Table of Contents:

Foreword by Wen-Chi Vivian Wu ................................................................. 6-9

1. Rasheed Al-Jarrah & Sayyah Al-Ahmad .............................................. 10-40
   - An Optimality-theoretic Account of Corrective Feedback in Process Writing

2. Mohamed Ridha Ben Maad ............................................................... 41-62
   - Towards a Discerning Image of Learners’ Estimation of Task Difficulty and
     Motivation: Goal Orientations Highlighted

3. Julian Cheng Chiang Chen ................................................................. 63-98
   - Designing a Computer-mediated, Task-based Syllabus: A Case Study in a
     Taiwanese EFL Tertiary Class

4. Katie Kim ............................................................................................ 99-177
   - L2 Phone-based Interaction (PBI) and Development of Communicative
     Competence: A Case Study of an Adult’s English Learning in EFL Context

5. Elizabeth Root ...................................................................................... 178-213
   - Participation in and Opposition to the Ideology of English in South Korea:
     Insights from Personal Narratives

6. Yo In’nami & Rie Koizumi ................................................................. 214-265
   - A Quantitative Reanalysis of Data on the Structure of L1 and L2 Language Ability
     in Multitrait-Multimethod Studies

7. Zahra Fotovatnia and Alireza Barouni Ebrahimi .............................. 266-300
   - Word Knowledge: Aspects, Viewpoints and Performances
8. Shizhou Yang  .............................................................. 301-338

- “A Small Clean Warm Plastic”: Understanding EFL Autobiographical Writing and Identity
Book Reviews

1. *Networked Collaborative Learning: Social Interaction and Active Learning (1st ed.)*
   Guglielmo Trentin
   Reviewed by Weihong Wang .................................................. 339-341

2. *Uncovering EAP: How to Teach Academic Writing and Reading*
   Sam McCarter and Phil Jakes
   Reviewed by Sally Ashton-Hay ............................................ 342-343

3. *Language and Learning in the Digital Age*
   J. P. Gee & E. R. Hayes
   Reviewed by Servet Çelik and Hasan Sağlamel ....................... 344-345

4. *Medical Interpreting and Cross-cultural Communication*
   Claudia V. Angelelli
   Reviewed by Pi-Chi Han ....................................................... 346-347
September 2012 Foreword
by Wen-Chi Vivian Wu

I would like to welcome all Asian EFL Journal readers to the September 2012 issue of AEJ. This third issue of AEJ consists of eight articles and four book reviews, covering a wide range of the most recent issues worthy of notice in the field of English as a second/foreign language teaching and learning, including computer assisted language learning, corrective feedback writing, autobiographical writing, learner motivation, and phone-based communicative competency, etc. In addition to the various topics, we have also included contributors from a broad variety of geographic locations, spanning Asia, Africa, and North America.

In the first paper, Rasheed Al-Jarrah and Sayyah Al-Ahmad conducted a theory-driven analysis to locate composition writing errors of English majors in both form and content in the Jordanian EFL context. The authors claim that by situating error correction within the framework of Optimality Theory, a constraint-based approach can provide EFL teachers with the mechanics to revise their pedagogy in teaching writing skills. This article can serve as guidance for those EFL writing instructors who wish to incorporate feedback correction in their instruction.

Mohamed Ridha Ben Maad examined in his study one of the most significant and most frequently researched issues in the EFL field, which is motivation. Based on an experimental study, goal orientation was treated by the author as a researchable learner difference factor, which can influence acquisition of second language learners in terms of language tasks and motivation. The findings suggest a need to revisit the established task research format that confines task variation to design and sequencing factors and eclipses individual differences.
The third article written by Julian Chen emphasizes the need to adopt a more vivid and real-life instruction to boost the motivation of Taiwanese EFL learners. Chen’s study mirrored the six key components of a Tasked-based Language Teaching (TBLT) design as well as capitalized on the cultural competence of Taiwanese students in L1 as a springboard for channeling their L1 knowledge into L2 production. Blogging was adopted by the author as a platform for students to jointly construct their sites, interact with peers, transfer their background knowledge from L1 to L2 in the process of task completion, and to develop their communicative and cross-cultural competence in a collaborative virtual community. Pedagogical implications for English teachers who are interested in adopting CMC task-based instruction in EFL contexts were also provided by the author.

To investigate the effectiveness of phone-based L2 interaction (PBI), which has rarely been researched in the current literature, Professor Kim studied a Korean EFL learner by engaging him into daily telephone conversations exclusively in English. By doing so, the author examined the development of an EFL learner’s L2 communicative competence and the increase of the learners’ willingness to communicate (WTC). In addition to assessing communicative competence, grammatical competence, sociolinguistic competence, discourse competence, and strategic competence, the learner’s changes in affective variables, such as motivation, L2 anxiety, and self-confidence, were taken into consideration. The result of this study suggested that PBI can be effective for developing low-WTC EFL learners’ communicative competence and for increasing those learners’ WTC.

