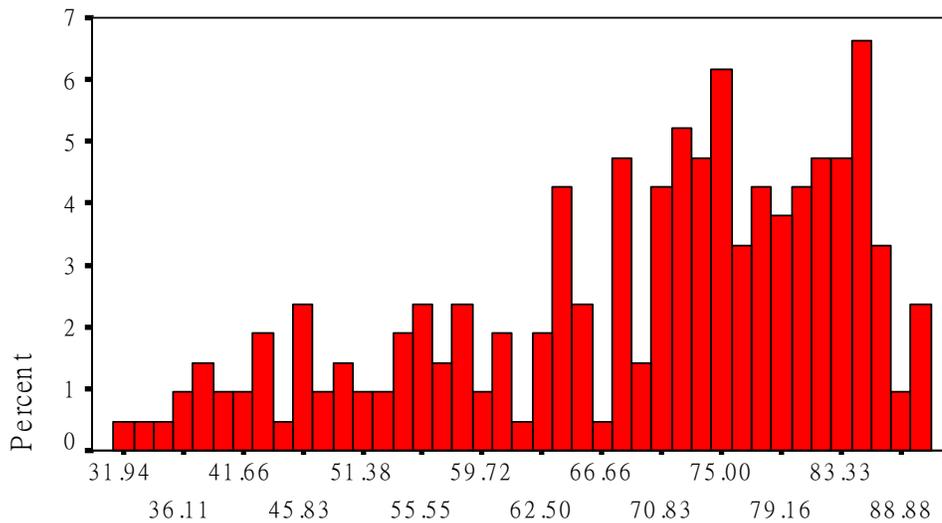
31.94. A score above 80.55% is in the upper 75 percentile while a score lower than 60.11% shows a performance in the lower 25 percentile. The table also indicates a negatively skewed distribution, as the mean (69.56%) is smaller than the median (72.22%) which is also smaller than the mode (86.11%). Graph 1 visually displays the negatively skewed distribution of the overall performance of the 211 participants. The distribution spreads out in a bell shaped curve showing that 50% of the student subjects were able to answer above 72.22% of the questions.

Table 2: Mean, Median, Mode, Range, Minimum, Maximum and Percentiles

N	Valid	211
Mean		69.56
Median		72.22
Mode		86.11
Std. Deviation		14.11
Range		58.33
Minimum		31.94
Maximum		90.27
Percentiles	25	61.11
	50	72.22
	75	80.55

Figure 1: Distributions by Percent

V A R 00001



V A R 00001

Ranking the Ability to Use and Interpret Graphics

These six types of graphics were used in this test: Graphs, Tables and Charts, Time Lines, Tables of Contents, Index, and Maps. To determine which types of graphics the participants were able or unable to use, the error responses to the questions in each of the six categories were tallied and computed separately by this formula: (total number of errors ÷ total number of questions)/100. Table 3 lists the results in descending order. The results show that the participants failed to answer more than 1/3 of the Graph questions, which include bar graphs, pie graphs, and line graphs. In addition, the participants also failed to answer nearly 1/3 of the questions on the following graphics: Index, Table of Contents, Charts and Tables, and Maps. The subjects were able to answer approximately 79% of the time line questions.

Table 3: Ranking of the Ability to Use and Interpret Graphics

Types	Percentage of Error
Graph	35.51
Index	32.74

Table of contents	31.67
Chart & Table	30.25
Map	30.21
Time line	21.33

Experience in Reading Graphics in English in High School

Of the 211 participants, 202 responded to the first survey question. Among these, 112 participants expressed that they had been taught how to use or interpret at least one or more than one of the six graphics in high school EFL classrooms, while the other 90 participants said they had never been taught how to use or interpret the graphics.

This formula was used to tally the percentage for each of the graphics: (total number indicated as having been taught ÷ 202)/100. The data show that among the six graphics, the Graph was the most frequently taught, with a percentage of 31.28; while the Index was the least taught, with a percentage of 6.16. The data also reveal that the least taught graphics were the following: Time Line (18.01%), Charts and Tables (12.32%), Maps (11.37%), and Table of Contents (10.43%). Table 4 shows the percentage of the graphics which participants said had been taught in high school EFL classrooms.

Table 4: Percent of the Graphics Taught in High School EFL Classes

Types	Total Percent of Being Taught
Graph	31.28
Time line	18.01
Chart & Table	12.32
Map	11.37
Table of contents	10.43
Index	06.16

Performances of Participants Rating Levels of English Language Difficulty

Of the 211 participants, 190 responded to the second survey question. The tally of the responses showed the following results: Difficult (78 participants), Somewhat Difficult (11 participants), and Not Difficult (101 participants). The data from the 11 participants were eliminated from the study as the number was small. The two categories, Difficult and Not Difficult, were used for computing whether or not a significant difference existed between the performances of the participants who found the level of English used in the test difficult and the performances of those who said the English was not difficult.

Table 5 shows that the mean performance of the group who considered the language level difficult is 62.05%; whereas the mean for the group who found the language level not difficult is 74.86%. Again, to bring out possible significant differences in the performances, a simple *t*-test was administered. A significant difference was found, $t(177) = 5.933, p < .001$, favoring the Not Difficult Group. Table 6 presents the result of the *t*-test.

Table 5: Mean, Median, Mode, Range, Minimum, Maximum and Percentiles

		Difficult	Not Difficult
N	Valid	78	101
Mean		62.05	74.86
Median		63.88	76.38
Mode		63.88	86.11
Std. Deviation		14.719	11.848
Range		54.17	58.33
Minimum		33.33	31.94
Maximum		87.50	90.27
Percentiles	25	50.00	70.83
	50	63.88	76.38
	75	73.96	83.33

Table 6: Result of the *t*-test by Difficulty × No Difficulty

English	N	Mean	Std. Deviation	<i>t</i>	df	Sig. (2-tailed)
Not difficult	101	74.86	11.848	5.933	177	< .001
Difficult	78	62.05	14.719			
Total	179					

Discussion

A feature of college textbooks is the abundant use of graphics. Graphics facilitate comprehension and the learning of information; therefore, the ability to read and interpret graphics compensates for the difficulties encountered by students in comprehending discipline-specific college texts. In Taiwan, however, there exist no studies that explore the current status of graphic literacy instruction in either high school or college freshman English classrooms. In this study, the ability to read and interpret graphics is measured by administering a criterion-referenced graphic test. The goals are fourfold: 1) to determine whether college freshmen are ready to use text aids and graphics to facilitate text comprehension, 2) to find out which graphics college freshmen are most likely to interpret correctly, 3) to discover whether college freshmen learned to read the graphics in English in high school, and 4) to evaluate to what extent English is a cause of the difficulties experienced by students in reading graphics.

In response to the first research question, the findings of this study reveal that only 13.27% of the participants reached the mastery level. Another 13.74% of the participants scored close to mastery level. The data also show that 50% of the student subjects were able to answer above 72.22% of the questions. The negatively skewed distribution further indicates that the majority of the participants tended to be able to read and interpret the graphics, even though the mastery rate was low.

In response to the second research question, the tally of the error rate reveals that the subjects were able to use, in descending order of competence, the following graphics: Timelines, Maps, Tables and Charts, Index, Table of Contents, and Graphs. In other words, the subjects were more capable of reading and interpreting Timelines than Maps, Tables and Charts, Index, and Table of Contents. They were least competent in reading and interpreting Graphs, which happen to come in greater variety: bar graphs, line graphs, pie charts, bar charts, and pictographs.

In the subjects' identification of the different graphics taught or not taught in high school English classes, the results show that the six graphics tested in the study were generally neglected. Obviously teaching students how to read and interpret graphics is not one of the major foci in high school English classrooms. Scrutinizing the data pertaining to graphics teaching ratio and graphics using error percentages, we observed an interesting incongruity. About 31.28% of the subjects reported that they had read Graphs in high school English classrooms. However, more than 1/3 of the subjects failed to read and interpret the Graph questions correctly. The data further show that the subjects were more successful in answering the Timeline questions, but Timelines were taught much less frequently than Graphs in English classes. Obviously, instruction or the lack of instruction in this case, does not influence the students' ability to read and interpret the graphics. Perhaps the complexity and the types of graphics, as well as the relation between these displays and the task demands and the learner's prior knowledge and cognitive abilities (Schnotz, 2002), do affect the students' success or failure in reading and interpreting graphics. Moreover, as the subjects are EFL learners, it is also possible that the subjects' fluency in English influences their performances.

As regards the last research question, the results show that the subjects who rated

English as Not Difficult outperformed those who rated English as Difficult. The gap between the two groups was large, with a mean difference of 12.81 (74.86 vs. 62.05). This result indicates that many more subjects who rated English as Difficult scored at the lower extremity of the normal curve. The significant difference leads to the conclusion that the ability to read English influences the ability to read and interpret graphics in English.

This study reveals that only 13.27% of the participants reached the mastery level in reading and interpreting graphics in English. A dearth of English proficiency and the ineffective high school EFL education in Taiwan may become one of the major causative factors. EFL education in Taiwan has been putting too much emphasis on rote-learning, and students are just asked to remember words and grammar and pass the tests (Gluck, 2007). Ineffective EFL learning and instruction have also been evidenced in many news reports in Taiwan and in international news media such as BBC News. For example, newspapers have reported that Taiwan ranked poorly in the International English Language Testing System (IELTS) given in 2006 and in 2007 (Gluck, 2007; Shieh, 2008). Taiwan ranked 17th in 2006 and 16th in 2007 among the top 20 Asian countries which had the highest number of people taking the language-proficiency test. Besides, Taiwan ranked the last among Mainland China, Korean, and Singapore in the Test of English as a Foreign Language (TOEFL-CBT) conducted from 2004 to 2006 (Shieh, 2008). In 2007, Taiwan again ranked the last among Mainland China, Korean, Hong Kong, and Singapore in the Test of English as a Foreign Language (TOEFL-iBT) (Hu, 2008; Shieh, 2008).

The factors leading to poor performances in reading and interpreting text adjuncts are numerous; for instance, Westelinck, Valcke, Craene and Kirschner (2005) point out that

individual differences in learning styles or spatial abilities, the demand of the task, time on task, and the quality of the external graphical representations may influence the graphics reading and interpretation. In addition, the learner's prior knowledge and cognitive abilities also affect the students' success or failure in reading and interpreting graphics (Schnotz, 2002). Moreover, it is possible that gender difference can also be one of the factors. In the current study, gender variations in the ability to read and interpret graphic displays were not the major research foci. However, it was observed that males and females did achieve differently. Among the 82 males and 129 females, it was found that the difference of mean between both sexes was 6.66 points (male: 65.494; female: 72.154). The result of *t*-test revealed that females significantly outperformed males, $t(209) = 3.41, p < .001$.

Conclusions

In Taiwan, it has long been a common practice among college subject teachers to assign English content textbooks to their students (Cheng, 1993; Cheng & Hung, 2002; Hu, Chen & Liu, 2008). These college texts are originally written for English speakers who "ideally" should possess at least a Grade 13 reading ability (Singer & Donlan, 1989). The possession of graphic reading skills, though, may compensate for the deficiencies in students' comprehension of the college discipline-specific texts. Nevertheless, as this study reveals, inefficient English ability aggravated by a lack of proficiency in graphic reading skills will deprive the reader of this channel to read to learn from the college content texts. As the findings suggest, the efficacy of reading to learn from these college content area texts in the EFL context is questionable among the majority of the college students in Taiwan.

The ability to read and interpret graphs, both static and animated, should be regarded as one of the basic life skills for students of different grade levels. Nearly three decades ago, graphic literacy educators pointed out that “the ability to read graphics is becoming increasingly important because they are more widely used in newspapers, magazines and textbooks than in days past” (Fry, 1981, p.388). Now, with the dramatic development of computers and the proliferation of internet technology, the ability to read and interpret graphic representations has extended far beyond the static textbook graphic displays. Such graphic reading proficiency has become imperative if we are to function in our daily lives.

Implications for Further Studies

As Bintz (1997, p.12) points out, “I am insecure because as an English teacher, somehow I am expected to know about reading, but at the college level I was only trained in English content.” Many studies, moreover, suggest that pre- and in-service teachers in English as native tongue setting might believe that they are not qualified to teach science or social studies reading to their students (Bintz, 1997). In future studies, researchers may want to pay special attention to EFL teachers. As most high school English teachers are majors in English literature or linguistics, their training in instructional strategy for other disciplines may be deficient. Studies might focus on such topics as attitudes towards teaching graphics in English and knowledge of and competency in interpreting graphics.

As this study is based mainly on quantitative data, to flesh out conclusive information regarding college students’ graphic abilities, it is necessary that qualitative studies should also be designed. Insight is needed for examining the frequency, attitudes, and strategies employed by college students beyond the freshman year in reading and interpreting

graphics in English subject textbooks. In addition, studies should also be conducted using high school students with regards to their knowledge of graphs and subsequent problems in reading graphs. Indeed, it is more important to train high school students to read and interpret graphics before they enter college. Teaching and learning guides can then be suggested to ready high school graduates to read and interpret graphics in English.

Other than differences in learning styles or spatial abilities, prior knowledge, and cognitive abilities in comprehending text adjuncts, gender differences in the ability to read and interpret text adjuncts may prove to be another area of research interest. The existing studies relating to gender variations have been focused on reading, vocabulary, motivations, reading and learning strategies; text-style preferences, cognition, ability in sciences or mathematics (Chan, 1994; Evans, Schweingruber & Stevenson 2002; Halpern, 2000; Hyde & Linn, 1988; Hyde & McKinley, 1997; Morgan & Douglas, 2007; Poole, 2005; Worrell, Roth, & Gabelko, 2007). However, no studies exist that explore gender difference in graphic comprehension. In future studies, differences in comprehending graphic displays between the two sexes should be first established; then, factors contributing to the differences should be explored and delineated.

Finally, the ability of high school and college students to design graphics should also be considered as one of the important skills in English writing class. In college, students are expected to be effective writers, that is, to be able to produce coherent and elegant academic texts that convey clear meaning. Designing and interpreting graphics are indispensable skills for this purpose. As college students are frequently required to write projects and papers either in Mandarin or in English, studies can be designed to explore how they use graphs in their projects and the quality of the graphs. Graphs used in both languages can be examined as well for differences and proficiency in usage. In addition, a

great number of college students are able to design web pages over the internet, though most of them are in Mandarin Chinese. Yet, studies should be developed to compare and contrast the quality of the graphs they design and, more importantly, the problems they meet and the quality of the graphs, tables and indices they design in English.

References

- Arnold, T., & Dewyer, F. (1975). Realism in visualized instruction. *Perceptual and motor skills*, 40, 369-370.
- Askov, E., & Kamm, K. (1982). *Study skills in the content areas*. Boston: Allyn & Bacon.
- Bintz, W. P. (1997). Exploring reading nightmares of middle and secondary school teachers. *Journal of Adolescent & Adult Literacy*, 41, 12-24.
- Bodemer, D., Ploetzner, R., Bruchmuller, K., & Hacker, S. (2005). Supporting learning with interactive multimedia through active integration of representations. *Instructional Science*, 33, 73-95.
- Booher, R. (1975). Relative comprehensibility of pictorial information and printed words in proceduralized instructions. *Human Factors*, 17, 266-277.
- CEEC - College Entrance Examination Center. (2008). Appoint exam profile. Retrieved September 20, 2008. from www.ceec.edu.tw
- Chan, L. K. S. (1994). Relationship of motivation, strategic learning and reading achievement in grades 5, 7 and 9. *Journal of Experimental Education*, 62, 319-339. (Retrieved August 16, 2008. from www.questia.com
- Cheng, T. Y. (1993). The syntactical problems Chinese college students meet in reading English technical textbooks. *Resource in Education*. (ERIC Document Reproduction Service No. ED 364 094).