Professor Root investigated how the ideology of English (defined as a set of beliefs about the necessity of English language skills) could influence English language classroom interactions in South Korea by studying 26 Korean university students and their 27 native-English speaking teachers (NESTs). Six acts of participation in, and seven acts of opposition to, the ideology of English, as well as six acts of participation and five acts of awareness were identified, respectively, from the students and the teachers. Critical
reflection was proposed by the author about how these acts influence classroom interactions between Korean students and NESTs.

To understand the structure of L1 and L2 language ability, Yo In'nami and Rie Koizumi conducted an extensive literature search to collect multi-trait, multi-method (MTMM) studies. Empirical evidence showed that the unitary and higher-order models best defined both L1 and L2 ability. The support for the unitary model was surprising but consistent with the findings of Davidson (1988). Moderator variable analyses failed to identify clear relationships between the examined models and moderator variables.

Zahra Fotovatnia and Alireza Barouni Ebrahimi in their study analyzed learner and teacher beliefs about depth of vocabulary knowledge in L1 and L2 to see if and to what extent these aspects were taken as important, and whether there was any relationship between the ratings of these aspects in L2 and the learners’ actual performances. The findings might help learners and teachers more easily recognize possible gaps in their depth of vocabulary knowledge and their beliefs about it in accordance with a sound theoretical framework.

The final article of this issue deals with the importance of autobiographical writing. Through autobiographical writing, the participant also became a more confident and expressive writer in English. These findings suggest that EFL autobiographical writing in an appreciative writing group may position struggling writers as knowers and capable learners and contribute to rich understandings of their social identities.

I hope you enjoy reading all the articles contained in this September 2012 issue and find them not only interesting but also stimulating. It is also the production editor’s hope that this issue will help provide new insights into conducting future studies for researchers, and will trigger fresh ideas about innovative English pedagogy for practitioners in the EFL field, both of which will, accordingly, contribute to the ever-changing nature of global English Language instruction. Finally, I would like to take this opportunity to
express my gratitude to the contributors as well as the reviewers of this issue for all the work you have put into to make this issue possible.
An Optimality-theoretic Account of Corrective Feedback in Process Writing

Rasheed Al-Jarrah, & Sayyah Al-Ahmad

Yarmouk University, Jordan

Bio Data
Rasheed Al-Jarrah is an Associate Professor of English, currently teaching English language and Linguistics in the English Department and Language Center at Yarmouk University/Jordan. I graduated from Ball State University – Muncie/Indiana with a PhD in Applied Linguistics in 2002. Areas of research interests include language learning, second language writing, optimal solutions, language testing, contrastive and discourse analysis studies, and Qur'anic studies.

Sayyah Al-Ahmad is an Assistant professor of English, teaching English Language and Linguistics in the English Department and Language Center at Yarmouk University/Jordan. He graduated from Indiana University of Pennsylvania/USA with a PhD in Rhetoric and Linguistics in 2003. His research interests include language teaching and learning, language acquisition, writing, and pragmatics.

Abstract
The present paper is a theory-driven analysis of English majors’ composition writing errors in the Jordanian EFL context over a period of one year. Concisely, it assesses corrective feedback (Truscott, 1996, 1999, 2007; Ferris, 1999a, 2004; Chandler, 2003; Bitchener et al. 2005; Bitchener, 2008; Guenette, 2007; Truscott & Hsu, 2008; Bruton, 2009; Alroe, 2011) in light of advancement in linguistic theory, namely Optimality Theory (Prince & Smolensky, 1993; McCarthy & Prince, 1993a, b). By situating error correction within the framework of Optimality Theory (henceforth OT), we show that a constraint-based approach provides us with the mechanics to revise our pedagogy in teaching writing skills. In particular, we hope to get to the level of detail to show (1) how to reset treatment priorities, and (2) which type of corrective feedback is most constructive at each stage of development. These can be very feasible objectives because OT proposes that structural requirements be presented as constraints (not rules) on surface representations in parallel fashion. In other words, OT refutes altogether the practice that structural requirements are presented as segmented rules in a particular and fixed order. Instead, it promotes the proposal that material presentation should show at least three things: (1) structural requirements are violable; (2) structural requirements oftentimes impose competing demands; and (3) settling the conflict between competing structural requirements in the most harmonic way requires relative ranking.
Keywords: Corrective Feedback, Multiple drafts, Writing Skills, Local and Global Errors, Optimality Theory, Language Learnability

Introduction
Research findings (Perl & Egendrof, 1979; Sommers, 1980; Faigly & Witte, 1981; Chandler, 2003; Guenette, 2007; Truscott, 2007; inter alia) have shown that there are at least two main types of errors which trigger instructor’s feedback: local errors and global errors (more commonly called form and content errors, respectively). Current research on the efficacy of corrective feedback has, however, dealt with the various types of errors as if they were segmented, unfocused, and unsystematic. Little research is published on how various types of error feedback are interrelated.