- Cheng, T. Y. & Hung, T. I. (2002). College content teachers' attitudes toward English reading instruction in the content areas in an EFL environment. *Journal of Hsing Kuo College of Management*, 1, 25-38.
- DeStefano D., & LeFevre, J-A. (2007). Cognitive load in hypertext reading: A review. *Computers in Human Behavior*, 23, 1616–1641.
- Dillner, M., & Olson, J. (1982). *Personalizing reading instruction in middle, junior and senior high schools*. New York: Macmillan.
- Evans, M. E., Schweingruber, H., & Stevenson, H. W. (2002). Gender differences in interest and knowledge acquisition: The United States, Taiwan, and Japan. *Sex Roles*, 47, 153-167.
- Fry, E. (1981). Graphical literacy. *Journal of Reading*, 24, 383-390.
- Gillespie, S. C. (1988). *The development of a criterion-referenced test to assess the ability of secondary social studies students to read and interpret selected graphics*. Ann Arbor, MI: U.M.I.
- Gluck, C. (2007). *English village open in Taiwan*. BBC News. Retrieved 13 September, 2007 from <http://news.bbc.co.uk/1/hi/education/6992823.stm>
- Gronlund, N. (1982). *Constructing achievement tests* (3rd Ed.). Englewood Cliffs, NJ: Prentice-Hall.
- Halpern, D. F. (2000). *Sex differences in cognitive abilities* (3rd Ed.). Mahwah, NJ.: Lawrence Erlbaum Associates.
- Heilman, A., Blair, T., & Rupley, W. (1981). *Principles and practices of teaching reading*. Columbus, OH: Charles E. Merrill.
- Hu, C. F. (2008). *Taiwan's test scores in English proficiency were compared with those of South Korea, Hong Kong and Singapore. Taiwan's scores were ranked the lowest*

- (in Mandarin Chinese). The Liberty Times. Retrieved October 10, 2008. From tw.news.yahoo.com/article/url/d/a/080920/78/168qb.html.
- Hu, C. H., Chen, S. Y., & Liu, L. R. (2008). In Taiwan, 10% of college subject classes are taught entirely in English (in Mandarin Chinese). *The Liberty Times*. Retrieved November 20, 2008: URL: [www.libertytimes.com.tw /2008/ new/ sep/15/ today-education1.htm](http://www.libertytimes.com.tw/2008/new/sep/15/today-education1.htm)
- Hunter, B., Crismore, A. & Pearson, P. (1987). Visual displays in basal readers and social studies textbooks. In D. Willows & H. Houghton (Eds.), *The psychology of Illustration: Volume 1*, (pp. 116~135). New York: Springer-Verlas.
- Hyde, J. S., & Linn, M. C. (1988). Gender differences in verbal ability: A meta-analysis. *Psychological Bulletin*, 104, 53-69.
- Hyde, J. S., & McKinley, N. M. (1997). Gender differences in cognition: Results from meta-analyses. In M. Crawford, J. S. Hyde, & J. T. E. Richardson (Eds.), *Gender differences in human cognition*, (pp.30~51). New York: Oxford University Press.
- Jaeger, R. (1979). Measurement consequences of selected standard-setting models. In M. Bunda and J. Sanders (Eds.), *Practices and problems in competency-based measurement* (pp. 48~58). Washington, DC: National Council on Measurement in Education.
- Lowe, R. K. (1993). Constructing a mental representation from an abstract technical diagram. *Learning Instruction*, 3, 157-179.
- Mateja, J., & Wood, K. (1982). Students' knowledge of textbook metastructure. In G. McNinch (Ed.), *Reading in the Disciplines: Second yearbook of the American reading forum* (pp. 97~100). Carrolton, GA: Thomason Printing.

- Mokros, J. R., and Tinker, R. F. (1987). The impact of microcomputer based labs on children's ability to interpret graphs. *Journal of Research in Science Teaching*, 24, 369–383.
- Montali, J. & Lewandowski, L. (1996). Bimodal reading: Benefits of a talking computer for average and less skilled readers. *Journal of Learning Disabilities*, 29, 271-279.
- Morgan, P. L., & Douglas, F. (2007). Is there a bidirectional relationship between children's reading skills and reading motivation? *Exceptional Children*, 73. Retrieved July 22, 2008 from www.questia.com
- Poole, A. (2005). Gender Differences in reading strategy use among ESL college students. *Journal of College Reading and Learning*. 36, 7-20. Retrieved July 22, 2008 from www.questia.com
- Schnotz, W. (2002). Towards an integrated view of learning from text and visual displays. *Educational Psychology Review*, 14, 101-120.
- Shieh, S. R. (2008). Inefficient English learning and instruction in Taiwan. Taiwan's English scores were ranked the lowest (in Mandarin Chinese). *United Evening News*. Retrieved February 10, 2009 from www.young-mint.com.tw/news.php?SerNo=556
- Singer, H., & Donlan, D. (1989). *Reading and learning from text*. Hillsdale, NJ.: Lawrence Erlbaum Associates.
- Thomas, K., & Moorman, G. (1983). *Designing reading programs*. Dubuque, IA: Kendall/Hunt.
- Tyler, R. (1973). Testing for accountability. In A. Ornstein (Ed.), *Accountability for teachers and school administrators*. Belmont, CA: Fearndon.
- Wainer, H. (1994). Understanding graphs and tables. *Educational Researcher*, 22, 14-23.

- Weidenmann, B. (1989). When good pictures fail: An information-processing approach to the effects of illustrations. In Mandl, H., and Levin, J. R. (Eds.), *Knowledge acquisition from text and pictures*, (pp. 157~170). North-Holland: Amsterdam.
- Weintraub, S. (1967). What research says to the reading teachers. *Reading Teacher*, 20, 345-349.
- Westelinck, K. D., Valcke, M., Craene, B. D., & Kirschner, P. (2005). Multimedia learning in social sciences: Limitations of external graphical representations. *Computers in Human Behavior*, 21, 555–573.
- Worrell, F. C., Roth, D. A., & Gabelko, N. H. (2007). Elementary reading attitude survey (ERAS) scores in academically talented students. *Roepers Review*, 29, 119-124.
- Retrieved August 16, 2008 from www.questia.com

Acknowledgements

The author would like to express his gratitude to Dr. Cindy S. Gillespie, now Cindy S. Hendricks, a professor and the director of the School of Teaching and Learning at Bowling Green State University, for granting me permission to use *Graphics Assessment for Social Studies*.

How English L2 Learners in China Perceive and Interpret Novel English Compounds

Xiaozhou Zhou

University of Warwick, UK

rola_bati@hotmail.com

Dr Victoria Murphy

University of Oxford, UK

Bio Data:

Xiaozhou Zhou received her Master's Degree in Applied Linguistics and Second Language Acquisition in University of Oxford. She is now a third-year-PhD student at the University of Warwick. This paper is based on the research conducted during her master's studies in Oxford.

Victoria Murphy received her PhD from McGill University in 2000 and is now University Lecturer in Applied Linguistics in the Department of Education, University of Oxford, and fellow of Kellogg College. Her research interests focus on aspects of lexical and literacy development in L2 learners, primarily with children.

Abstract

This study aims to investigate how English learners in China interpret novel English noun-noun compounds. Relevant research literature is for the most part limited to L1 children's interpretations of noun-noun compounds. Therefore the current study extends the research area into the L2 domain with a view to comparing interpretations of L2

learners with those of L1 children in the study of Krott, Gagné and Nicoladis (2008). Fifty-two students from two universities in Shanghai, China participated in the research. They were given an English compound test consisting of 30 novel noun-noun compounds. The results indicated that, generally speaking, the participants displayed an overall competence in understanding and interpreting novel English compounds. The differences in performance success between the intermediate and advanced groups were not found to be statistically significant. Possible factors affecting participants' interpretations and vocabulary teaching implications are discussed at the end of the study.

Keywords: noun-noun compounds, Chinese learners, vocabulary teaching, L1 transfer

Introduction

The acquisition of vocabulary has assumed an important, not to say crucial, role in second language learning theory over the past decade (Lewis, 1993). Within the area of vocabulary acquisition, one type of vocabulary merits special focus – noun-noun compounds. On the surface, these are simple constructions composed of two parallel nouns. However, they are not semantically parallel since there is always one type of (semantic) relation connecting the two constituents. Understanding a noun-noun compound thus involves knowing the two nouns (known as the modifier and the head) and selecting an appropriate relation type. Take “paper salad” as an example: in this compound, “paper” is the modifier and “salad” is the head. One likely relation between these two constituents might be “made of”, meaning “salad which is made of paper”. Previous research in this area has investigated how English native-speaking children interpret English noun-noun compounds.

As a compounding language, Mandarin Chinese has a much more complex system of constructing noun-noun compounds. Unlike English, it is unnecessary to analyse a Chinese noun-noun compound by segmenting the two characters, because Mandarin emphasizes meaning more than form. Therefore, if people understand the meanings of the

separate characters, they will naturally obtain the meaning of the compound. There are vast differences between the natures of English and Mandarin which are very likely to cause different understandings of English noun-noun compounds by Chinese people compared to those of English native speakers. A clearer understanding of how Chinese learners interpret English noun-noun compounds, together with a comparison of results obtained from Chinese participants to those gained from previous research conducted on L1 children will hopefully be of use to learners and teachers alike.

The current research is a descriptive study in which quantitative methods are employed. It is an extension of the research conducted by Krott et al (2008). Krott et al investigated L1 children and adults' interpretations of novel English noun-noun compounds. The findings suggest that there are general similarities but some specific differences in the strategies used by children and adults when interpreting novel compounds.

Little research has so far been conducted concerning English L2 learners in the compounding area, especially involving learners whose native language is also a compounding language. It will thus be innovative as well as meaningful to extend the study of Krott et al (2008) into the Chinese EFL domain to investigate how L2 learners in China understand and interpret English novel compounds and compare their interpretations with those of L1 children.

General Discussion on L2 Vocabulary Acquisition

N. Ellis (1997) argued that learning a new word involves more than understanding a novel sound pattern and recognizing and writing its orthographic pattern. It also requires learning about its syntactic properties, its place in lexical structures, its semantic and referential properties etc. However, these different types of vocabulary acquisition are all

underpinned by two distinct types of learning mechanism: learning a word's form and learning its meaning (N. Ellis, 1994). Unlike L1 children who acquire their native language in a more *natural* and *implicit* way, L2 adults have already developed rich conceptual and semantic systems linked to their L1. For them, learning a new word is a process of mapping: either "mapping the word form onto pre-existing conceptual meanings or onto L1 translation equivalents" (N. Ellis, 1997, p.134). During this mapping process, various references and collocations of that new word are likely to be neglected. Since the translation equivalence is often made explicitly at the early stage of learning, it is very easy for the learner to immediately build up an association with an L1 word. Thus, many of the equivalences set up at that time have to be modified during later learning processes (Ringbom, 1986).

L2 vocabulary development is also influenced by the organization of the mental lexicon. It has been argued that human memory is very flexible and can "process" a large quantity of data, provided that it is systematically organized (Takač, 2008). To gather data on the organization and functioning of the mental lexicon, researchers have been studying various speaker behaviors such as tip-of-the-tongue phenomena and people suffering from aphasia (Aitchison, 1990). Recently, the dynamic characteristic of the mental lexicon has become more and more prominent in second language learning research. The dynamic characteristics in the mental lexicon can be reflected in the concept of *spreading activation* (Hulstijn, 2000), in which two lexical items which are stored without any interconnection can be linked via some formal or semantic features. Noun-noun compounds are good examples of spreading activation. During the process of spreading activation information on existing lexical items is expanded and completed. For example, the compound *banana boat* was a combination of two previously independent nouns but

can be explained as the boat for shipping bananas or a popular snack made of bananas and resembling a boat. It therefore follows that research into noun-noun compounds is helpful in providing us with further insights into the mental lexicon. So far, studies in noun-noun compounds have been limited to predominately native speakers of English. The current study thus extends the scope of previous research by looking at the L2 domain, in the hope that the research findings will contribute to further understanding of the L2 mental lexicon.

The fact that the native language has a considerable effect on the way a second language is learnt and used has been confirmed more and more strongly by research (Kellerman, 1984; Perdue, 1993). Errors in second language learning such as *interlingual* confusion are believed to be caused by interference and transfer from the native language (Swan, 1997, p.161). Therefore in terms of second language vocabulary acquisition, a key problem is to identify what sort of effect a learner's native language will have on their acquisition of a second language. Will the native language help, hinder or exercise no effect?

The influence of the native language is not just limited to language itself but also can spread to the broader cultural sphere. If the native and target languages do not have much in common in the cultural domain, there may be little overlap between the concepts they express. This can be far more frustrating for learners than learning a related foreign language.

In second language vocabulary acquisition, the L1 can either facilitate or create obstacles during the process of recalling and using previously learnt words. This also applies to the construction of a complex lexical item that has not been previously used as a unit (Takač, 2008). Therefore, this current study is also exploring, though not directly

investigating, if the Chinese learners' L1 is exerting any influence on their understanding of English compounds. First of all, it is necessary to know what English noun-noun compounds are and what involves an understanding of a noun-noun compound.

An Overview on Research in Compounding Area

Research into the area of compounds has been developing steadily over a long period as researchers explore various theories. In the L1 domain, some studies investigated the strategies used by children when understanding and producing novel compounds. Murphy and Nicoladis (2006) raised and tested the hypothesis that the frequency of different complex lexical forms to which children were exposed influences their development in terms of interpreting those forms. In addition, researchers were interested in the native speakers' inference process in terms of novel compounds, i.e. what knowledge they are using when interpreting these compounds.

With regard to the L2 domain, researchers applied compounding to the investigation of underlying representation of language in L2 inflectional morphology. Results from the research of Murphy (2000) showed a dissociation between regular and irregular morphology as manifested in the compounding task in the study, which indicated that the level-ordering model (Kiparsky, 1983), which described ordering morphological processes at different levels, might not be relevant to second language morphological acquisition. However, the dual-mechanism model (Pinker and Prince, 1988, 1992), which believed that there were two distinct systems for language learning – one for processing regulars and one for irregulars did not adequately explain the dissociation phenomenon in this study either.

In terms of the process of L1 children acquiring noun-noun compounds, a usage-based

theory of language acquisition has been evidenced by research such as that of Krott and Nicoladis (2005). This theory posited that children initially acquire and use a compound as a whole unit without knowledge of the underlying linguistic structure. With the acquisition of more and more words, children tended to make conservative generalizations about the constructions of the particular words they heard and used (Tomasallo, 1992, 2000). Furthermore, “the higher the number of constructions in which those particular words are used, the more likely children are to generalize the use of those words in similar constructions” (Krott & Nicoladis, 2007, p.205). Consequently, it could be argued that if children were influenced by the family³ size when interpreting noun-noun compounds, it was likely that they were using analogies with other compounds in their vocabularies to interpret the one before them (Krott & Nicoladis, 2007). In other words, they generated the rules of the linguistic features of compounding by deriving meanings from existing and familiar compounds.

In addition, researchers are interested in factors that play a role in children’s acquisition of the general principles of compounding. These factors can be learned through comparing compound acquisition across languages and language families (Krott & Nicoladis, 2007). Cross-linguistic studies have shown that productivity and frequency of compounds are crucial to children’s acquisition of compounds, and in languages in which compounding is highly productive, such as English, Chinese and Dutch, children begin to coin novel noun-noun compounds at a very early age (Becker, 1994; Clark, 1993, 2003). The current research therefore aims to investigate how Chinese learners of English,

³ A family refers to the set of compounds that share the same constituent. e.g. The modifier family for “chocolate” includes compounds sharing the form “chocolate X” such as “chocolate cake”, “chocolate milk” and “chocolate muffin” and the head family for “cake” consists of compounds of the form “X cake” such as “cheese cake” and “fruit cake”.

whose first and second languages are both compounding languages, interpret novel English noun-noun compounds in order to view whether or not L2 learners acquire and understand the compounds in the same way as L1 children do, and if not, to what extent they differ.

Introducing the Original Paper Krott et al (2008)

The Krott et al (2008) aimed to investigate which factors affect children's selection of thematic relations when interpreting noun-noun compounds. The researchers started by introducing compounds and explaining what makes the understanding of novel compounds challenging. Descriptions and illustrations of the semantic structure of English noun-noun compounds, the definitions of the "modifier" and the "head" and possible thematic relations between these two roles were provided. It was argued that interpreting novel compounds relies largely on real-world knowledge. The researchers argued that real-world knowledge includes how constituents within a compound have been used in other compounds. This seems to be particularly relevant to children's understanding of compounds.

A usage-based framework assumes that children start to learn words in specific strings before generalizing to an abstract pattern. It suggests that children learn compound words initially as individual words, with no knowledge of their internal structure. Further evidence also shows that parsing compounds as chunks lasts until early school years, and probably even adulthood.

However, results from other studies indicate that children are able to understand the sub-categorization function of compounds in their early years. They can even create

novel compounds on their own.

Researchers seem to be more convinced by the usage-based account in which children start by having no knowledge of the internal structures but gradually build up patterns related to modifier or head families. It is believed that at some stage children are able to use the modifier family of one word to predict the possible relations in other compounds containing this word.

In this study, 27 L1 children aged 4;9 to 5;8 took part in the experiment. In addition, 36 adults who were all native speakers of English aged 17 to 26 were recruited as participants. Thirty novel noun-noun compounds were constructed for the children. Adult participants were asked to write down interpretations for these 30 compounds and the relations which dominated their responses were selected as the most likely interpretations.

The results from Krott et al (2008) showed that 39.7% of all responses from children and 70% from adults were the dominant interpretation. In addition, in 18.5% of the responses, children reversed the roles of modifiers and heads in their explanations.

In terms of effects of modifier and head families, the research findings indicated that children's interpretations were affected primarily by the relational bias in head families. Supportive head families were likely to result in dominant interpretations while supportive modifier families seemed not to have salient effect on their responses. On the contrary, adults based their interpretations primarily on their knowledge of modifiers in familiar combinations.

There are two salient figures in the findings of this research: "success rate with dominant relations" and "usage of non-dominant relation". The former one measured the extent to which they were successful with one type of relations. "Successful" in this context meant choosing the relation type which was chosen by L1 adults as the dominant relation in Krott et al (2008). The latter referred to cases in which their choices of relation

types were inconsistent with dominant ones.

Through comparing “success rate with dominant relations” and distribution of the most frequent relations of compounds in the CHILDES database, it was found that that children’s success in responding with dominant relations was not affected by the overall frequency of those relations in the database. In other words, finding a relation more often in their vocabulary did not necessarily mean they would understand it better, thus did not guarantee a higher success rate when interpreting compounds with this relation.

With regard to relation preference for children and adults, taking both *success rate with dominant relation* and *usage if non-dominant relation* together into consideration, the results indicated that children had a higher success rate for relation HAS than for the FOR and MADE OF relations. Adults seemed to have similar success rates for HAS, MADE OF and FOR. Furthermore, children overused the relation HAS but underused FOR and MADE OF.