What is worth noting here is that research findings have produced conflicting evidence not only for the efficacy of (or lack thereof) error correction (see Guenette, 2007), but also for the type of feedback. Lalande (1982), Frantzen (1995), and Ferris, Chaney, Komura, Robert, and McKee (2000) argued that indirect feedback could be more effective than direct feedback. However, Chandler (2003) found completely the opposite. Still, Ferris and Roberts (2001) found no significant differences between the two correction methods. In one of Chandler's (2003) experiments, it turned out that all correction methods (including coded and uncoded) were relatively equally effective. One interesting finding in Bitchener's (2008) study is that the direct feedback group who received no-meta linguistic explanation above each targeted error outperformed the direct feedback group who received meta linguistic explanation above each targeted error. However, Bruton (2009) believes that "there is no guarantee that the students revisions will be correct … if the error correction required is beyond the students' capacity." (p. 137).

As for content feedback and feedback on form (the two broad categories of errors) are concerned, Chandler (2003, p. 276) argues that teachers have available for them three patterns of response in the classroom. These are:

(1) the conventional response (giving feedback on content first and feedback on form in a later draft)
(2) the reverse pattern
(3) one in which form and content feedback are mixed
The only conclusion that Chandler comes up with is that, irrespective of the pattern, it should be superior to giving no feedback. Although Chandler never shows which of these response patterns would be more fruitful, especially in foreign language learning contexts, it could be inferred that she sees no harm in the conventional response (giving feedback on content first and feedback on form in a later draft). One main contribution of the present paper is to present argument in favor of the third pattern (one in which form and content feedback should be mixed). In the main body of this research paper, we hope to show that this claim is both theoretically and empirically motivated.

In order to do just that, 120 EFL second year English majors’ composition writings at a Jordanian state university were investigated over a period of one year in terms of global errors (e.g. content and organization) and local errors (e.g. grammatical and mechanical errors). By analyzing the corrections students have made to the final drafts from an OT perspective, the researchers sketch out a tutorial to show how treatment priorities can be reset to promote foreign language learners’ writings to higher levels of proficiency. The present study will then attempt to answer the following two research questions:

- What is the relationship between feedback on form and feedback on content?
- Why do students who receive combined feedback on form and content feedback make larger gains in grammatical accuracy?

In section 2 below, a review of the related literature about corrective feedback will be presented. The competing models of analysis are provided in section 3, with special focus on the basic premises and conventional notions of an OT analysis, the theoretical framework within which the present paper is coached. In the Methodology section, we will shed light on the context, participants, treatment and design of the study. In section 5, we attempt an OT discussion of the major findings, chiefly by linking the tutorial with the major findings of previous research.
Review of Related Literature

Corrective Feedback in a Process-Oriented Classroom

Writing researchers attempt to investigate writing and the composing process using a variety of techniques to explore what goes on in the mind of writers as they are involved in composing. However, a point worthy of mention here is that ESL/EFL studies of students’ response to their instructors’ written feedback on their essays used to be in single-draft rather than multiple-draft, contexts (Cf. Cohen, 1987; Cohen & Cavalcanti, 1990; McCurdy, 1992). With the advent of the process approach to writing (Zamel, 1976, 1983; Flower & Hayes, 1981; Rose, 1980), new insights have been gained in L1 and L2 composition research and pedagogy. Research has shown, contrary to previous perceptions, that writing is a recursive, cyclical, nonlinear, as well as a decision-making process in which writers discover new ideas as they look for meaning (Perl, 1980; Rose, 1980; Hughey, Wormuth, Hartfiel, and Jacobs, 1983; Zamel, 1983). That is, when writers compose, they do not follow a rigid linear line of planning, writing, and revising, but they follow a recursive pattern in which they move back and forth discovering, analyzing, and synthesizing ideas (Hughey, et al., 1983). They go back to reread, add, delete, or move some sentences or paragraphs forward or backward. Reid (1993) describes the writing process as “two steps forward, one step back,” a process of ebb and flow as they write (p. 8).

As this paper concerns itself with corrective feedback, we aim to shed more light on the revision process per se. In the history of the writing process, there have been two views of revision. The former view conceives of revision as a cleaning-up activity aimed at correcting surface or local errors in grammar, diction, punctuation, and spelling. However, the latter view has perceived revision as a recursive process in which writers go back and forth in the midst of generating a text looking for global errors in the meaning and correcting them, employing certain techniques such as deletion, addition, substitution, and rearrangements (Sommers, 1980; Flower & Hayes, 1981). In studying experienced and inexperienced writers, Faigly & Witte (1981) and Sommers (1980) have found that inexperienced writers tend to revise locally and their revision does not improve the text. Perl & Egendrof (1979) succinctly state inexperienced writers “frequently lose track of what they mean by becoming caught up in correcting details on
grammatical or logical ground before they have clearly sensed and expressed in some form what they mean to say” (p. 127).

On the other hand, experienced writers tend to revise globally and they achieve much progress in revising for meaning. Most studies (Cf. Semke, 1984; Fathman & Whalley, 1990; Ashwell, 2000) have shown that students receiving form and content feedback gained most. However, it is not as yet clear how to administer both types of feedback in the classroom. The proposed analysis suggests that feedback on form and content feedback be given simultaneously on each revision draft. The present paper is addressing this issue from a constraint-based perspective. In the next section, we aim to situate the proposed analysis in the context of other accounts.

**Competing Models of Analysis**

Due to the shift from product to process approaches, many lines of research have started to emerge. The first line, led by the expressive school, conceives of writing as self-expressive, self-discovery, and self-actualizing (Elbow, 1973; Macrorie, 1976; Faigly, 1986). The second line has been advocated by the cognitive school in which writing is perceived as a problem-solving process whereby researchers stress the importance of audience, purpose, and the situation for writing (Britton, Burgees, Martin, McLeod, and Rosen, 1975; Flower & Hayes, 1981). Both expressivists’ and cognitivists’ work is based upon the assumption that writing is mainly an individual act which takes place in the individual mind. Early writing process research studies that examine the psychological and behavioral perspectives of writing implement case study research as an instrument to explore the cognitive behavior of the writer as she is involved in composing. With the advent of Optimality Theory (McCarthy & Prince, 1993a, 1993b; Prince & Smolensky, 1993), another line of research has started to emerge.

The basic premise on which we capitalize in this paper is that Optimality Theory refutes simple linear sequencing in favor of parallel presentation, for language learning is an integrative process. As for the machinery, we hope to show that (1) errors are violations of structural requirements (presented as violable constraints not rules) on surface representation; (2) satisfying the demands imposed by some of these constraints is only possible at the expense of satisfying the demands of other constraints (i.e. some
constraints are higher ranking while some are lower-ranking); and (3) settling the conflict between competing violable constraints requires relative ranking.

We will shortly show that the mechanics of OT, in addition to revealing the simplicity behind systems that initially seem complex, increase our understanding of how language is best taught/learned in the classroom. However, before embarking on this, let us lay out the basic premises and conventional notations of an OT analysis.

**Optimality Theory: An Overview**

Although most of its reputation is due to being applied to phonology, Optimality Theory (McCarthy & Prince, 1993a, 1993b; Prince & Smolensky, 1993) has proved successful in accounting for phenomena beyond that scope. Optimality Theory proposes the concept of conflicting universal constraints to account for the belief that languages share the same set of universal principles, and yet allow for variation. It capitalizes on the claim that learners have available to them innate Universal Grammar, which is no more than a language-particular ranking of a set of universal, violable constraints. In this theory where “grammaticality equals optimality” (Prince & Smolensky, 1997, p. 1604) what the grammar of each language does is resolve the conflict between competing universal constraints by language-specific constraint rankings.

Unlike probably all previous rule-based theories (e.g. Chomsky & Halle, 1968; Liberman & Prince, 1977; McCarthy, 1979a, 1979b) which map inputs onto outputs through linearly ordered rewrite rules, optimality theory proposes that structural requirements (now preferably called constraints) are violable; and they apply to surface forms. As a basic premise of optimality, constraints can never be satisfied perfectly (i.e. no output form can meet the structural requirements of all constraints simultaneously). Constraint satisfaction can only be optimal; that is, some constraints are satisfied at the expense of others: a higher-ranking constraint is satisfied at the expense of a lower-ranking constraint. This is so because unlike the violation of a lower-ranking constraint, the violation of higher-ranked constraints yields ungrammatical output forms. However, because some output forms violate some lower-ranking constraints and yet surface as the winning surface forms, constraint violation is never binary; there are degrees of violation.
Due to the claim that “every language considers exactly the same set of options for realizing an input” (Prince & Smolensky, 1997, p. 1605), which is technically called *richness of the base*” (Tesar & Smolensky, 1996, p. 33), OT proposes two generators, namely Gen (generator), and Eval (evaluator). Whereas Gen acts on the lexicon to provide all possible output structures, Eval (evaluator) evaluates all candidate forms against the set of universal constraints uniquely. In other words, whereas Gen is language-independent, Eval is language–dependent. Archangeli (1997) sketches this out as in figure (1) below.