This research concluded with a suggestion that a usage-based framework was adequate for describing children’s development in interpreting novel compounds. However, the fact that five-year-olds still made mistakes when interpreting compounds challenged the previous assumption that two- and three-year-olds had developed a general noun-noun schema: this further demonstrated that children’s general knowledge about compounds was developed on an item-by-item basis and at a fairly slow pace. There were also a number of possible variables that need to be addressed in future research.

The Krott et al (2008) study extended the approach of usage-based theories to investigating children’s understanding of novel English compounds and demonstrates that a usage-based framework is applicable for this investigation. It deepened the research in compounding area with a combination of using L1 children and novel English

compounds, which was an innovation in this research area. In addition, the methodology of this research and the results generated from it served as a major contribution for future replication and extension.

This Krott et al (2008) research prompted me, the author, to design a similar study of L2 speakers, since previous research into this area has only focused on L1 learners, especially children. It would therefore seem potentially beneficial to extend and develop these research findings by investigating how L2 learners whose native language is also a compounding language approach novel compounds. The key differences in the ways in which compounds are constructed in Chinese and English can lead to errors in comprehension when Mandarin L1 speakers are presented with compounds in English. Thus, an extension of Krott et al (2008) is conducted with English L2 speakers in order to investigate how they interpret English novel compounds and to what extent their interpretations differ from those of L1 children in Krott et al (2008).

One focus is gaining a general and clear idea of the how L2 speakers interpret novel English compounds by asking them to write down their understandings towards the compound items selected from Krott et al (2008). It is hoped that through reading their answers we would know whether or not they understand the basic principle of the English compounds' structure and whether or not they can generate an appropriate pattern when interpreting these compounds. Speculations on the interpretations of current participants will be given on the possible factors that could have been affecting their answers. The other focus is on comparison – between L1 children in the study of Krott et al (2008) and L2 adults in the current research. The findings and evidence will be searched and further explained to support the usage-based approach to the acquisition of compounding, as that in Krott et al (2008).

Vocabulary acquisition is regarded as one of the most crucial processes in learning a second language. English L2 learners are often criticized for using unhelpful strategies when learning vocabulary. Therefore, it is hoped that this research will have implications for vocabulary learning and teaching, particularly with regard to the acquisition of compounds by English learners in China.

Research Questions

This is an exploratory study in which descriptive statistics are presented to demonstrate how they interpret these compounds. Two groups (intermediate and advanced groups) of participants were asked to complete a language background questionnaire and an English compound test. There are two research questions:

- 1) How do Chinese learners of English interpret novel English compounds?
- 2) To what extent are these interpretations different from the way in which native-speaking children interpret novel English noun-noun compounds?

It is hoped that the answers to these questions could provide implications to L2 vocabulary teaching in classroom.

Methodology

The sample consisted of 27 first year undergraduates majoring in science and 25 first year graduate students majoring in English Language and Literature from two different universities in Shanghai, China. They were cast into two different groups according to their previous English proficiency tests results.

As no research has been carried out on how Chinese learners of English interpret compounds like this, a pilot study was conducted beforehand with three Chinese graduate

students studying at Oxford University. On average, they achieved 60% of the total possible mark. A feedback session was conducted afterwards and some amendments were made based on their reports.

The research was first conducted with the intermediate group. The time limit for completing the test was 30 minutes. The same procedures were repeated with the advanced group a week later during a class session. Both data collection processes ensured participant anonymity.

Data Analyses

The analyses of the English compound test started with firstly coding all the answers that participants wrote on their test paper. Each answer for each compound included two parts: the interpretation and the relation. The interpretation referred to their understandings of the meaning of the compound and the relation referred to their understandings of the specific relation type between the two words in this compound. Therefore, the analyses of the English compound test fell into two sections: analyses of their interpretations as well as their choices of relations. Both interpretations and relations were coded before further analyses were conducted. In the process of coding, due to the limited English proficiency levels of the participants, for one compound, there could be various expressions containing the same meaning. Therefore, the researcher figured out the meaning underlying their expressions and categorised them.

In terms of the analyses of the interpretations, the frequencies of responses which were dominant interpretations were calculated. Dominant interpretations were those consistent with the interpretations that L1 adults decoded in the original paper. Calculations were conducted with 52 participants altogether first, and two groups respectively. In addition,

in order to view whether or not there were significant differences between performances of two groups, each interpretation of each participant was coded into either consistent or inconsistent with the dominant one and a chi-square test was administered on the sum of these codes.

With regard to the analyses of their choices of relations, firstly, the average frequency of each type of relations was calculated, followed by calculations of two figures “success rate with dominant relations” and “usage if non-dominant relations”. “Success rate” measured to what extent they were successful with one type of relation. Comparison was made on these two figures between L2 adults and L1 children.

Results

The following table shows the interpretations that were made by most of the participants among altogether 52 of them, i.e. among all the different interpretations for one particular compound, this interpretation in the table was the most popular one. (Please refer to Table 1 at the end of the text.)

Among the 30 items collected from all the 52 participants, 21 of the highest-frequency-answers are consistent with the dominant interpretations which referred to interpretations that most L1 adults provided in Krott et al (2008), whereas within the remaining 9, the majority of the participants produced interpretations that differ from the dominant ones. Expressed as a percentage of all responses, 59.87% are dominant interpretations.

Table 2 shows the average frequency of the choice of each type of relation. Only the most frequently chosen relations, which refer to those relations that appear in the answers of more than one item, are listed here. (Please refer to Table 2 at the end of the text.)

The range of the average relation frequency varies from 2.5% to 61.5%. Among the 18

types of relations in the table, the relation “during” enjoys the highest frequency (61.5%), followed by “use” and “part” (both 59.6%), whereas the relations “from” (2.5%) and “of” (7.7%) are among the least frequent ones.

Considering the participants from the two groups of L2 learners as a whole, it can be seen that L2 adults display an overall competence in understanding and interpreting novel English compounds. The data show that, generally speaking, they are able to distinguish the “modifier” from the “head”. This is demonstrated by their use of the pattern “B is ...A” for an A + B noun-noun compound, indicating that they have developed a sensitivity to the roles of modifiers and heads.

Their responses, which are inconsistent with the dominant interpretations in the original paper (Krott et al, 2008), can be divided into two patterns. On the one hand, there were a small number of cases in which the participants had difficulty in identifying which noun was the “modifier” or the “head” and in interpreting the compound in a way such as “A is ...B”. On the other hand, they occasionally used the word’s verb form (provided that it has a verb form) or mistook the noun for a verb in their interpretation of the compound.

Of all responses in the original research, 39.7% were dominant interpretations made by L1 children while in the current study, 52 L2 adults achieved 59.87% on the same scale. Overall, L2 adults thus performed better than L1 children.

In addition, the comparison between L2 adults and L1 children is statistically focused on their relation preference. The results for L1 children are taken from the Krott et al (2008) with the view of extending this compounding research area into the L2 adult domain. The following table provides a comparison of the “success rate with dominant relations” and “usage if non-dominant relation” between L2 adults (all the participants from the current study) and L1 children. In this context, “successful” refers to participants

being able to provide a relation type which was consistent with the answer that was decoded as dominant relation in the original paper. If they chose a relation type which was not the dominant one for that compound, this choice would be regarded as “usage if non-dominant relation”.

From Table 3, (Please refer to Table 3 at the end of the text.) it can be seen that L2 adults are quite successful with the relations “for” and “located” while less successful with “has” and “be”. On the contrary, L1 children appear to be successful with the relations “has” and “for” but less successful with “made of”.

In addition, the data show that L2 adults tend to use the relations “made of” and “during” more often than L1 children when producing an unexpected response, whereas they tend to use “has” and “located” less frequently than L2 adults. It should be noted that these comparisons, however, are descriptive only as statistical comparisons were not possible since the data from the L1 children were not part of the present study.

Discussion

Comparison of L2 adults and L1 children

The data from the Results showed that L2 adults outperformed L1 children in interpreting novel English compounds. In other words, they interpreted the compounds more closely to the dominant interpretations than the L1 children. However, although L1 children understood that modifier and head nouns play different roles; “they show weak signs of using knowledge about the modification that the modifier can create when they determine how a modifier and a head should be linked (Krott et al, 2008, p.30).” On the contrary, generally speaking, L2 adults demonstrated their abilities in producing an appropriate relation between a modifier and a head.

In Krott et al (2008), it was mentioned that in 18.5% of the responses L1 children reversed the roles of modifiers and heads in their explanations, suggesting that “even five-year-olds still have some problems to distinguish between heads and modifiers when exposed to novel compounds”. (Krott et al, 2008:21) This phenomenon existed for L2 adults too but with a much lower frequency: 2.3%. Although there is no evidence to determine whether or not this is coincidence, it is more reasonable to conclude that the reasons behind this phenomenon differ between L1 children and L2 adult learners. One possible explanation could be that the cognitive ability of a five-year-old child is not sufficiently developed to distinguish modifiers and heads correctly every time. On the other hand, L1 transfer might explain the reversing behaviour of L2 learners, since obviously the cognitive ability of an intermediate L2 learner is already sufficiently developed to enable this structure to be understood if explained directly and clearly. However, if they have never previously been requested to complete a specialized compound test before or verbalize their interpretations in such a grammatically precise way, they will have few guidelines to govern their inference processes. More details on L1 transfer are provided in the next section.

Plausible Reasons for Participants’ Interpretations

L1 transfer

It was generally known that in terms of vocabulary acquisition, the native language has a considerable effect on how a second language is learnt and used (Kellerman, 1984; Perdue, 1993). It is generally believed that if second language learners are more aware of the similarities and differences between their L1 and L2, it is more likely for them to find effective vocabulary learning strategies. Nevertheless, for those learners whose L1 is

closely related to the target language, initially mapping a new word to a counterpart in the mother tongue does not guarantee full understanding of that word, since various collocations and references are missing during the process of mapping (Ringbom, 1986). On the other hand, as for those in whose L1 barely any cognates can be found in relation to L2, it is much more difficult for them to build an association with not only the new word but also the cultural concepts (Swan, 1997).

In the current study, the data suggests that participants' L1 exerts three significant effects on their perception of novel compounds: 1) participants reverse the roles of modifiers and heads 2) participants try to identify a deeper metaphoric meaning beneath the surface of the compound 3) participants perceive two words as coordinate/parallel relations when having difficulty in obtaining the meaning.

Logically, participants from the intermediate group should already be equipped with some strategies for accessing the meaning of English compounds, (i.e.) they should be familiar with the basic concepts and positions of “modifier” and “head”. However, participants' interpretations indicate that they sometimes reverse the roles of “modifier” and “head” especially in compounds composed of more challenging lexical items. Here are a few examples of this tendency to misinterpret the compounds:

- ✧ paper salad – a mixture of papers / paper for salad / paper which looks like salad
- ✧ windmill – wind from the mill / wind that works for mill
- ✧ honeybee – honey made by bee
- ✧ bubble gum – bubble made by gum / bubble of gum
- ✧ light bulb – a light with a bulb
- ✧ cardboard box – cardboard in box
- ✧ rabbit-hole – rabbits staying in a hole

Some responses indicated a general concordance with the dominant interpretation in a general sense, but with the “modifier” being headed. For example:

- ✧ book basket – book stored in basket
- ✧ daylight – day has light
- ✧ cheese burger – cheese in the burger

In these examples, the interpretations provided do not affect or contradict the meanings of these compounds: the core meaning was understood but there was a failure to identify positions of the “modifier” and the “head” with accuracy.

This phenomenon suggests that compounds in which the two constituents are high-frequency words such as lemon juice, picnic table, banana boat, were relatively easy for the participants to parse. Conversely, with respect to compounds containing low frequency words (e.g. bulb, cardboard, mill), participants negotiated an interpretation which was the combination of the meanings of the two constituents regardless of the relation between them.

A certain number of Chinese noun-noun compounds consist of two morphemes/words with coordinate relations. This is a result of the formation of certain ancient Chinese words in which a synonym or an antonym was added to an existing one-syllable-morpheme. Hence, parsing an English noun-noun compound by breaking it down into two parallel separate words seems to be the easiest solution when participants are faced with completely unfamiliar compounds. The following are examples of this tendency:

- ✧ fairy story – story which is fairy (It seems that the participant did not know the exact meaning of the word “fairy”.)

- ✧ light bulb – bulb which is light
- ✧ jellyfish – fish with the name of jelly

It can be seen from the examples that participants resorted to viewing the “modifier” and the “head” as the same thing with different names when they had difficulties in understanding the meaning as a whole. If asked for the definition of each constituent, they might be able to give correct answers. However, when two constituents are put together, the direct and strong influence of Chinese has impelled them to regard them as coordinate parts.

Confusion of grammatical function

Although the researcher made it very clear at the beginning of the test that all the words in the test were noun-noun compounds, there were still a few cases in which the participants misinterpreted the grammatical function of the constituents of the compounds. The following examples illustrate this:

- ✧ April fool – fool people on April 1st
- ✧ snakebite – bitten by snake
- ✧ haircut – cut the hair

In these examples, the responses indicate that the “head” is interpreted as a verb. In fact, the words “fool”, “bite” and “cut” can be verbs or nouns, but in this test it was made clear that they were nouns: the answers generally take the form of using the first noun to

modify the second⁴. However, the above answers reveal that the participants have more or less understood the meaning of the compound. However, in the following examples, incorrect interpretations can be noted:

- ✧ baby bear – give birth to new life
- ✧ apple peel – to peel the apple

In the first example, it is obvious that the participant confused the noun “bear” and the verb “bear”. Furthermore, the order of the constituents in the compound seemed to be reversed as well. It is likely that he / she interpreted it as “bear (v.) a baby”. Similarly, in the second example, the noun “peel” which means the protective outer layer of fruits was mistakenly understood as the verb “peel” meaning the action of “removing the outer layer”. In other words, a noun-noun compound was interpreted as a “verb+noun” phrase.

Lack of environmental and cultural exposure

Knowing a word does not just mean knowing its semantic meaning: the learner needs to be familiar with the pronunciation, spelling, collocations etc. Moreover, language is inextricably linked to the culture and society of the target country. The assembly of simple letters or characters not only maps with the superficial meaning of the word itself but also with the different types of culture of the country where this target language is spoken (Swan, 1997). Sometimes the word has a specific meaning in a certain geographical area: for example, if an English learner in China who has never been to the UK or other English- speaking countries is asked what “egg cup” means, it is very

⁴ i.e. a fool on April 1st, a bite caused by snake; a cut done to the hair

unlikely that the learner will be able to guess the meaning as this compound is not generally taught to English learners in China. On the contrary, a Chinese child who has been living in the UK is far more likely to have encountered this compound. In China, egg cups are not familiar items. Unfamiliarity with the environment, culture and lifestyle in English-speaking countries resulted in participants misinterpreting some compounds. Here are some further examples:

- ✧ cheese burger – burger made of cheese
- ✧ fruit cake – cake made from / of fruit

In Mandarin, the counterparts of “cheese burger” and “fruit cake” are both borrowed words from English as they are not original Chinese food items. Without experiencing these things that belong to the culture of the target language, it is unlikely to provide accurate interpretations for them.

Misunderstandings caused by limited English proficiency levels

These misunderstandings have different possible causes: some are due to unfamiliarity with the vocabulary while others might be due to lack of exposure to the language. Let us consider the following responses:

- ✧ side ache – ache accompanied by a syndrome
- ✧ apple peel – tool to peel apple
- ✧ haircut – tool used to cut the hair

It is very likely that in the first example, the participant misunderstood the compound “side ache” by mentally comparing it to “side effect”. Therefore, he / she wrote down

“ache accompanied by a syndrome” which is similar to the definition of “side effect”. However it is not possible that he / she did not know the two words “side” and “ache” – “ache” was even referred to in the answer. One plausible explanation could be that the two different compounds (“side ache” and “side effect”) were stored in the same place in the participant’s mental lexicon. Looking at “side ache” triggered the recall of “side effect”. Or the meaning “ache accompanied by a syndrome” was matched incorrectly with the term “side ache” in his / her mind.

In the second example, the participant simply mistook “peel” for “peeler”. The third example is less easy to explain. Both constituents are very basic lexical items – maybe the participant is familiar with “to have one’s hair cut”, but not with haircut as a compound.

Limitations and Suggestions for Future Research

One limitation in the current study is that there is no clear and definite “dominant interpretation” for L2 learners. The “dominant interpretation” gained from Krott et al (2008) was relevant to only L1 children but applied to L2 adults as well. However, the English input that the L2 adults in the current study are exposed to is qualitatively different from the input an L1 child is exposed to. Therefore, what constitutes a “dominant interpretation” for an L1 child may not be the dominant interpretation for L2 adults.

Insights Provided for Vocabulary Learning and Teaching

Some of the answers participants provided (e.g. failing to explain “pillow-fight” and mistaking “side ache” for “side effect”) show that to some extent they lack authentic L2

input of English noun-noun compounds. Especially for those intermediate learners whose major course is not English, the opportunity to encounter new words is only limited to text books. However there are texts rewritten or modified by Chinese people from original copies, and if they are extracted from the original books or articles, they are out of date already. It is thus suggested that learners should find more authentic and updated learning materials for themselves such as listening to English news on the radio, reading newspapers and magazines online, watching recent English films. Only by doing so can they have a good chance of being exposed to more noun-noun compounds and other new words in fashion. Teachers could also integrate some new words into the course or assign students articles with authentic English as reading homework and check their understandings of the unknown words afterwards.