Figure (1)

OT demonstrates the idea of conflicting constraints as in tableau (1) (where the constraints are arranged horizontally, and the outputs vertically, and where * means violation of a constraint): Tableau (1)

<table>
<thead>
<tr>
<th>Constraints</th>
<th>W</th>
<th>X</th>
<th>Y</th>
<th>Z</th>
<th>ETC.</th>
</tr>
</thead>
<tbody>
<tr>
<td>AB</td>
<td></td>
<td>*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AC</td>
<td></td>
<td></td>
<td>*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AD</td>
<td></td>
<td></td>
<td></td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>BA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>*</td>
</tr>
<tr>
<td>ETC.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Tableau (1) makes clear that whereas one constraint acts against the realization of B after A as the winning configuration, the other constraints militate against the realization of anything other than B after A, and so on. What this basically means is that constraints (faithfulness and markedness) impose conflicting demands that cannot be met simultaneously. To resolve the conflict, OT makes use of the notion of dominance. That is, constraint X dominates Y if X is a higher ranking constraint and Y is a lower ranking constraint. Constraints are represented in the Tableau in a descending fashion from left to right. Violation of a higher-ranking constraint yields a sub-optimal form (indicated by the exclamation point ‘!’ in the cell corresponding to the violation). The pointing hand ⇩ indicates the winning configuration as shown in Tableau (2). By reversing the domination relationship between X and Y, the output configuration AC loses the competition as shown in Tableau (3).

Tableau (2)

<table>
<thead>
<tr>
<th>Input</th>
<th>X (*AB)</th>
<th>Y (*AC)</th>
</tr>
</thead>
<tbody>
<tr>
<td>AB</td>
<td>*!</td>
<td></td>
</tr>
<tr>
<td>⇩ AC</td>
<td></td>
<td>*</td>
</tr>
</tbody>
</table>

Tableau (3)

<table>
<thead>
<tr>
<th>Input: AB</th>
<th>Y (*AC)</th>
<th>X (*AB)</th>
</tr>
</thead>
<tbody>
<tr>
<td>⇩ AB</td>
<td></td>
<td>*</td>
</tr>
<tr>
<td>AC</td>
<td>*!</td>
<td></td>
</tr>
</tbody>
</table>

Finally, constraints vary in terms of content and formulation. As for content, constraints are partitioned into two major families: faithfulness constraints and markedness constraints. Whereas markedness constraints are structural requirements on the well-formedness of output forms, faithfulness constraints match the input forms with their corresponding output forms. In addition to keeping the symmetry between a linguistic form and its different realizations or variants (faithfulness), constraints compete to choose the least marked output form (markedness). Needless to say, both types of constraints are needed: whereas markedness constraints only favor output forms that are structurally simple, faithfulness constraints allow for cross-linguistic variations.
As for formulation, research has shown that constraints are subject to at least two major universal constraint schemata: the Generalized Alignment schema of McCarthy & Prince (1993b) and the Correspondence Theory of McCarthy & Prince (1993a). The schemata provide a vocabulary to state constraints which can roughly be formulated as in (2) below:

(2)  
Markedness \*VERB-s verbs are unsuffixed  
Faithfulness (3rd person noun, 3rd person verb)

One point worthy of mention here is that the markedness and faithfulness constraints formulated in (2) above impose conflicting demands. Whereas the markedness constraint militates against the realization of “s” at surface form, the faithfulness constraint requires the realization of this “s” such that the output matches the input. In order to resolve this classic antagonistic relationship between markedness and faithfulness constraints, the notion of constraint ranking comes into play. In order that each winner be more harmonic than all its competing losers, the constraint violated by the loser must dominate the constraint(s) violated by the winner. As far as language acquisition is concerned, markedness constraints are often said to dominate faithfulness constraints in early stages of language acquisition (for details see Gnanadesikan, 1996). Later, a process of unhurried readjustment whereby faithfulness constraints dominate markedness constraints gets underway. All in all, grammar re-adjustment (or language acquisition progress) involves constraint re-ranking. And this has to be a big part of what teachers should let themselves in for in the classroom.

Methodology

The Context

The present study investigates English majors’ composition writings in two writing classes at one of the Jordanian state universities (namely Yarmouk University) over a period of one year. Before being admitted to university education, students sit for the school exit exam called the General Certificate of Secondary Education Examination – Tawjih. Those who pass qualify to pursue their higher education. The American credit-hour system is implemented at most Jordanian state and private universities.
During their school education, English language classes meet four times a week, a 45-minute class each. Writing is integrated along with the other language skills, so rarely are there special meetings devoted to writing teaching. In their first year at the university, students do not take independent writing classes. What this basically means is that writing continues to be integrated with the other language skills. It is not until the second year that they actually start to be exposed to writing in an independent language skills course.

For corrective feedback to be maximally effective, it has not only to be comprehensive (correcting all errors), but it has also to be focused (correcting all mistakes of each error type) (For details see Alroe, 2011). Integrating writing with the other skills results in scant attention being given to writing in general and to corrective feedback in particular. This could be due to the diverse realities of the Jordanian context such as large class size, students' lack of motivation, teachers' heavy workloads, as well as teachers' and students' attitudes towards the learning environment. In addition, in the absence of "appropriate" teaching methods and teaching materials, writing has become no more than "a series of pseudo-communicative assignments," to use Tarnopolosky's (2000, p. 212) words, that rarely heighten learners' interest, engagement, and motivation.

The Subjects
The subjects of this study were 120 EFL second year students majoring in English and taking two writing classes (Writing-1 and Writing-2) at Yarmouk University- Jordan. Each class met for 75 minutes twice a week over 14 weeks. All participants were undergraduate students from the Department of English. They registered for English Writing-1 and English Writing-2 as part of their academic load at the beginning of the first and second semester of the academic year 2008/2009. The instructor-researcher taught two sections; 37 were registered in one section and 35 were registered in the other section. About 60% of the participants were females and 40% were males.

Treatment
In the module Writing-1, students were supposed to write a well-developed and organized paragraph with clear topic sentence and an appropriate concluding sentence. In
the module Writing-2, students were expected to write a coherent, cohesive, complete and unified essay with clear thesis statement, and suitable concluding paragraph. Both courses were taught by an instructor researcher who is one of the authors of the present paper. He implemented the process approach to writing in which he used a multiple-draft syllabus (at least three drafts per assignment for each course). He always stressed content and organization in his feedback on the first draft for at least two reasons. First, we are aware that our students prefer to work on grammatical and mechanical errors (probably due to long-established tradition of Classical and Qur'anic Arabic language instruction). Second, it is of the basic premises of the process approach to view writing as a recursive, cyclical, nonlinear, as well as a decision-making process in which writers discover new ideas as they look for meaning (see literature review section above). The goal had always been to maintain a balance between form and content attention-grabbing, or “perceptual saliency”- to use Dabaghi & Tavakoli's (2009, p. 107) words of error corrections. Many invention strategies from the process approach to writing such as free writing and brainstorming were used to generate ideas. Students were then asked to write the first drafts and bring them in the next class.

The instructor researcher provided two types of feedback on every draft: coded written feedback and oral feedback. As for coded written feedback, the instructor pointed to the exact location of an error and the type of error involved and marked it with a code but did not provide corrections, thereby leaving the door open for students to diagnose and correct for themselves (for example, WF means an error in the use of the word form). To help students understand the meaning of the feedback symbols written on their drafts and use them in correcting their errors, they were provided with a correction symbols sheet. Regarding oral feedback, the instructor assigned each student a five-to-ten-minute one-on-one instructor-student conference where each student had the opportunity to ask questions about their errors and the corrections they had received as well as the chance to receive additional explanations and examples. Each conference session started with the instructor asking the student about corrections s/he had not understood and on which further assessment was still needed. In assessing students’ writing performance, the holistic approach was used to grade the students’ final drafts of each assignment only. Students’ drafts were analyzed in terms of the corrections students
had made. Concisely, each draft was evaluated whether local errors (including sentence structure, sentence arrangement, word choice) or global errors (such as level of thought, development, and organization) had been made. It turned out that every single student had made some local corrections, but very few students altered the global build-up of their drafts (See Appendix 1 for sample drafts).

The OT analysis carried out by the second researcher was then a post-learning scrutiny of students’ error corrections on their multiple drafts. If a student had made some corrections on the first draft, these same corrections were tracked down on subsequent revisions by the second researcher. What this basically means is that that OT was not integrated as a classroom treatment during the process of writing. Rather, it was a post-learning evaluation of students' products in a process-oriented environment over the course of time.

As our investigation was a theory-driven (top-down) analysis of the corrections students had made to their drafts, no quantitative analysis of each error type was made. The present study was only geared to provide a tutorial to show how by situating process writing within a constraint-based approach, treatment priorities could be reset to promote foreign language learners’ writings to higher levels of proficiency. In particular, we aimed to show that OT provides us with the mechanics (namely demotion mechanism) on how to get students postpone corrections of local errors in favor of global ones.

**Findings and Discussion**

*Optimality Theory and Process Writing*

It has already been stated that process writing consists of several stages of development, and these stages are not organized in a linear, lockstep fashion (Rose, 1980; Hughey, et al., 1983; Zamel, 1983). Lamonica (2006) argues for at least the following stages: process, global revision, local revision and editing. In terms of constraint ranking, whereas the ranking of process and editing relative to each other and to the other constraints is to some extent fixed, most of the problem lies in the relative ranking of global (content) and local (form) revision, and this is exactly what we are letting ourselves in for in this paper. In terms of formulation and content, these two constraints can roughly be sketched out as in (3) below, where each constraint is a specification on what students *should not do*:
Despite all the efforts made by the instructor to promote global error correction (especially on the first drafts), students are most of the time caught up revising for local errors at the expense of global errors. This has made us believe that foreign language learners' relatively limited success in attaining higher levels of proficiency, especially at beginning and lower-intermediate levels, lies in having *GLOBAL outrank *LOCAL as shown in (4) below:

(4)  *GLOBAL >> *LOCAL

The task of the instructor should not just be focused on what material is presented (i.e. content knowledge) but, more importantly, on how that material is presented to second language learners (i.e. pedagogy). OT, we hypothesize, provides the machinery: Feedback on local and global errors should be presented as violable constraints on surface representations concurrently, as in (5) below:

(5)  *LOCAL: (NO LOCAL) students should not attempt corrections of local errors

*GLOBAL: (NO GLOBAL) students should not attempt corrections of global errors

In less technical jargon, what this basically means is that students are instructed not on what to do, but on what not to do. To illustrate, instead of instructing them to correct some local error (i.e. passivization, relativization, if-clauses, etc.), they are instructed not to correct them at least momentarily in favor of more compelling demands. In other words, OT requires setting priorities of treatment (some constraints are higher ranking while some are lower ranking). In order to achieve maximum effectiveness in the learners, constraints are ranked relative to each other in that higher-ranking constraints are satisfied at the expense of lower ranking constraints. From a pedagogical point of view, what a constraint-based analysis does is promote the proposal that at some stage of learning not correcting a local error could be more helpful than not correcting a global error. In order to do just that, the domination relationship between global and local errors of (4) above should be reversed by giving enough classroom instruction to have *LOCAL outrank *GLOBAL as in (6) below:

(6)  *LOCAL >> *GLOBAL
An OT analysis then focuses not only on the provision of (of lack thereof) instructor’s feedback, but also on how this feedback is presented at each stage of learning. Given the two main types of constraints (*LOCAL and *GLOBAL), let us demonstrate the pedagogy that an OT analysis promotes in the classroom. At the level of detail we are considering here, *LOCAL and *GLOBAL should be decomposed into families of sub-constraints that require relative ranking. Lamonica (2006) decomposes local revision into sentence structure, word choice, and sentence arrangement; in the meantime, global revision is decomposed into level of thought, development, and organization:

(8) Sentence Structure (SS), Word Choice (WC), Sentence Arrangement (SA)  
Level of Thought (LT), Development (DE), and Organization (OR)

Going into further detail, each constraint can be further decomposed into subconstraints. For example, Sentence Structure (SS) is a cover term that includes the whole set of the English grammar such as Passive, If-Clause, Relative-Clause, etc. structures; and Word Choice encompasses the entire of the English Lexicon such as Share With, Reason For, Cause Of:

(9) SS (PASSive, IF-CLause, RElative-CLause, etc.)  
WC (Share With, Reason For, Cause Of, etc)

As for formulation, these are all worked out as constraints detailing what students should not do as in (10) below:

     *GLOBAL*LT, *DE, *OR

On close scrutiny of our students’ revised drafts, it becomes evident that learners have all attempted corrections regarding sentence structure, sentence arrangement and word choice, yet very few modifications are made regarding level of thought, development of ideas and overall organization. In order to remedy the situation, the task of the instructor is twofold. First, he should provide feedback in the form of constraints as in (11) below:

     *GLOBAL *LT, *DE, *OR
Second, his teaching should make clear that *LOCAL constraints should dominate *GLOBAL constraints, so that students would pay more attention to global errors (and thus resetting their priorities). In some technical jargon, the writing teacher should provide enough data which will have the consequent effect of the promotion of *LOCAL constraints, and the demotion of *GLOBAL constraints.


OT then provides us with the machinery to get even to particular examples, whereby our sole task is to find out the ranking of each constraint relative to the whole set of constraints. And this is what we mean by pedagogy (i.e. how the material is presented to second language learners both in the textbook and by the teacher).

It has already been noted that one big advantage of this line of reasoning is that it gives us the means to pattern errors, and thus prioritize treatment. That is, due to the claim that constraints (and ultimately error patterns) are ranked relative to each other, we get to the happy conclusion that certain errors are then predicative of others (Dinnsen et al., 1997). The demotion of some error patterns may pull along some other error patterns for free. In other words, fixing one error pattern may obviate the need for fixing some other error pattern(s). For instance, once learners are instructed that *PASS dominates *LT, then *LT is demoted not only one step below *PASS but one step below all constraints encompassed by *SS (namely. *PASS, *IF-CL, *RE-CL, etc.):

(13) Constraint ranking


Similarly, when the local *WC dominates the global *DE, then *DE is demoted at least one step below the whole set of constraints encompassed by *WC. When the learner is instructed that the local constraint *SA should outrank the global constraint *OR, we get to the desired constraint ranking as in (14) below, whereby the whole set of constraint encompassed by *SS, and *WC, and *SA outrank *LT, *DE and *OR:

(14) Constraint ranking


What the constraint ranking in (14) above shows is that writing teachers have available to them the mechanics to shift priorities. This corroborates the findings of some research
(Hillocks, 1986; Leki, 1990, 1991; Ferris, 1995) that postponing local errors in favor of
global errors on preliminary drafts could be more constructive than providing
spontaneous corrections of all types of errors once and for all. This constraint-based
approach reinforces our belief in the explanatory power of some notational devices. In the
next section, in addition to bringing evidence from the tutorial to support our previous
claim that variation in students' performance arises because of a unique constraint
ranking, we will attempt to link this tutorial with some findings of previous investigations
on corrective feedback.