Learning English should not be just learning the language itself but extended to learning the history, culture and social customs of English-speaking countries as well. Some interpretations against the dominant relations are caused due to unfamiliarity with the language environment and cultural atmosphere of the target countries. Most of the students in China learn English in class and receive no further language instruction elsewhere. They learn the vocabulary, the grammar and ways to communicate in order to better serve in their examinations. Since no content concerning culture and social customs will appear on test papers, students feel no need to learn these “extras”. Even though the learning process in class can guarantee them expressing ideas in English without difficulties, it cannot guarantee that they could fully understand input from native speakers because it takes a much bigger receptive than productive vocabulary to help one engaged in an interaction in L2 (Nation, 2001). It is also the reason why Chinese students on average are poor at understanding English idioms. This cultural gap can be filled up

by including cultural courses in the curriculum or adding cultural elements in class.

With respect to the narrow approach of learning English noun-noun compounds, it is argued that basic rules of construction should be explained clearly to the students once they have their first exposure of compounds of this type. It is better for teachers to introduce the noun-noun compounds in the appropriate contexts. For instance, teachers should raise the awareness of the students to the characteristics of this type of word formation, rather than ignore them or treat them just as common nouns. Teachers could present a certain amount of examples of noun-noun compounds with different relationship types embedded so that the students would understand the underlying patterns of noun-noun compounds better. In addition, comparisons with Chinese compounds should be made clear to elicit the students to realise the differences and similarities between English and Chinese compounds. Otherwise, it is very likely that students take English compounds for granted and understand them in the way they understand their first language. The teaching and learning of noun-noun compounds are often neglected in the English language teaching in China. There are plenty of other aspects within this context that often result because of a lack of attention by teachers. Many invisible misconceptions have been formed by the students, which if not corrected in time and once settled in their minds, could be carried along with them for a long time in the future.

Conclusion

This current study helps to bridge a gap in research into how L2 learners interpret English noun-noun compounds. It extended this research area by collecting data from English learners in China so as to enable a meaningful comparison between learners whose native

languages are both compounding languages.

The findings of this study have two main implications: one finding concerns the interpretations of novel English compounds by English learners in China. The results demonstrate a general ability to grasp the underlying pattern and that the learners are capable of defining the meaning of compounds using a word or a phrase to form a relation between the two constituents. From their answers in the compound test, it can be seen that they are aware of the roles of modifiers and heads and most of the time they are able to use the modifier to describe the head. However, they are unable to interpret the meanings of the compounds appropriately every time. This could be attributed to their limited levels of English or the influence of their native language. The other implication is related to the comparison. Comparing the results of the L1 children and L2 adults shows that these groups differ in relation preference but share a common trend in terms of their performances – a number of participants in both groups reversed the roles of modifiers and heads when interpreting some compounds.

To sum up, this study is only the first step in researching second language learners' interpretations of English noun-noun compounds. It focuses on how they interpret the compounds and to what extent their interpretations differ from those of L1 children. Future research could be conducted to further explore whether L2 learners' interpretations of English compounds are more affected by compounds with the same modifier or head. Findings could be compared with results of similar studies of L1 adults and children in order to determine what knowledge L2 learners are using when they interpret these compounds. This could lead to further developments in the field of teaching and learning compounds.

References

- Aitchison, J. (1990). *Words in the Mind*. Oxford: Basil Blackwell.
- Becker, J. A. (1994). “Sneak-shoes”, “sworders” and “nose-beards”: A case study of lexical innovation. *First Language*, 14, 195-211.
- Clark, E. V. (1993). *The Lexicon in Acquisition*. Cambridge: Cambridge University Press.
- Clark, E. V. (2003). *First Language Acquisition*. Cambridge: Cambridge University Press.
- Ellis, N. (1994). Implicit and explicit vocabulary acquisition. In N. Ellis (Ed.) *Implicit and Explicit Learning of Languages* (pp. 211-282). London: Academic Press.
- Ellis, N. (1997). Vocabulary acquisition: word structure, collocation, word-class, and meaning. In Schmitt & McCarthy. *Vocabulary: Description, Acquisition and Pedagogy*. Cambridge: Cambridge Language Teaching Library.
- Hulstijn, J. (2000). Mnemonic methods in foreign language vocabulary learning: Theoretical considerations and pedagogical implications. In J. Coady and T. Huckin (eds) *Second Language Vocabulary Acquisition* (pp. 203-224). Cambridge: Cambridge University Press.
- Kellerman, E. (1984). The empirical evidence for the influence of the L1 in interlanguage. In A. Davies, C. Cramer and A. Howatt (Eds.) 1984: *Interlanguage*.
- Kiparsky, P. (1983). From cyclic phonology to lexical phonology. In H. Van der Hulst & N. Smith (Eds.), *The Structure of Phonological Representations* (pp. [131-175]). Dordrecht, the Netherlands: Foris.
- Krott, A. and Nicoladis, E. (2005). Large constituent families help children parse compounds. *Journal of Child Language*, 32, 139-158.
- Krott, A. and Nicoladis, E. (2007). Word family size and French-speaking children’s segmentation of existing compounds. *Language Learning*, 57:2, [201-228].

- Krott, A., Gagné, C. and Nicoladis, E. (2008). How the parts relate to the whole: A usage-based account of children's interpretations of Novel Compounds. *Journal of Child Language*.
- Lewis, M. (1993). *The Lexical Approach*. Hove and London, England: Language Teaching Publications.
- Murphy, V. (2000). Compounding and the representation of L2 inflectional morphology. *Language Learning*, 50(1), 153-197.
- Murphy, V. and Nicoladis, E. (2006). When answer-phone makes a difference in children's acquisition of English compounds. *Journal of Child Language*, 33, 677-691.
- Nation, I.S.P. (2001). *Learning Vocabulary in Another Language*. Cambridge: Cambridge University Press.
- Perdue, C. (1993). *Adult Language Acquisition: Cross-linguistic Perspectives*. Cambridge: Cambridge University Press.
- Pinker, S. and Prince, A. (1988). On language and connectionism: Analysis of a parallel distributed processing model of language acquisition. *Cognition*, 28, 73-193.
- Pinker, S. and Prince, A. (1992). Regular and irregular morphology and the psychological status of rules of grammar. In L. A. Sutton, C. Johnson, & R. Shields (Eds.), *Proceedings of the 17th annual meeting of the Berkeley Linguistics Society* (pp. 230-251). Berkeley: Berkeley Linguistics Society.
- Ringbom, H. (1986). Crosslinguistic influence and the foreign language learning process. In E. Kellerman and M. Sharwood Smith (Eds.) *Crosslinguistic Influence in Second Language Acquisition*. New York and Oxford: Pergamon.
- Swan, M. (1997). The influence of the mother tongue on second language vocabulary

acquisition and use. In Schmitt & McCarthy. *Vocabulary: Description, Acquisition and Pedagogy* (pp. 156-180). Cambridge: Cambridge Language Teaching Library.

Takač, V. (2008). *Vocabulary Learning Strategies and Foreign Language Acquisition*. Clevedon, UK: Multilingual Matters Ltd.

Tomasello, M. (1992). *First Verbs: A Case Study of Early Grammatical Development*. Cambridge: Cambridge University Press.

Tomasello, M. (2000). First steps towards a usage-based theory of language acquisition. *Cognitive Linguistics*, 11, 61-82.

Do Recasts Promote Noticing the Gap in L2 Learning?

Hideki Sakai

Shinshu University, Japan

caff@nifty.com

Bio Data:

Hideki Sakai is an associate professor in the Faculty of Education at Shinshu University, where he teaches undergraduate and graduate courses in TEFL methodology and second language acquisition. His current research interests include second language interactional studies, classroom second language acquisition, and psycholinguistic aspects of listening and speaking.

Abstract

This paper reports the effects of implicit negative feedback in the form of recasts on noticing. Twenty Japanese-speaking learners of English were assigned to an experimental group (the recast group, $n = 10$) or a control group (the no-feedback group, $n = 10$). They engaged in an information gap task, during which the recast group received recasts to their erroneous utterances or non-corrective repetition of their targetlike utterances, whereas the no-feedback received no corrective feedback. The participants' verbal reports about noticing about their errors were elicited through stimulated recall. Both groups noticed errors or problems at the moment of production about 10% of the time. Further, the recast group noticed their errors through recasts 21.6% of the time. Results suggest that recasts alone did help L2 learners notice their errors.

Keywords: implicit negative feedback, recasts, noticing, stimulated recall, negative evidence

Introduction

Over the past few decades, conversational interaction and its effects on second language (L2) learning have drawn considerable attention in the field of language acquisition (e.g., Mackey, 2007). According to Long's (1996) interaction hypothesis, "environmental contributions to acquisition are mediated by selective attention and the learner's developing L2 processing capacity" (p. 414). In other words, conversational interaction may facilitate L2 learning because it constitutes a context in which L2 learners' attention is directed toward some formal aspects of the language. In particular, Long argued that "Negative feedback obtained during negotiation work or elsewhere may be facilitative of L2 development, at least for vocabulary, morphology, and language-specific syntax, and essential for learning certain specifiable L1-L2 contrasts" (p. 414). Recasts, on which this study focuses, are one such type of negative feedback. By utilizing a stimulated recall technique to elicit introspective data, the present experimental study addresses the issue of how L2 learners interpret recasts for the learning of morphology, in particular, irregular past tense forms.

Recasts as Positive and Negative Evidence

According to Nicholas, Lightbown, and Spada (2001), recasts are "utterances that repeat a learner's incorrect utterance, making only the changes necessary to produce a correct utterance, without changing the meaning" (pp. 732-733). Recasts are considered to provide the linguistic data of both positive evidence (i.e., what is grammatical in the target language) and negative evidence (i.e., what is ungrammatical in the target language) at the same time. The following shows an example of a recast, taken from the data gathered for the current study.

(1) An Example of a Recast

Researcher: *What did he do at four in the evening?*

Koichi: *He shaked the bottle.* [an erroneous utterance]

Researcher: *He shook the bottle?* [a recast]

Koichi: *Yes.*

The learner, Koichi (a pseudonym), produced an overgeneralized past form of *shake*, which was followed by a recast in the next researcher's turn. The recast maintained the proposition of the original utterance but contained a reformulation of the erroneous element into the correct one, *shook*. In this example, the recast was supposed to inform the learner of both positive evidence (i.e., the correct form was *shook*) and negative evidence (i.e., the form *shaked* was ungrammatical in English). The semantic and discursal properties of recasts that repeat the information generated by learners and that are juxtaposed with the erroneous utterances make it easier for L2 learners to make cognitive comparisons between their interlanguage and the target language (Long, 1996, 2007; see also Saxton, 1997, 2000).

Thus far, a number of descriptive studies have shown that recasts are the most frequent negative feedback types (e.g., Doughty, 1994, Lyster, 1998a, 1998b; Lyster & Ranta, 1997; Nassaji, 2007; Oliver, 1995; Panova & Lyster, 2002; Roberts, 1995; Van den Branden, 1997). Further, recasts have been found to be effective for L2 learning (e.g., Ayoun, 2001; Carroll & Swain, 1993; Doughty & Varela, 1998; Ellis, 2007; Han, 2002; Ishida, 2004; Iwashita, 2003; Loewen & Nabei, 2007; Long, Inagaki, & Ortega, 1998; Mackey, 1995, 1999; Mackey & Philp, 1998; for a meta-analysis, see Mackey & Goo, 2007). Recently, L2 researchers have addressed the issue of how recasts benefit L2 learning (Gass & Mackey, 2007; Leeman, 2003; Long, 2007). One probable reason for much interest in recasts is that if the mechanisms underlying the functions of recasts are made clearer, recasts will be utilized more effectively in L2 classrooms. Recasts are considered to be implicit in that they do not provide an overt indication of the existence of errors to the learner (Ellis, 2007; Ellis, Loewen, & Erlam, 2006). Thus, they have been argued to provide negative evidence while avoiding interrupting the flow of communication (Long, 2007; Long, Inagaki, & Ortega, 1998).

However, at the same time, the very implicitness of recasts as negative feedback may bring about one problem for L2 learning. For example, Lyster (1998a) argued that recasts are ambiguous as negative feedback because recasts and non-corrective repetition, that is to say, repetition of learners' correct utterances, are difficult to distinguish from each other. On the other hand, Long (1996) argued that L2 learners do notice the corrective nature of recasts. As indirect evidence, research has shown that L2 learners respond to

recasts and non-corrective repetitions differently (e.g., Doughty, 1994; for a review of L1 studies, see Long, 1996). For example, Doughty (1994) reported that L2 learners of French imitated recasts more frequently than the non-corrective repetition (21.5% vs 2.3%). However, it has been pointed out that differing responses to recasts and non-corrective repetitions do not always imply that L2 learners perceive recasts as negative feedback (e.g., Mackey & Philp, 1998). Recently, a more direct introspective technique (i.e., stimulated recall) has been employed in order to elicit noticing data from the learners (Egi, 2007a, 2007b; Mackey, Gass, & McDonough, 2000; Nabei & Swain, 2002; Roberts, 1995).

Recasts and Noticing for L2 Learning

Stimulated recall is “one subset of a range of introspective methods that represent a means of eliciting data about thought processes involved in carrying out a task or activity” (Gass & Mackey, 2000, p. 1). In stimulated recalls, learners are asked to report what they were thinking at the time of the target event. In other words, stimulated recalls are retrospective, not concurrent, introspective methods in that there is some interval between the time of the event and the time of verbal reports. The data elicited through stimulated recall concern what Schmidt (1990) called *noticing*, that is to say, only a subset of detected information in Tomlin and Villa’s (1994) sense. Tomlin and Villa (1994) defined *detection* as “the cognitive registration of sensory stimuli” (p. 192) and argued that detection does not entail awareness (p. 199). On the other hand, Schmidt (1990) defined *noticing* as “the availability for verbal report” (p. 132). Because stimulated recall is an elicitation of verbal reports, not every item of detected information will be elicited through stimulated recall. (For a discussion of theoretical issues of noticing and awareness, see Robinson, 1995; Simard & Wong, 2001.)

Previous studies using stimulated recall have showed that L2 learners had difficulty in perceiving the corrective intention in recasts precisely, in particular, when the recasts concerned grammatical-syntactic errors (Gass & Lewis, 2007; Mackey, Gass, & McDonough, 2000; Roberts, 1995). Recently, Egi (2007a) made an attempt to explore how L2 learners interpret recasts. She investigated the relationship between L2 learners’ interpretation of recasts and factors such as linguistic targets, recast length, and number

of changes. Her participants were 49 learners of Japanese. They performed two communicative tasks with native speakers in dyads. Of the 49, 31 performed the tasks under immediate recall conditions; the other 18 attended a stimulated recall session after the task. Their recall protocols were categorized as (a) responses to content, (b) negative evidence, (c) positive evidence, and (d) negative and positive evidence. Responses to content were defined as “comments indicating that the learner interpreted the recast as the interlocutor’s replies to the conversational content” (p. 522). Negative evidence was operationalized as “comments indicating that the learner was aware of error production or correction without a clear indication that the learner noticed the targetlike form provided in the recast” (p. 524). Positive evidence referred to “comments indicating that the learner attended to the targetlike model without reporting having produced errors or having been corrected” (p. 534). The last category, negative and positive evidence, was operationalized as “comments that indicated the learner’s (a) recognition of error production or correction and (b) attention to the targetlike form in the recast” (p. 524). Of the 476 morphosyntactic feedback episodes, her participants interpreted recasts as responses to content (21.22%), negative evidence (34.03%), positive evidence (18.70%), and negative and positive evidence (26.05%); of the 77 lexical feedback episodes, they interpreted recasts as responses to content (18.18%), negative evidence (33.77%), positive evidence (18.18%), and negative and positive evidence (29.87%). No statistical difference was found between the linguistic targets and the interpretations of recasts. On the other hand, the recast length and the number of changes were found to affect the participants’ interpretations of recasts. Shorter recasts helped L2 learners interpret the recasts as negative and positive evidence more frequently than longer recasts; recasts with fewer changes were more effective for learners to interpret them as negative and positive evidence than those with more changes. What is important in Egi’s study is that she developed a categorization system for coding verbal protocols on the basis of types of linguistic data (negative evidence, positive evidence, or both). However, her system still seems to be insufficient to address the issue of whether or how recasts facilitate learners’ noticing.

In the literature, noticing has been categorized mainly into (a) noticing a form in the input, (b) noticing one’s interlanguage deficiencies (or a hole), and (c) noticing the gap

between the interlanguage and the target language (Doughty & Williams, 1998, p. 228; Swain, 1998, p. 66). In her output hypothesis, Swain (1985, 1993, 1995) proposed that output in the target language may promote the second type of noticing: “the activity of producing the target language may prompt second language learners to consciously recognize some of their linguistic problems” (1995, p. 126). Thus, it is necessary to distinguish noticing one’s errors through the activity of production from noticing errors through recasts. In Example (1) mentioned above, Koichi produced an overgeneralized form *shaked* for the past tense. In the case that, at the moment of production, he noticed his error (that *shaked* was wrong), the role of the recast may have been only to provide positive evidence about how to say the past form of the verb. By contrast, in the case that he did not recognize his error at the moment of production, the recast may have had the dual function of indicating that *shaked* was wrong and of showing the correct form. In Egi’s (2007a) categorization system, no clear distinction was made between L2 learners’ noticing of errors by themselves and L2 learners’ noticing of errors with the help of recasts. For example, she provided an example of the verbal protocol coded as negative and positive evidence as follows: “I had no idea how to say this verb, so I just said *hayai* [“fast”], which turned out to be wrong, and then you told me how to say ‘hurry,’ *isoide*” (p. 524). The first part of the comment indicates that this learner had already recognized her problem at the moment of production. Thus, the recast may not have provided the learner with negative evidence; rather, it may have functioned only as positive evidence. In other words, Egi’s categorization did not make it clear whether recasts facilitate noticing errors or not.

To sum, although several studies (Gass & Lewis, 2007; Mackey, Gass, & McDonough, 2000; Roberts, 1995) pointed out the difficulty of noticing the corrective intention in recasts, Egi’s (2007a) detailed analysis of verbal protocols about L2 learners’ interpretations of recasts suggests that recasts are indeed interpreted as negative evidence (accounting for 34.03% for morphosyntactic feedback episodes and 33.77% for lexical feedback episodes) or negative and positive evidence (accounting for 26.05% for morphosyntactic feedback episodes and 29.87% for lexical feedback episodes). However, because there have so far been a very limited number of studies on L2 learners’ interpretations of recasts, further research is needed. Furthermore, as pointed out earlier,

in order to address the function of recasts as negative feedback, it is necessary to isolate the effects of recasts from those of the activity of output itself. Thus, the present study was designed to separate the confounding effects of recasts and language production. This study involved using a no-feedback group as a comparison group. By comparing the recast group with the no-feedback group, this study attempted to examine the effects of recasts, excluding the noticing effects of production. This study is significant in that by utilizing a stimulated recall technique, it attempted to explore the role of recasts in details.

The research question posited for this study was: Do recasts alone provide negative evidence about the grammaticality of L2? In other words, this study examined whether L2 learners, who did not notice their linguistic problems or errors at the moment of production, notice their errors through recasts. On the basis of Egi (2007a), it was hypothesized that the recast group will notice more errors in production than the no-feedback group.

Method

Participants

The participants were initially 26 Japanese university student volunteers from a required general English course at a national university in Japan. Since the results of an exit questionnaire indicated that six participants had experience abroad and/or had started to learn English prior to entrance to junior high school, those six participants were eliminated from the analysis of the current study in order to make the two groups comparable in terms of prior L2 learning experiences. Finally, the number of participants was 20 (10 male and 10 female). The twenty participants were randomly assigned to the recast group ($n = 10$) or the no-feedback group ($n = 10$) as a control group. Table 1 shows the pertinent information (sex, age, and years of learning English) about the participants. All of them started to learn English in their 1st year of junior high school, that is to say, in the 7th grade. Prior to the experiment, they had had received about 7 years of formal instruction in English at junior and senior high school as well as at university.

Table 1***Information about Participants (Sex, Age, and Years of Learning English)***

Recast Group (<i>n</i> = 10)				No-Feedback Group (<i>n</i> = 10)			
Learner	Sex	Age	English study	Learner	Sex	Age	English study
Tomoko	F	20;01	7;11	Maki	F	20;05	7;11
Yuki	F	20;00	7;11	Ayumi	F	20;05	7;11
Rinko	F	21;07	8;11	Chitose	F	18;04	6;03
Yuko	F	18;04	6;04	Kaori	F	19;06	7;03
Momo	F	19;02	6;04	Shuko	F	19;03	6;04
Toshiya	M	20;07	7;11	Taka	M	19;01	6;11
Yusuke	M	20;03	7;11	Ichiro	M	18;06	6;04
Taro	M	20;09	7;11	Seiji	M	18;11	6;04
Akira	M	20;03	7;11	Hideki	M	18;09	6;04
Koichi	M	19;10	7;11	Kotaro	M	18;09	6;04
<i>M</i>		20;01	7;08	<i>M</i>		19;02	7;00

Note. All names are pseudonyms.

Target Grammatical Structure

The target structure chosen for this study was irregular past tense forms for mainly two reasons. First, several researchers had dealt with this structure in studies on the effects of interaction (e.g., Ellis, 1987; Nobuyoshi & Ellis, 1993; Takashima, 1995). Second, it is well pointed out that L2 learners may produce overgeneralized forms, that is, *-ed* forms, instead of the appropriate irregular past forms (e.g., Dulay, Burt, & Krashen, 1982, p. 158; Lightbown & Spada, 1993, p. 56). In order to investigate whether recasts provide negative evidence for the learner, it is necessary to create a condition in which the learner will produce interlanguage forms different from the targetlike forms. For example, an L2 learner produces *drinked* as the past form of the verb *drink* and is provided with a recast including the targetlike form *drank*. This case may enable researchers to examine whether the learner will not only notice the correct form *drank*, but also recognize that the overgeneralized form *drinked* is not acceptable.

Prior to this study, another eleven university students who had similar backgrounds to the participants in this study were asked to write the past forms of verbs. This pilot study yielded ten verbs for which most learners produced generalized forms: *feed*, *beat*, *cast*, *strike*, *draw*, *swing*, *bite*, *shake*, and *fight*. These ten words were then used as the target

verbs for the tasks used in the main study.

Procedure

The procedure of the study is summarized in Table 2. The participants in both groups took part in the experiment individually. The experiment consisted of three tasks, in which they communicated with the researcher, and a stimulated recall session. First, the participants performed Task 1, Yuki’s Story, which served as a pretest. Then, they performed Task 2, Ken’s Story, the treatment task. In Task 2, the participants in the recast group received recasts of their erroneous utterances or non-corrective repetition of their targetlike utterances from the researcher. On the other hand, the participants in the no-feedback group received no corrective feedback or positive evidence. Then, both groups carried out Task 3, Yuki’s Story,¹ followed by the stimulated recall session, in which participants were asked about what they had thought during Task 2. The interval between Task 2 (Ken’s story) and the stimulated recall was about 10 minutes.

Table 2: Procedures

	Time	Recast group	No-Feedback group
Explanation	5 min	- Consent form	- Consent form
Task 1 (pretest)	10 min	- Yuki’s Story	- Yuki’s Story
Task 2 (treatment task)	10 min	- Ken’s Story received recasts or non-corrective repetition	- Ken’s Story received no feedback but received some positive encouragement like “Yes” and “OK.”
Task 3 ^a	10 min	- Yuki’s Story	- Yuki’s Story
Verbal reports	15 min	- Stimulated recall	- Stimulated recall
Questionnaire	10 min	- Exit questionnaire	- Exit Questionnaire

a. The results of Task 3 are not reported in the current study (see Note 1).

Thus, in this study, the independent variable was the provision of recasts in Task 2; the dependent variable was obtained from the verbal protocols in the stimulated recall.

Tasks

To serve as a pretest, a story-making task was provided before the treatment session. In

this task, the participants were asked to make a story about what Yuki did. First, they were given a chart that contained a time line from 8:00 to 22:00 and ten small picture cards that depicted the actions of the target verbs, which had been selected through a pilot test. They were told to shuffle the picture cards and to place each card into a time slot on the chart with the card faces down. Then, they were allowed to turn the cards over. They were told to narrate in English what Yuki (a female character) did yesterday. The planning time of 1 minute was given, during which they were not allowed to make written notes. Then, they were told to start with the sentence, “Yuki got up at seven yesterday,” so that the past time reference was obligatory. Through this task, the participants’ production data was obtained.

The treatment task was an information gap task, Ken’s Story. The participants were given a timeline and ten small picture cards (see Appendix). The pictures depicted the actions of ten target verbs, although the actions were different from those of Yuki’s Story. For example, a picture for the verb *shake* showed that Yuki was shaking hands in Yuki’s Story and that Ken (a male character) was shaking a bottle in Ken’s story. First, they were told to shuffle the picture cards and place them on the chart. Then, they were instructed to turn the cards over. The researcher held the same set of ten picture cards and the time chart and asked the participants, “What did Ken do at ten in the morning?” so that the cards could be placed in the same order as on the participants’ chart. For the recast group, responses to the participants’ elicited utterances in the form of recasts were provided when the participants made an error or in the form of non-corrective repetitions when the participants’ utterances were targetlike.² The reason for the provision of non-corrective repetition was that, without repetition of correct utterances, the participants would notice the corrective nature of recasts easily. Participants in the no-feedback group were provided with responses like “Yes.” “I see.” and “OK.” to every utterance. Thus, no information about the grammaticality of the target structures was given to the no-feedback group. The following examples illustrate the two conditions.

(2) The Recast Group

Yusuke: *He ... he swung the bat.* [an erroneous utterance]
Researcher: *Ah, he swung a bat?* [a recast]
Yusuke : *Yes.*

(3) The No-Feedback Group

Kaori: *He shaked a bottle.* [an erroneous utterance]

Researcher: *Um. OK. I got it.*

Thus, the participants were provided with almost the same number of opportunities to produce the target structures. However, there was no provision to control for opportunities for the participants to react to the feedback.

Stimulated Recall

The stimulated recall technique was employed as a method of eliciting noticing data.³ The instruction of the stimulated recall was created on the basis of Egi (2004) and Gass and Mackey (2000). The participants were asked to report what they had been thinking about for every utterance. As prompts to stimulate their recall of noticing during the interactions, they listened to audiotaped interactions of the treatment session and looked at the pictures and timeline. The recall session was conducted in Japanese, the L1 for the participants and the researcher.

Analysis

The researcher transcribed the participants' performance in all the tasks and their verbal protocols in the stimulated recall. A graduate assistant checked all the transcriptions. The following analyses were based on the transcriptions.

The production of the past tense forms in Task 1 was analyzed as a measure of knowledge of the target structures. More specifically, the accuracy rate of the past tense forms was calculated using two methods, following Pica (1983; see also Takashima, 1995, pp. 72-75): supplied in obligatory context (SOC) analysis and target-like use (TLU) analysis. The formula for SOC analysis is: $[(\text{the number of correct suppliance in obligatory contexts} \times 2) + (\text{the number of misformations in obligatory contexts} \times 1)] / (\text{total obligatory contexts} \times 2)$. The formula for TLU analysis is: $(\text{the number of correct suppliance in obligatory contexts}) / [(\text{the number of obligatory contexts}) + (\text{the number of suppliance in non-obligatory contexts})]$. This study reports type counts, which are not influenced by repetitive use of the same words.

To obtain the indices of the SOC and TLU analyses, the researcher coded the participants' utterances into (a) correct use of the past tense form (Correct), (b) use of wrong past forms (Misformation), (c) non-use of past tense forms in the obligatory contexts (Incorrect), and (d) wrong use of past tense forms in contexts that do not require past tense forms (Overuse). A graduate student assistant coded 25.0% of the data, and the intercoder agreement reached 96.4% (269 out of the 279 instances).

To code the verbal reports in the stimulated recall session, an individual analysis proceeded in two steps: (a) analyzing the participants' production in the treatment task (Task 2, Ken's Story) and (b) coding their verbal reports. First, the production in the treatment task was coded into five categories: correct, misformation, incorrect, overuse, and other. The category *Other* refers to cases in which the participants were not provided with recasts containing their verbs, as in the following example.

(4) The Category *Other*

Researcher: *What did he do at three in the afternoon?*

Toshiya : *Horse he take a carrot for horse.* [Other]

Researcher : *Ah, he fed a carrot to the horse?* [a recast]

Toshiya : *Yes.*

Recasts like this example contained corrective information about the past tense reference and the choice of verb. In this study, such cases were excluded from the analysis. This analyzing procedure of Task 2 was similar to that of Task 1; the former was intended to be used for a further analysis of verbal protocols, whereas the latter was utilized for the calculation of accuracy rate of the past tense forms. Another assistant coded all the productions in Task 2 as well. The intercoder agreement was 95.0% (266 out of the 280 instances). The disagreements were resolved after discussion.

Second, the verbal reports regarding the participants' non-targetlike use of past tense forms (that is, Misformation and Incorrect) were analyzed and coded according to four categories as shown in Table 3 (see also Sakai, 2004). These categories were based on the relevant literature (Doughty & Williams, 1998, p. 228; Swain, 1998, p. 66). First, their reports about noticing were divided into (a) noticing a hole (noticing what they cannot

express in their interlanguage system) and (b) noticing the gap (noticing that what they say is different from the target language). The latter type of noticing was further divided into noticing a gap (a) at the moment of production and (b) through a recast.

Table 3

Coding Categories for Verbal Protocols

Category	Noticing Type	Definition
Problem	Noticing a hole	- make comments about what they cannot express or what they only partially know
Production	Noticing a gap at the moment of production	- notice errors in their utterance, or that their utterance is wrong, at the moment of speaking
Recast	Noticing a gap through a recast	- notice errors, or that their utterance is wrong, by receiving a recast
Unnoticing	Noticing at the interview / No report	- notice errors, or that their utterance is wrong, at the moment of the verbal report, or fail to make a report

The following are examples for each category. The original comments of the stimulated recall were in Japanese.

(5) Noticing a hole (coded as *Problem*)

Task 2: Ken's Story

Momo: *He dro.. drinken coffee. He have a coffee.* [an erroneous utterance]

Researcher: *He drank coffee?* [a recast]

Momo: *Yes.*

Researcher: *OK.*

Stimulated Recall:

Momo: *Well, about this, I stopped because I was thinking of what the past form of the word drink was. Actually, I wanted to say a cup, full of coffee, but I did not know the expression. That was*

before speaking. Yes. I did not know the past form of the word drink before starting to speak.

(6) Noticing a gap at the moment of production (coded as *Production*)

Task 2: Ken's Story

Researcher: *What did he do at six in the evening?*

Taro: *Ah, he taken care horse.* [an erroneous utterance]

Researcher: *He took care of a horse?* [a recast]

Stimulated Recall:

Taro: *Well, about this, I wanted to express "feeding" when I saw the picture. But I didn't know how to say it. Ah, so while I was thinking, I came up with the expression "take care of," so I thought I would say it instead. I don't know why, but I said taken for took. This may be because I was upset under pressure to say something. In my brain, I wanted to say "took care" correctly. But I said "taken care," I don't know why.*

Researcher: *When did you notice about taken?*

Taro: *After saying it, I thought I would correct it. But at that time you responded. So I thought it was OK. That's all.*

(7) Noticing a gap through a recast (coded as *Recast*)

Task 2: Ken's Story

Researcher: *Next what did he do at eight in the evening?*

Yuki: *He shaked umm he shaked a ...* [an erroneous utterance]

Researcher: *Ah. He shook a bottle?* [a recast]

Yuki: *Yes.*

Researcher: *OK. I got it.*

Stimulated Recall:

Yuki: *Well, let's see, I thought the past form of shake was shook, and, and, I did not come up with the word bottle. I thought that I knew it, but I did not remember it. That's all.*

Researcher: *Yes. When did you notice about shook?*
Yuki: *Well, when I was told by you, I thought that was it.*

(8) Noticing at the interview or no report (coded as *Unnoticing*)

Task 2: Ken's Story

Yusuke: *And ... he, he swunged the bat.* [an erroneous utterance]
Researcher: *Ah, he swung a bat?* [a recast]
Yusuke: *Yes.*
Researcher: *OK. I got it.*

Stimulated Recall:

Yusuke: *Well, before starting to speak, I thought of a word to mean "swinging a bat." That's all.*

The researcher coded all the verbal reports. The assistant who coded the participants' production in Task 2 and another assistant coded the verbal reports obtained from the recast group and the no-feedback group, respectively. The intercoder agreements were 96.1% (49 out of the 51 instances) and 98.7% (76 out of the 77 instances). Disagreements were resolved after discussion.

Statistical Analysis

First, non-parametric Mann-Whitney tests and chi-square test were performed to check the comparability of the two groups in terms of the accuracy of the past tense forms and the distribution of past tense forms.

To examine whether the two groups differed for each category of noticing, three non-parametric Mann-Whitney tests were performed with Holm's sequential Bonferroni method⁵ to control for Type 1 error.

Results

Comparability of the Two Groups

The participants were randomly assigned to one of the two conditions; so to check the comparability of the two groups, the accuracy of the past tense forms in the pretest (Task

1) was examined. The results of SOC were $Mdn = .65$, $IQR = .46$ for the recast group and $Mdn = .64$, $IQR = .40$ for the no-feedback group; the results of TLU were $Mdn = .54$, $IQR = .55$ for the recast group and $Mdn = .53$, $IQR = .45$ for the no-feedback group.⁴ Non-parametric Mann-Whitney tests did not detect statistically significant differences in the accuracy of the past tense forms in Task 1 between the two groups ($p = .869$ for the SOC analysis; $p = .782$ for the TLU analysis).

Table 4 shows the distribution of each category of past tense forms for the two groups. The chi-square test revealed no statistically significant difference between the two groups ($\chi^2(2) = 0.658$, $p = .720$). In other words, the recast group and the no-feedback group did not differ in the production of past tense forms on Task 2. Therefore, these results suggest that the participants in the two groups had relatively similar proficiency levels in terms of past tense forms.

Table 4

Distribution of Each Category for the Past Tense Forms in Task 2

	Recast Group		No-Feedback Group	
Correct	64	(48.5%)	71	(48.0%)
Incorrect	50	(37.9%)	61	(41.2%)
Misformation	18	(13.6%)	16	(10.8%)
Total	132	(100%)	148	(100%)

Note. For the recast group, the number in the category *Other* was 20 out of 132 (3 for Correct, 14 for Incorrect, and 3 for Misformation).

Verbal Reports

Table 5 shows the frequencies of each noticing category for the two groups (see also Figure 1). For the recast group, the category *Unnoticing* accounted for 70.5% (36 out of the 51 instances); for the no-feedback group, this category was observed more frequently (89.6%, 69 out of the 77 instances). Table 6 indicates that the medians of frequency rates of *Unnoticing* for the recast and the no-feedback groups were .63 (i.e., 63%, $IQR = .36$) and 1.00 (i.e., 100%, $IQR = .18$) respectively. Tables 5 and 6 show that for both the recast and the no-feedback groups, the categories *Problem* and *Production* were observed in similar percentages (7.9%, 4 instances for the recast group; 10.4%, 8 instances for the no-

feedback group) and that the medians of frequency rates of *Problem* and *Production* were both .00 (*IQRs* = .14 and .18 respectively). For the recast group, noticing a gap though a recast accounted for 21.6% (11 out of the 51 instances, *Mdn* = .27, *IQR* = .50). Thus, the results suggest (a) that since both groups reported *Problem* and *Production* to a similar degree, the participants noticed errors or what they could not express in English irrespective of the provision of recasts, and (b) that since the recast group noticed errors at a higher rate than the no-feedback group, the provision of recasts facilitated the participants' noticing greatly. In other words, the main differences between the two groups came about by means of noticing through recasts.

Table 5

Distribution of Noticing Types by Group

Verbal Report	Recast Group				No-Feedback Group			
	Incorrect	Mis	Total		Incorrect	Mis	Total	
Problem	0	1	1	(2.0%)	2	0	2	(2.6%)
Production	2	1	3	(5.9%)	5	1	6	(7.8%)
Recast	5	6	11	(21.6%)	0	0	0	(0.0%)
Unnoticing	29	7	36	(70.5%)	54	15	69	(89.6%)
Total	36	15	51	(100.0%)	61	16	77	(100.0%)

Note. Mis = Misformation.

Table 6

Medians and Interquartile Ranges of Frequency Rates of Noticing Types by Group

Verbal Report	Recast Group		No-Feedback Group	
	<i>Mdn</i>	<i>IQR</i>	<i>Mdn</i>	<i>IQR</i>
Problem/Production ^a	.00	.14	.00	.18
Recast	.27	.50	.00	.00
Unnoticing	.63	.36	1.00	.18

Note. Frequency rates were obtained by dividing the frequency for each noticing type by the total number of noticing. They were calculated for each participant. Then, medians and interquartile ranges of frequency rates were calculated for each group.

a. The frequency rates of *Problem* and *Production* were combined for a further statistical analysis that tests the effects of output.

These findings were supported by statistical analyses. First, statistically significant differences were obtained between the two groups for the *Recast* category and the *Unnoticing* category ($p = .008$ for *Recast* and $p = .022$ for *Unnoticing*). Thus, the hypothesis that the recast group will notice more errors in production than the no-feedback group was confirmed. The comparison of the two groups for *Problem* and *Production* was not significant ($p = .720$). Therefore, the participants in the two groups noticed their errors or what they could not express in English to a similar degree.

Moreover, Table 5 provides breakdowns by error types (Incorrect and Misformation). Regarding the results for Misformation, the recast group noticed 8 of the 15 instances, whereas the no-feedback group noticed only 1 of the 16 instances. The results suggest that it may be difficult for the participants to notice errors coded as Misformation, such as *shaked* and *fighted*, by themselves and that recasts may facilitate such noticing of overgeneralized forms.

It is important to ascertain whether only a few individuals contributed to the high rate of noticing of errors through recasts. A post hoc analysis of the data showed that such individual biases were not observed. Of the ten participants in the recast group, three (Rinko, Momo, and Yusuke) did not report noticing of errors through recasts; four (Yuko, Toshiya, Akira, and Koichi) reported their noticing of errors through recasts for misformation (1 instance) and incorrect (1 instance); two (Tomoko and Yuki) reported their noticing of errors through recasts for misformation (1 instance); and one (Taro) reported his noticing of errors through recasts for incorrect (1 instance). Although the number of cases for each individual was quite small, the positive effect of recasts on noticing was observed for most participants (7 out of 10).

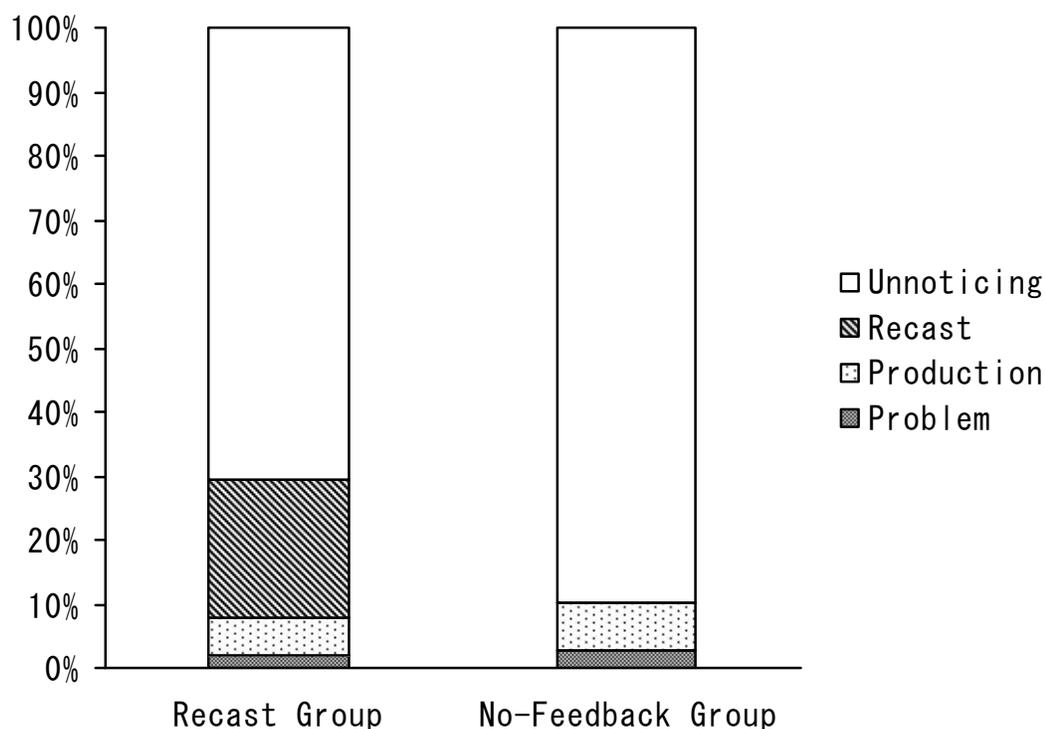


Figure 1. Distributions of noticing types by groups.

Discussion

Analyses of verbal reports showed that the provision of recasts was shown to be effective in promoting L2 learners’ noticing, in particular, of overgeneralization errors. Also, they showed that the activity of production led to a certain degree of noticing of errors or linguistic problems irrespective of the provision of recasts. This study attempted to isolate the effects of recasts from those of the activity of production on L2 learners’ noticing. By doing so, this study provides separate empirical support for the claims of the interaction hypothesis (Long, 1996, 2007) and the output hypothesis (Swain, 1985, 1993, 1995) with regard to noticing.

In his interaction hypothesis, Long (1996, 2007) claims that negative feedback facilitates L2 acquisition by drawing L2 learners’ attention to nonsalient and problematic forms. He pointed out that “Heightened attention makes detection both of new forms and of mismatches between input and output more likely, and such mismatches may also provide at least some of the information a learner needs about what is *not* permissible in a

language” (p. 453). The findings of this study lend support to his claim. First, the results showed that the recast group noticed significantly more errors than the no-feedback group and suggested that recasts are effective in promoting L2 learners’ noticing. Furthermore, as shown in the Results section, the recast group noticed overgeneralized forms (i.e., Misformation) as errors far more frequently than the no-feedback group (8 out of the 15 instances for the recast group; 1 out of the 16 instances for the no-feedback group). As Long argued, the results suggest that recasts are effective in drawing L2 learners’ attention to otherwise nonsalient forms and in informing L2 learners that overgeneralized forms such as *shaked* and *fighted* are ungrammatical in the target language.

The findings of this study also support Swain’s (1985, 1993, 1995) output hypothesis, which claims that by producing the target language, L2 learners may notice what they cannot express in the target language. According to Swain, this type of noticing, that is to say, noticing a hole, may make L2 learners more attentive to further input, leading to more opportunities for L2 acquisition. The results showed that both the recast and the no-feedback groups noticed their linguistic problems or errors through the activity of production (7.9% for the recast group; 10.4% for the no-feedback group). However, it is important to note that the low rates suggest that the noticing function of output may be limited, in particular, for overgeneralized morphological errors like the irregular past tense forms. In other words, L2 learners who produce overgeneralized past forms of irregular verbs by adding the *-ed* suffix to a verb stem may have difficulty in noticing such errors by themselves (i.e., without any feedback) at the moment of production. Because the rule of the addition of the *-ed* suffix is grammatical for some set of verbs and productive in English, L2 learners may need negative evidence that applying that rule to some set of verbs (i.e., irregular verbs) is not grammatical. The activity of production itself may not provide such negative evidence.

Finally, coding systems for the verbal protocols are discussed. Previous studies used systems developed by the researchers for the purposes of their study. For example, Mackey, Gass, and McDonough (2000) and Gass and Lewis (2007) attempted to investigate whether L2 learners will notice the linguistic targets (lexical, morpho-syntactic, or phonological aspects) of negative feedback correctly. Egi (2004) aimed at examining the reactivity and veridicality of the stimulated recall method and therefore

included such categories as *no memory* and *no thoughts* in her system. Importantly, Egi (2007a, 2007b) categorized the verbal protocols on the basis of the types of linguistic data (negative evidence, positive evidence, and negative and positive evidence). Thus, her categorization system made it possible to examine what type of data L2 learners may receive from recasts. Finally, this study divided the verbal protocols according to the noticing types proposed by Swain (1998) and Doughty and Williams (1998); in particular, this study attempted to isolate the effects of production alone from the effects of recasts. In the categorization system used in this study, the category *noticing a gap through a recast* covers both (a) negative evidence and (b) negative and positive evidence in Egi's categorization system. Therefore, the categorization system of this study examined whether L2 learners received negative evidence, that is, whether they noticed their errors or not; nevertheless, it did not make clear whether they received positive evidence, that is, whether they noticed the target forms. The categories for coding must be developed with great care in accordance with the purpose of the study. One implication for future study is that it might be advantageous to combine the coding systems of Egi (2007a, 2007b) and the present study. More specifically, verbal protocols can be categorized on the basis of noticing types as in the current study so that the effects of output alone will be eliminated; and they can be further categorized according to whether L2 learners receive positive evidence, negative evidence, or both. A post hoc analysis following this procedure revealed that, out of the 11 verbal reports coded as *noticing a gap through a recast*, five were coded as negative evidence; six were coded as negative and positive evidence. Although the numbers were small, this post hoc analysis suggests that recasts provide both negative and positive evidence for L2 learners.

Conclusion

The research question posited for this study was: Do recasts alone provide negative evidence about the grammaticality of L2? By analyzing the verbal protocols of L2 learners' noticing elicited through a stimulated recall technique, this study provided a confirmative answer to this question: In other words, the results suggest that recasts do facilitate L2 learners' noticing of errors even when they did not notice their linguistic problems or errors at the moment of production. Some researchers have argued that

recasts are ambiguous as negative feedback (Lyster, 1998a) and that it is difficult for L2 learners to notice the corrective intentions of recasts accurately (Gass & Lewis, 2007; Mackey, Gass, & McDonough, 2000; Roberts, 1995). On the other hand, by analyzing verbal protocols of L2 learners' noticing in detail, Egi (2007a) found that L2 learners do receive negative evidence from recasts. Along this line, this study provided additional evidence for the function of recasts to provide negative evidence.

Based on the findings of this study, two pedagogical implications can be made. First, although production itself may lead to some degree of noticing, providing L2 learners with opportunities for interactions in which they can receive negative feedback such as recasts from a more competent interlocutor or teacher may be more effective in order to promote noticing of errors. In other words, when learners produce an overgeneralized and interlanguage form such as *shaked* or *drinked*, the findings of this study suggest that they may have difficulty in noticing their errors by themselves (i.e., through the activity of production). That is the point in which recasts may be effective. Thus, as a second pedagogical implication, it is suggested that teachers can provide recasts, depending on error types. Recasts may be effective for overgeneralized past forms of irregular verbs.

In interpreting and generalizing the findings, however, it is important to take into account a few limitations of the study. First, this study was experimental in nature. The participants in the recast group received the feedback intensively; the context was quite different from natural conversational interactions. The findings that recasts promote L2 learners' noticing need to be corroborated by more studies with different contexts including natural settings. Second, verbal reports about noticing cannot avoid suffering from memory loss (Egi, 2004; Gass & Mackey, 2000; Robinson, 1995). For example, Robinson (1995) stated that "the experience of noticing may be fleeting and thus difficult to recall" (p. 299). To overcome this problem, this study used audiotaped recordings and task materials as prompts to stimulate the recall of noticing. However, future studies may require different methods to measure noticing, such as think-aloud protocols and immediate reports, within their research design.

Notes

1. The phase of Task 3 was intended to examine the learning effects, that is, the improvement from Task

1 (pretest) to Task 3 (posttest). However, the results of Task 3 are not reported in the current study because they are beyond the scope of this study that focuses on the analysis of verbal reports on noticing.

2. The recast group was provided with either recasts or non-corrective repetition. Because all the erroneous utterances were followed by recasts, the recast group did not receive any repetition of their erroneous utterances from the researcher.
3. Two issues, reactivity and veridicality, were taken into consideration (Egi, 2004; Gass & Mackey, 2000). Reactivity concerns the influence of stimulated recall on the target performance. Adams (2003) showed that stimulated recall can provide additional learning opportunities if it is positioned before the posttest (see also Nabei & Swain, 2002). In this study, stimulated recall was carried out after all the tasks were completed in order to avoid influencing the performance on Task 3. Veridicality concerns the accuracy of verbal reports in stimulated recall. Egi (2004) not only pointed out the memory decay problem associated with stimulated recall, but also the misinformation effect (p. 256). In the current study, the misinformation effect might be observed because Task 3, positioned between the treatment task (Task 2) and the stimulated recall, might have influenced the verbal reports; in other words, the L2 learners might have confused their thought processes in the treatment task with those in Task 3. To reduce the misinformation effect, two different tasks with different pictures were prepared for Task 2 and Task 3 even though both tasks were intended to elicit the same verbs as mentioned in the Tasks section.
4. Because the sample size was small and the data were not normally distributed, non-parametric procedures, Mann-Whitney tests, were selected (Hatch & Lazaraton, 1991, p. 270). Accordingly, instead of means and standard deviations, medians (*Mdns*) and interquartile ranges (*IQRs*) were used as a measure of central tendency and dispersion respectively. The interquartile ranges indicate the ranges of the middle 50% of the data.
5. Holm's sequential Bonferroni method is one of the methods to avoid making a Type I error, referred to as "the probability of rejecting a null hypothesis when it is true" (Green & Salkind, 2005, p. 417), in multiple comparisons. This method was used for this study according to Green and Salkind (2005, pp. 418-419). The family wise alpha for this study was set at the .05 level. Because three comparisons between the two groups were tested in this study, the three-step procedure was as follows: First, the smallest *p* value among the three comparisons was evaluated against the .016 level, obtained by dividing .05 by the number of comparisons; second, the second smallest *p* value was evaluated against the .025 level, obtained by dividing .05 by 2 (the number of comparisons – 1); and finally, the largest *p* value was evaluated against the .05 level. If the null hypothesis is not rejected at a particular step in the procedure, the analysis is stopped at that point. In that case, the remaining comparisons are assessed as nonsignificant.

References

- Adams, R. (2003). L2 output, reformulation and noticing: Implications for IL development. *Language Teaching Research*, 7, 347-376.
- Ayoun, D. (2001). The role of negative and positive feedback in the second language acquisition of the passé composé and imparfait. *The Modern Language Journal*, 85, 226-243.
- Carroll, S., & Swain, M. (1993). Explicit and implicit negative feedback: An empirical study of the learning of linguistic generalizations. *Studies in Second Language Acquisition*, 15, 357-386.
- Doughty, C. (1994). Fine-tuning of feedback by competent speakers to language learners. In J. Alatis (Ed.), *GURT 1993: Strategic interaction and language acquisition* (pp. 96-108). Washington D.C.: Georgetown University Press.
- Doughty, C., & Varela, E. (1998). Communicative focus on form. In C. Doughty & J. Williams (Eds.), *Focus on form in classroom second language acquisition* (pp. 114-138). Cambridge: Cambridge University Press.
- Doughty, C., & Williams, J. (1998). Pedagogical choices in focus on form. In C. Doughty & J. Williams (Eds.), *Focus on form in classroom second language acquisition* (pp. 197-261). Cambridge: Cambridge University Press.
- Dulay, H., Burt, M., & Krashen, S. (1982). *Language two*. NY: Oxford University Press.
- Egi, T. (2004). Verbal reports, noticing, and SLA research. *Language Awareness*, 13, 243-264.
- Egi, T. (2007a). Interpreting recasts as linguistic evidence: The roles of linguistic target, length, and degree of change. *Studies in Second Language Acquisition*, 29, 511-537.
- Egi, T. (2007b). Recasts, learners' interpretations, and L2 development. In A. Mackey (Ed.), *Conversational interaction in second language acquisition* (pp. 249-267). Oxford: Oxford University Press.
- Ellis, R. (1987). Interlanguage variability in narrative discourse: Style shifting in the use of the past tense. *Studies in Second Language Acquisition*, 9, 1-20.
- Ellis, R. (2007). The differential effects of corrective feedback on two grammatical structures. In A. Mackey (Ed.), *Conversational interaction in second language acquisition* (pp. 339-360). Oxford University Press.

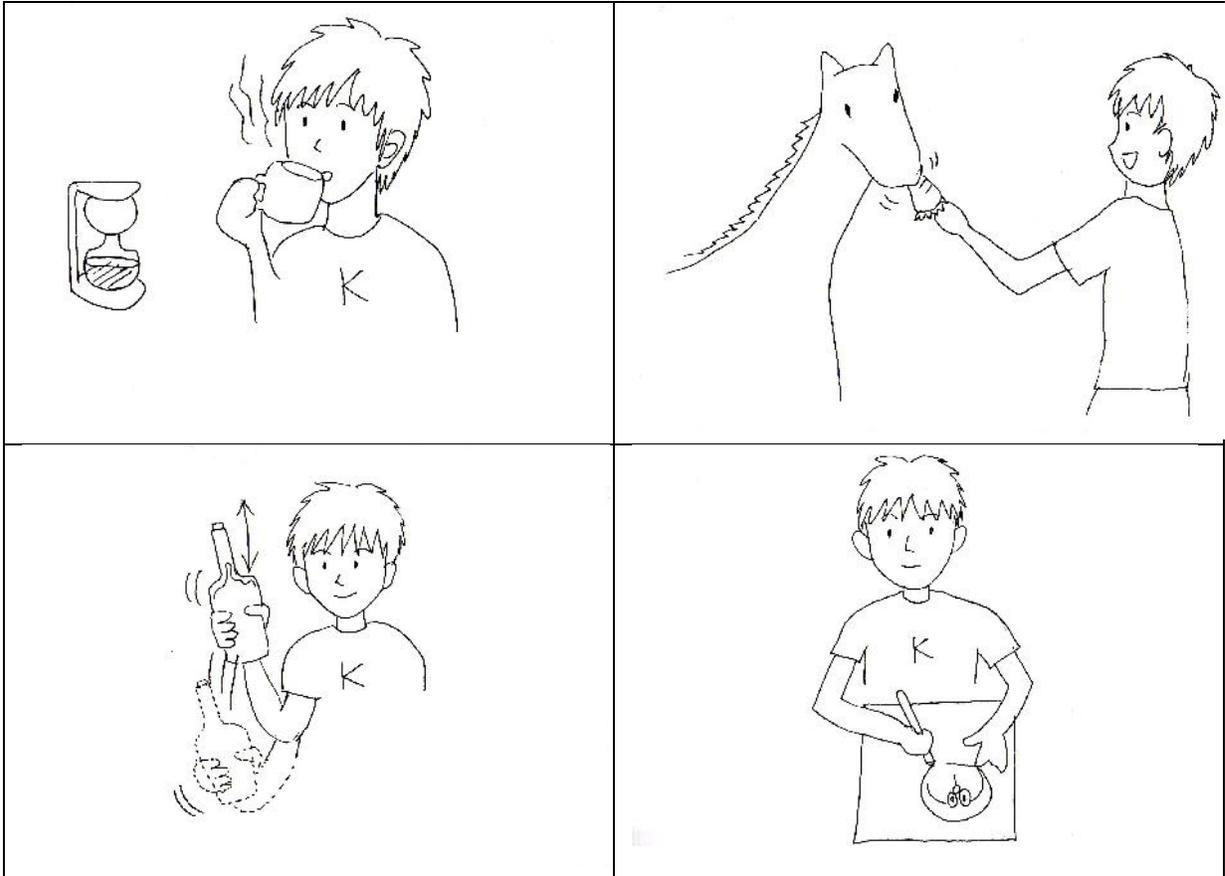
- Ellis, R., Loewen, S., & Erlam, R. (2006). Implicit and explicit corrective feedback and the acquisition of L2 grammar. *Studies in Second Language Acquisition*, 28, 339-368.
- Gass, S. M., & Lewis, K. (2007). Perceptions about interactional feedback: Differences between heritage language learners and non-heritage language learners. In Mackey A. (Ed.), *Conversational interaction in second language acquisition* (pp. 79-99). Oxford University Press.
- Gass, S. M., & Mackey, A. (2000). *Stimulated recall methodology in second language research*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Gass, S. M., & Mackey, A. (2007). Input, interaction, and output in second language acquisition. In B. VanPatten & J. Williams (Eds.), *Theories in second language acquisition* (pp. 175-199). Mahwah, NJ: Lawrence Erlbaum.
- Green, S. B., & Salkind, N. J. (2005). *Using SPSS for Windows and Macintosh: Analyzing and understanding data* (4th ed.). Upper Saddle River, NJ: Pearson Education Inc.
- Han, Z. (2002). A study of the impact of recasts on tense consistency in L2 output. *TESOL Quarterly*, 36, 543-572.
- Hatch, E., & Lazaraton, A. (1991). *The research manual: Design and statistics for applied linguistics*. New York: Newbury House.
- Ishida, M. (2004). Effects of recasts on the acquisition of the aspectual form –te i- (ru) by learners of Japanese as a foreign language. *Language Learning*, 54, 311-394.
- Iwashita, N. (2003). Negative feedback and positive evidence in task-based interaction: Differential effects on L2 development. *Studies in Second Language Acquisition*, 25, 1-36.
- Leeman, J. (2003). Recasts and second language development: Beyond negative evidence. *Studies in Second Language Acquisition*, 25, 37-63.
- Lightbown, P. M., & Spada, N. (1993). *How languages are learned*. Oxford: Oxford University Press.
- Loewen, S., & Nabei, T. (2007). Measuring the effects of oral corrective feedback on L2 knowledge. In A. Mackey (Ed.), *Conversational interaction in second language acquisition* (pp. 361-377). Oxford University Press.

- Long, M. H. (1996). The role of the linguistic environment in second language acquisition. In W. C. Richie, & T. K. Bhatia (Eds.), *Handbook of second language acquisition* (pp. 413-468). San Diego, CA: Academic Press.
- Long, M. H. (2007). *Problems in SLA*. Mahwah, NJ: Lawrence Erlbaum.
- Long, M. H., Inagaki, S., & Ortega, L. (1998). The role of implicit negative feedback in SLA: Models and recasts in Japanese and Spanish. *The Modern Language Journal*, 82, 357-371.
- Lyster, R. (1998a). Recasts, repetition, and ambiguity in L2 classroom discourse. *Studies in Second Language Acquisition*, 20, 51-81.
- Lyster, R. (1998b). Negotiation of form, recasts, and explicit correction in relation to error types and learner repair in immersion classrooms. *Language Learning*, 48, 183-218.
- Lyster, R., & Ranta, L. (1997). Corrective feedback and learner uptake: Negotiation of form in communicative classrooms. *Studies in Second Language Acquisition*, 19, 37-66.
- Mackey, A. (1995). *Stepping up the pace: Input, interaction and interlanguage development: An empirical study of questions in ESL*. Unpublished doctoral dissertation, University of Sydney, Australia.
- Mackey, A. (1999). Input, interaction, and second language development: An empirical study of question formation in ESL. *Studies in Second Language Acquisition*, 21, 557-587.
- Mackey, A. (Ed.) (2007). *Conversational interaction in second language acquisition*. Oxford University Press.
- Mackey, A., Gass, S., & McDonough, K. (2000). How do learners perceive interactional feedback? *Studies in Second Language Acquisition*, 22, 471-497.
- Mackey, A., & Goo, J. (2007). Interaction research in SLA: A meta-analysis and research synthesis. In A. Mackey (Ed.), *Conversational interaction in second language acquisition* (pp. 407-462). Oxford University Press.
- Mackey, A., & Philp, J. (1998). Conversational interaction and second language development: Recasts, responses, and red herrings? *The Modern Language Journal*, 82, 338-356.

- Nabei, T., & Swain, M. (2002). Learner awareness of recasts in classroom interaction: A case study of an adult EFL student's second language learning. *Language Awareness, 11*, 43-63.
- Nassaji, H. (2007). Elicitation and reformulation and their relationship with learner repair in dyadic interaction. *Language Learning, 57*, 511-548.
- Nicholas, H., Lightbown, P. M., & Spada, N. (2001). Recasts as feedback to language learners. *Language Learning, 51*, 719-758.
- Nobuyoshi, J., & Ellis, R. (1993). Focused communication tasks and second language acquisition. *ELT Journal, 47*, 203-210.
- Oliver, R. (1995). Negative feedback in child NS/NNS conversation. *Studies in Second Language Acquisition, 17*, 459-481.
- Panova, I., & Lyster, R. (2002). Patterns of corrective feedback and uptake in an adult ESL classroom. *TESOL Quarterly, 36*, 573-595.
- Pica, T. (1983). Methods of morpheme quantification: Their effect on the interpretation of second language data. *Studies in Second Language Acquisition, 6*, 69-78.
- Roberts, M. A. (1995). Awareness and the efficacy of error correction. In R. Schmidt (Ed.), *Attention & awareness in foreign language learning* (pp. 163-182). Honolulu, HI: University of Hawaii at Manoa.
- Robinson, P. (1995). Attention, memory, and the "noticing" hypothesis. *Language Learning, 45*, 283-331.
- Sakai, H. (2004). Roles of output and feedback for L2 learners' noticing. *JALT Journal, 26*, 25-54.
- Saxton, M. (1997). The contrast theory of negative input. *Journal of Child Language, 24*, 139-161.
- Saxton, M. (2000). Negative evidence and negative feedback: Immediate effects on the grammaticality of child speech. *First Language, 20*, 221-252.
- Schmidt, R. (1990). The role of consciousness in second language learning. *Applied Linguistics, 11*, 129-158.
- Simard, D., & Wong, W. (2001). Alertness, orientation, and detection: The conceptualization of attentional functions in SLA. *Studies in Second Language Acquisition, 23*, 103-124.

- Swain, M. (1985). Communicative competence: Some roles of comprehensible input and comprehensible output in its development. In S. M. Gass & C. Madden (Eds.), *Input in second language acquisition* (pp. 235-253). Rowley, MA: Newbury House.
- Swain, M. (1993). The output hypothesis: Just speaking and writing aren't enough. *The Canadian Modern Language Review*, 51, 158-164.
- Swain, M. (1995). Three functions of output in second language learning. In G. Cook & B. Seidlhofer (Eds.), *Principle and practice in applied linguistics: Studies in honor of H. G. Widdowson* (pp. 125-144). Oxford University Press.
- Swain, M. (1998). Focus on form through conscious reflection. In C. Doughty & J. Williams (Eds.), *Focus on form in classroom second language acquisition* (pp. 64-81). New York: Cambridge University Press.
- Takashima, H. (1995). *A study of focused feedback, or output enhancement, in promoting accuracy in communicative activities*. Unpublished doctoral dissertation presented to Temple University Japan.
- Tomlin, R., & Villa, H. (1994). Attention in cognitive science and second language acquisition. *Studies in Second Language Acquisition*, 16, 183-203.
- Van den Branden, K. (1997). Effects of negotiation on language learners' output. *Language Learning*, 47, 589-636.

Appendix
Sample Pictures for Task 2 (Ken's Story)



Book Review

Language Teacher Research in Africa

Thomas S. C. Farrell and Leketi Makalela (Eds). Virginia: Teachers of English to Speakers of Other Languages, Inc., 2009. Pp. viii + 124.

Reviewed by Iris F. Levitis

Max Planck Institute for Demographic Research

Rostock, Germany

Of interest to English teachers, preservice teachers, and TESOL students and professionals, *Language Teacher Research in Africa* brings together eight articles written by English language teachers in Africa which can be grouped according to three pedagogical foci: oral language (chapters 3, 4), writing skills (chapters 6, 7), and academic writing (chapters 2, 5, and 8).

The first pedagogical area, oral language skills, is explored by Kadenge, Mabugu, and Dube (chapter 3), wherein they pose the question: What pronunciation difficulties do Shona speaking undergraduates have when learning English? To address this question the researchers established an inventory of pronunciation errors. Corrective feedback provided to students was based on this inventory. The researchers concluded that giving students corrective feedback tailored to the errors that they committed resulted in more native-like pronunciation. Kasanga (chapter 4) investigates the value of incorporating

debate into an English for Academic Purposes (EAP) course. Using ethnographic data collection and reflection, Kasanga studied the impact of debate as an academic exercise. This exploratory study draws on teacher and undergraduates assessment and determines that carefully planned debate can be a useful method of practicing English.

Another pedagogical area explored is basic writing skills. Njoroge Ndung'u's article (chapter 6) explores the importance of vocabulary instruction in a secondary school. Students completed cloze exercises and wrote sentences using twenty synonymous nouns to determine their knowledge of the word's lexical information. Results indicated that vocabulary instruction, to be successful, requires more emphasis on lexical role and grammatical knowledge than it currently receives. Ojwang' (chapter 7) addresses the issue of fossilized spelling errors in undergraduate student writing. Ojwang' analyzes undergraduate essays and dictations to classify types of spelling errors and explores different learning methods adopted by students. Additionally, students completed surveys about their experiences with spelling instruction. The article concluded that current spelling instruction is inadequate and leads to later writing problems.

The issue of academic writing is tackled by three of the articles in this volume. Banda (chapter 2) provides an ethnographic exploration of ten black African university students, their personal history of English education, and their subsequent literacy practices. Student writing samples as well as individual and group interviews were gathered and analyzed. Banda concluded that one legacy of apartheid is the inadequate academic writing skills of the current generation of rural black South Africans. Makalela (chapter 5) asks what impact form-focused feedback has on the writing of first-year university students. To determine what aspects of writing were problematic, Makalela analyzed student essays and quantified grammatical errors and categorized the errors that made it

possible to focus on providing students with relevant feedback. Makalela concluded that this dynamic approach to feedback resulted in better student writing. Elizabeth Steinbach (chapter 8) examines the relationship of pre-service English teachers to writing. Steinbach uses a combination of ethnographic and experimental methods in order to survey the writing skills and the beliefs of participants enrolled in a process-writing course. Steinbach concluded that the technique of process writing did not match this particular context and concludes that writing must be taught in context to be effective.

The strengths of this book are the accessibility of the research and the range of methodologies employed (e.g ethnography, quantitative methods, and interventionist). The one, flaw, however, is that it suffers from a lack of geographical breadth. The claimed geography covered is the African continent, but this is misleading since all but three articles (one from Zimbabwe, and two from Kenya) are from South Africa. The overrepresentation of South Africa causes the educational problems to appear to have more significance in this volume than if additional countries had been included. The issues faced by English teaching researchers in other African countries remain unexplored. Despite this flaw, this book is valuable both for its insight into problems faced by African teacher researchers as well as for the methods used to solve these problems.

Book Review

An Introduction to Conversation Analysis

A. J. Liddicoat. London: Continuum, 2007. Pp. ix + 319.

Reviewed by Hayriye Kayi

University of Texas, Austin

USA

An Introduction to Conversation Analysis by Anthony Liddicoat is an excellent source which provides a comprehensive overview of conversation analysis (CA). Although the book is intended for undergraduate and graduate students of sociolinguistics, discourse analysis, sociology, and applied linguistics, anyone interested in CA in talk analysis will find something in this text, as the author adopts a broad perspective to CA by saying “Conversation analysis legitimately investigates all areas of socially motivated talk.” (p. 6).

The book consists of ten chapters. After a very brief introduction to conversation and CA in the first chapter, the author traces the history of CA and locates it as a method for studying interaction. In chapter 2, Liddicoat provides a detailed presentation of how conversation is represented in the form of transcripts (e.g. words, prosody, speech sounds, contiguous and simultaneous talk, pauses, problems of hearing and comprehension, non-

verbal elements of talk, and translation). Each feature is introduced with a sample transcript (e.g. stress, long sounds, quiet talk, etc).

The next five chapters cover three broad areas of conversational organization: turn taking, sequence organization and expansion, and repair. Turn taking, which the author explains is both context-free and context-sensitive, is discussed in chapters 3 and 4 with a specific focus on possible models and features of turn-taking, turn allocation, turn taking errors and violations, and gaps and overlaps. How turns at talk become sequences and how these sequences are coherently expanded are the main focus of the chapters 5 and 6. The following chapter, chapter 7, covers repair system which deals with breakdowns in conversation. The author concludes that repair is “an interactionally sensitive mechanism which is constrained by social as well as linguistic considerations” (p. 212).

The rest of the book is devoted to three areas of conversational difficulty. Chapters 8 and 9 address how people start and end a conversation. The final chapter, chapter 10, focuses on story-telling with a specific focus on beginning and ending stories, story structure, second stories, and stories of shared experiences. The author argues that “stories are not simply told by tellers who in some sense take a long turn at talk; rather, they are collaboratively achieved by the participants through and in the telling of stories” (p. 302).

An Introduction to Conversation Analysis, with its comprehensible yet accessible overview of CA, is an invaluable addition to the existing literature. There are a few places which might be challenging for the novice reader as he/she engages the terms and skills applicable to this field. Acknowledging this, the author offers helpful suggestions along the way. Accepting this, readers will find the book to be to the point, engaging, and an excellent reference for applying conversational analysis in any social setting.

Book Review

Feedback in Second Language Writing: Contexts and Issues.

Ken Hyland & Fiona Hyland (Eds). Cambridge: Cambridge University Press, 2006.

Pp. xi + 291.

Aysha Viswamohan

Indian institute of Technology Madras

Chennai, India.

ESL teachers of writing often face several challenges in their work and providing appropriate feedback is a major issue. *Feedback in Second Language Writing* is a comprehensive work designed to assist scholars in this direction. Divided in three sections with a total of fourteen chapters, the three sections are appropriately titled, “Situating Feedback: Sociocultural Dimensions,” “Shaping Feedback: Delivery and Focus Dimensions,” and “Negotiating Feedback: Interpersonal and Interactional Dimensions.”

The book begins with “Contexts and issues in feedback on L2 writing: An introduction.” This is co-authored by the editors and includes teacher feedback to L2 writers, interactive writing conferences, the nature of peer interactions, automated or computer-mediated feedback and contexts and issues of feedback. The next chapter (chapter 2), “Sociocultural theory: A framework for understanding the socio-cognitive dimensions of

peer-feedback” by Villamil and de Guerrero gives an overview of sociocultural theory and its key concepts. Nelson and Carson’s “Cultural issues in peer response: Revisiting ‘culture’” provides a detailed review of various issues related to culture in peer response. The discussions of the key issues by various writers are critically questioned. And in chapter 4, the last from Section I, “Appropriation, ownership, and agency: Negotiating teacher feedback in academic settings,” Tardy accounts for culture from different sociocultural, sociohistoric and sociopolitical perspectives.

Chapter 5, “Does error feedback help student writers? New evidence on the short- and long-term effects of written error correction” by Ferris, starts the second section of the book. The author discusses teacher feedback strategies and their effectiveness in improving L2 undergraduate students’ writing, which leads to the next chapter by Ware and Warschauer, “Electronic feedback and second language writing.” In this chapter, the writers observe that electronic feedback can be used for developing metacognitive and metalinguistic awareness among students. The next chapter of this section, Milton’s “Resource-rich web-based feedback: Helping learners become independent writers,” describes methods for giving students access to online resources for feedback on lexicogrammatical errors. Hamp-Lyons in chapter 8, “Feedback in portfolio-based writing courses”, states that a portfolio is a collection of students’ written work over a semester or year, and she endorses their efficacy. “Students and research: Reflective feedback for I-search papers” is the next chapter. In the first part, the author, Johns, discusses ways to improve learner autonomy through carefully scaffolded activities; in the second, she notes that I-Search is a tool to encourage continuous student reflection and self-evaluation.

Section III of the book begins with the chapter 10, “Feedback and revision in second

language writing: Contextual, teacher, and student variables” by Goldstein. Through her two case studies, the author advocates the need to look at each student and his or her context individually. The editors come together again in chapter 11, “Interpersonal aspects of response: Constructing and interpreting teacher written feedback.” Here, they point out that teachers’ feedback is “a response to a person rather than to a script” (p 206). The next chapter, “Formative interaction in electronic written exchanges: Fostering feedback dialogue” is by Hewings and Coffin. Analyzing feedback in three different tutor groups in asynchronous computer-mediated communication, the authors argue that tutor participation is required but without playing the dominant role. Weissberg’s “Scaffolded feedback: Tutorial conversations with advanced L2 writers,” explains Scaffolding as “the verbal support provided to the learner by the tutor that enables the learner to complete a new task” (p 247). In the final chapter, ““You cannot ignore”: L2 graduate students’ response to discipline-based written feedback”, Ilona Leki emphasizes that L2 students expect to have more feedback, although they remain silent during interaction, and this aspect must be understood by the disciplinary teachers.

The range of articles in the book offers several scholarly perspectives on the much needed skill of writing. Thus *Feedback* is a useful tool for teachers/researchers of writing as it provides in-depth awareness of the theory and praxis of the skill. It is also a valuable addition to the corpora of writings on feedback for teachers and researchers in this field.

Book Review

Diagnosing Foreign language Proficiency: The Interface Between Learning and Assessment

Alderson, Charles J. London: Continuum, 2006. Pp. xii + 284.

Deepti Gupta

Panjab University

Chandigarh, India

This volume is mainly an account of the DIALANG project. Its description of the project is very thorough and, in places, quite critical and objective. The reader gets acquainted with the project while getting updated on self assessment in language development.

DIALANG is a European project for the development of diagnostic tests in 14 European languages, made available on the Internet free of charge. It is computer-based and Internet-delivered and offers separate tests for reading, writing, listening, grammatical structures and vocabulary from beginner to advanced levels. It directs users so that they can assess their level of proficiency before offering them tests in the area selected; finally giving them feedback on their performance and advice to improve their proficiency.

Alderson was the scientific coordinator of the DIALANG project team and in the seventeen chapter volume under review, he presents a rationale for, description of, and

advantages of diagnostic tests in language learning and teaching. He uses DIALANG as an instance of such path-breaking pedagogy.

The first eight chapters describe the DIALANG project, devoting a chapter to each aspect. Chapter one examines the teaching/testing interface, building up a case for diagnostic testing in language teaching and learning by describing its advantages. The second chapter looks at the literature available and the various varieties of diagnostic testing created so far. It then concludes by making a case for computer-based diagnostic testing. Chapter three introduces the project and describes the steps that the ALS (Administration Language System) is composed of on the project website. The chapter also guides the reader through the process of using this diagnostic tool. The fourth chapter traces the history of the project while clearly stating that it is yet to stand the test of time. Chapter five presents the statistics collected so far to prove its efficacy. The sixth chapter examines the ticklish issue of standard-setting and how the procedures of the test tackle it. Chapter seven describes the vocabulary size placement test of the project. A noteworthy aspect of DIALANG is that it works equally smoothly as a diagnostic tool in fourteen European languages and across reading, listening, writing, vocabulary and grammar. It works as a supranational system, so far free of cost. Perhaps these benefits arise out of the inclusion of experts from all 14 language areas. Chapter eight discusses the role of feedback and self-assessment in learners' profile-development to round off the description. It also examines closely the whole paradigm of self-assessment in the project and how it works. The strength of the project is the empirical evidence provided by way of data collected in every chapter of the volume under review.

Having established a comprehensive background to DIALANG, the next five chapters examine the five areas of language proficiency tested by the self-assessment system:

reading, listening, writing, grammar, and vocabulary. A common framework supports each chapter, aiming towards an ease in readability. Every chapter first considers the construct of the area under discussion and describes the aspect tested in DIALANG; second, the test specifications and their target; third, the individual variables in the area and how they influence development. To conclude, each chapter looks at the pending requirements of dependable assessment in the area. The chapters on grammar and vocabulary take into account the relationship between the macro and micro skills of language as well.

The last four chapters (14 – 17) establish the dynamism quotient of DIALANG. The place of feedback and advice in the project is considered, quoting extensively from the available data. The essentially experimental nature of the project is defined in terms of the items used in it, with tasks from its website to give a feel of the test items. The pedagogic experiments used in self-assessment in the project and a detailed discussion of the future of diagnostic testing and its relationship with foreign language development make up the last chapter.

This volume is a comprehensive account of the project and is informative. The reader's expectation of a descriptive and detailed account of DIALANG is fulfilled but a negative aspect is that it lacks a critical review of foreign language proficiency tests. The information is provided without any attempt at evaluation or judgement. Another difficulty is that the text sometimes becomes rather dense and challenging in the representation of facts, but perhaps this is a feature of cyber territory.

Publisher

Asian EFL Journal

Chief Editor

Dr. Roger Nunn

The Petroleum Institute

Abu Dhabi, UAE

Senior Associate Editor

Dr. Mike Cribb
Senior Lecturer in English
Coventry University
UK

Senior Associate Editor

Dr. John Adamson
University of Niigata Prefecture
Japan

Associate Editors

Dr. Esmat Babaii

University for Teacher
Education
Iran

Mr. Roger Cohen

Defense Language Institute
San Antonio Texas
USA

Dr. Eva Bernat

School of Education
University of New South Wales
Australia

Dr. Shaoan Zhang.

College of Education
University of Nevada,
Las Vegas
USA

Dr. Sivakumar

Sivasubramaniam
The Petroleum Institute
Abu Dhabi
U.A.E.

Dr. Jeremy Cross

English Language and
Literature National Institute of
Education Singapore

Dr. Ahmet Acar

Dokuz Eylül University
Turkey

Dr. Aly Anwar Amer

Sultan Qaboos University
College of Education
Sultanate of Oman

Mr. Neil Heffernan

Ehime University
Matsuyama
Japan

Dr. Xiuping Li

Newcastle University
UK

Stuart D. Warrington

Waseda University
Tokyo
Japan

Dr. John A. Unger

Department of Transitional
Studies
University of New Mexico
New Mexico
USA

Associate Production Editors

Alison Smith
United Arab Emirates University
UAE

Robert Baird
Southampton University
UK

Dr. Toshinobu Nagamine
Prefectural University of Kumamoto
Japan

Dr. Wen-chi Vivian Wu
Providence University
Taiwan

Dr. Aradhna Malik
Vinod Gupta School of Management
Indian Institute of Technology
India

Dr. Margaret Hearnden
Department of Educational Studies
University of York

Editors

Naoki Fujimoto-Adamson
University of Niigata Prefecture
Japan

Mr. John Baker
Chung Hwa College of
Medical Technology
Taiwan, China

Mr. Peter Ilic
Asia University
Tokyo, Japan

Dr. Kota Ohata
International Christian University,
Tokyo, Japan

Dr. Malcolm Benson
Hiroshima Shudo University, Japan

Dr. Ming Cheung
Department of English and Communication
City University of Hong Kong
Hong Kong

Professor Chen Yong
School of Foreign Languages
Chongqing University
China

Dr. John Thurman
Hokkaido University,
Japa

Dr Jia Li
Ontario Institute for Studies in Education
University of Toronto
Canada

Dr. Yu Ling Cheun
National University of Education
Taiwan, China

Dr. Radha M.K. Nambair
School of Language Studies & Linguistics
Universiti Kebangsaan Malaysia
Malaysia

Dr. Deepti Gupta
Department of English
Panjab University
Chandigarh
India

Annie Hong Qin Zhao
Education Department
University of Bath
UK

Dr. Toshiyuki Takagaki,
Faculty of Art & Culture
Onomichi University,
Japan

Dr Aysha Viswamohan
Dept of Humanities &
Social Sciences
Indian Institute of Technology
Madras

Dr. Budsaba Kanoksilapatham
Silpakorn University
Thailand

Dr Jane Mok Fung Yee
Department of English,
Hong Kong Polytechnic University
Hong Kong

Dr. Ahmed Shakir El-Kilabi
Department of English,
Nizwa College of Education,
Oman

Dr. Michael Thomas
Nagoya University of Commerce &
Business,
Japan

Dr Nooreiny Maarof
Dept of Methodology & Educational
Practice
Faculty of Education
Universiti Kebangsaan
Malaysia

Dr. Benedict Lin
Centre for Research in
Applied Linguistics
Division of English Studies
University of Nottingham
Ningbo
China

Dr. Arda Arikan
Hacettepe University
Faculty of Education
Turkey

Nolan Weil
Intensive English Language Institute
USA

Meenakshi Raman, Ph.D.
BITS, Pilani,
Rajasthan,
India

Dr. Mohammad Ali Salmani-Nodoushan
University of Zanjan
Iran

Mr. David Brown
The University of the Thai Chamber of
Commerce
Thailand

Dr. Tan Bee Tin
Department of Applied Language Studies
and Linguistics
University of Auckland
New Zealand

Mr. David Litz
United Arab Emirates University
UAE

Dr. Joan Cutting
The Moray House
School of Education
The University of Edinburgh

Dr. Mohamed El-Okda
Dept. of Curriculum
College of Education
Sultan Qaboos University
Sultanate of Oman

Marcus Otłowski
Department of International Communication
Kochi University
Japan

Dr Bilal Kirkici
Faculty of Oriental Studies
Oxford University
UK

Dr. Shamala Paramasivam
Faculty of Modern Languages
& Communications
Universiti Putru
Malaysia

Veronica Wynne
Boulder Valley Schools
Boulder, CO,
USA

Will Baker
Southampton University
UK

Dr. Peter Petrucci
Massey university
New Zealand
homepage

Dr Stefanie Pillai
Director
Centre for Industrial Training & Relations
(CITRA)
University of Malaya

Dr. Keiko Sakui
Kobe Shoin Women University
Japan

Dr. John Thurman
Hokkaido University,
Japan

Maria Belen Diez-Bedmar
Department of English Studies
University of Jaén
Spain

David Hall
Macquarie University
N.S.W.
Australia

Dr. Berna Çoker Koloğlu
Dokuz Eylül University
Turkey

Dr. Pin-hsiang Natalie Wu
Chien-kuo Technology University
Taiwan, China

Huli Wang
School of Foreign Languages
Dalian University of Technology
China

Nat Carney
Kwansei Gakuin University
Japan

Farhad Mazlum Zavarag
Faculty of Literature and Foreign
Languages,
Tarbiat Moallem University,
Tehran,
Iran

Dr. Karin Zotzmann
University of Queretaro
Mexico

Dr. Ruth MH Wong
English Department
Hong Kong Institute
of Education

Dr. Elke Stracke
Faculty of Arts and Design
University of Canberra
Australia

Dr. Ying-Ying Kimberly Chuang
Department of Applied Foreign
Languages,
Cheng Shiu University,
Taiwan, China

Vander Viana
Catholic University of
Rio de Janeiro
Brazil

Jim. E. King
School of English Studies
University of Nottingham, UK

Sean Sutherland
King's College
London
UK

Stefanie Shamila Pillai
Director - Centre for Industrial
Training and Relations
Faculty of Languages & Linguistics
University of Malaya

Peter Burden (Ed.D)
Okayama Shoka University
Japan

Scott Menking
English Linguistics and
Literature Dept.
Shimane University
Japan

Dr Suganthi John
University of Birmingham
U.K.

Damien Rivers
Kanda University
of International Studies
Japan

Dr. Patrisius Istiarto Djiwandono
Education Dept'
Ma Chung University
Indonesia

Dr Huahui Zhao
PhD Bristol, UK.
China

- 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 Balcikanli**
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 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,
Nakhornpathom,
Thailand
- Dr. Zahariah Pilus**
International Islamic University
Malaysia
- 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 Fran Cisco State Uni
Project Manager
Hawaii Pacific University, USA
- Prof. Paul Nation**
Victoria University
New Zealand

Senior Advisors

<p>Professor Claire Kramsch University of California Berkeley, U.S.A.</p>	<p>Dr. Amy Tsui Faculty of Education University of Hong Kong</p>	<p>Dr. James P. Lantolf Centre for Language Acquisition Pennsylvania State University U.S.A.</p>
<p>Dr. Francis Mangubhai University of Southern Queensland Australia</p>	<p>Professor Robert Phillipson Faculty of Languages, Communication and Cultural Studies Copenhagen Business School Denmark</p>	<p>Dr. Alan Tonkyn The University of Reading Applied Linguistics Dept UK</p>
<p>Professor Terry Shortall Birmingham University Birmingham UK</p>	<p>Marc Helgesen Miyagi Gakuin Women's University Sendai, Japan</p>	<p>Dr. Robert David Carless Hong Kong Institute of Education Hong Kong</p>
<p>Dr. Jeong-Bae Son University of Southern Queensland Australia</p>	<p>Dr. Robert J. Dickey Gyeongju University Korea</p> <p>Dr. Luke Prodromou Leeds University UK</p>	<p>Dr. Yasuko Kanno Asst' Professor English University of Washington U.S.A.</p>

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>

APA Citation Style: <http://www.liu.edu/cwis/CWP/library/workshop/citapa.htm>

APA Style Workshop:

<http://owl.english.purdue.edu/workshops/hypertext/apa/index.html>

v) Keywords: All articles must include Keywords at the beginning of the article. List 4-6 keywords to facilitate locating the article through keyword searches in the future.

vi) Graphs and Charts - either in the body of the document or at the end. In certain cases, a graphic may not appear in the text of the web version of the *Asian EFL Journal* but a link to the graphic will be provided.

vii) Paragraphs. Double space between paragraphs. Indent the beginning of each paragraph with three strikes of the space bar except those immediately following a heading, quotation, example, figure, chart or table. Do not use the tab key.

viii) Keep text formatting (e.g., italics, bold, etc.) to the absolute minimum necessary. Use full justification. All lines to be against Left Hand Side Margin (except quotes - to be indented per APA style).

ix) Abstract

The abstract should contain an informative summary of the main points of the article, including, where relevant, the article's purpose, theoretical framework, methodology, types of data analysed, subject information, main findings, and conclusions. The abstract should reflect the focus of the article.

x) Graphs – to fit within A4 size margins (not wider)

Thank you for your cooperation.

asian_efl_journal@yahoo.com

Please include the following with your submission:

Name

School affiliation

Address

E-mail

Phone number

Brief Bio Data noting history of professional expertise

Qualifications

An undertaking the work has not been published elsewhere

Abstract

Any questions regarding submission guidelines, or more detailed inquiries about less common citation styles, may be addressed to the Editorial Board.

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:

asianefljournals@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>.