Discussion
The thesis advocated here is that when composing second language learners fail to
exploit the relative ranking of the set of violable constraints in a native-like manner. This
is probably due, we hypothesize, to the common practice of splitting error correction
types (e.g. local versus global errors) on students' written work. As a remedy strategy,
teachers need to rework the relationship between feedback on form and content feedback
by providing enough corrective feedback that would have the inevitable effect of
demoting some structural requirements and promoting others. Hence, study findings
show that accuracy gains are best when both types of errors are targeted simultaneously
(see literature review section above).

In order to analyze second language learners’ error patterns, three things have
been done. First, we demonstrated the learners’ actual rankings of the set of violable
constraints at a particular stage of learning. Second, we identified some error patterns that
are due to learners’ failure to exploit the target language native-like constraint ranking.
Finally, we sketched a remedy strategy by which we set priorities of treatment; i.e. which
error patterns should be targeted first. This last step is of great significance as it relates to
Smolensky, 1993, 1996; inter alia). The underlying assumption is like this: given the
thesis that certain errors are predicative of others (Dennsen, Barlow, and Morrisette,
1997) and that constraints (and ultimately error patterns) are ranked relative to each other,
the demotion of some error patterns may pull along some other error patterns for free. In
other words, fixing one error pattern may obviate the need for fixing some other error pattern(s).

**How does this relate to previous research findings?** Whereas some studies (Cf. Fathman & Whalley, 1990; Ashwell, 2000; Ferris, 1999; Ferris & Roberts, 2001; Chandler, 2003; Russell & Spada, 2006) advocate the view that corrective feedback is effective, others (Cf. Sheppard, 1992; Polio, Fleck, and Leder, 1998; Truscott, 1996, 1999) believe that teachers could make do without error correction in writing classes. The conflicting findings of the limited research on the efficacy of corrective feedback stimulate further investigations and theory-building. Despite the impressionistic views of writing teachers, research in both L1 and L2 writing has shown little evidence of the impact of teachers’ feedback on students’ writing (Leki, 1990). For example, Knoblauch & Brannonm (1981) and Hillcocks (1986) have investigated different types of instructors’ feedback and have found out that none of these different comments have affected the desired impact on students’ subsequent drafts. However, some scholars (Krashen, 1984; Hillcocks, 1986; Freedman, 1987) have questioned the essence of the theoretical claims on which these studies were conducted. Alternatively, they have stressed that feedback provided on intermediate drafts (first and second) is more constructive than comments on final drafts. Berger (1991) reviews both L1 and L2 studies on the impact of teacher and self-feedback on helping students revise and improve their writing, reporting that teachers’ feedback on preliminary drafts helps students revise effectively (Chaudron, 1983; Zhang, 1985; Zhang & Halpern, 1988).

Our suggested answer to the state of contradiction in previous research findings is that none of the studies, we believe, have as yet considered the sequencing of the different types of instructors’ feedback, i.e. which kind of feedback is given at each stage of development. What most previous studies have stressed is the importance of providing students with some feedback. However, they have not got to the level of detail to explain which type of feedback is most constructive at each stage of development, a question that brings the relationship between feedback of form and content feedback under scrutiny. This question is just beginning to be addressed (Cf. Ellis, Loewen, and Erlam, 2006; Dabaghi & Tavakoli, 2009). By quoting Chandler (2003) in the introduction section, we attempted to show that most teachers in EFL context still see no harm in following the
conventional practice of providing feedback on content first and feedback on form in later drafts. However, by situating writing within the framework of OT, we hope to have shown that all competing demands (whether global or local) should be produced on a parallel fashion, so that the conflict could be resolved only optimally. This claim is supported by the findings reported in a number of studies on corrective feedback. Some studies (e.g. Robb, Ross, and Shorteed, 1986; Lizotte, 2001; Chandler, 2003) have shown that providing corrective feedback on form contributed positively not only to 'grammatical gains', but also to 'fluency gains' even if no feedback is provided to that effect. In other words, fixing form has pulled along content for free. Second, a number of studies (e.g. Fathman & Whally, 1990; Ashwell, 2000) have shown that gains in formal accuracy are optimal when feedback on form and content feedback are provided simultaneously. For example, Alroe (2011) has stressed that correction can be more corrective when it is "integrated with content comment" (p. 34). This constitutes argument against the conventional practice regarding splitting up corrective feedback (Sommers, 1982; Zamel, 1985). Not only this, but building a tutorial like the one we suggest here could also help us describe, at least tentatively, some aspects of the learners' transitional grammar.

For example, Dabaghi & Tavakoli (2009) found that corrective feedback was more effective in learning irregular past verb forms and the definite article than in learning regular past verb forms and the indefinite article in some EFL context. Whereas their findings confirm the findings of some previous research (e.g. Dietrich, Klein, and Noyau, 1995), they run counter to the findings of others (e.g. Dulay & Burt, 1975; Larsen-Freeman, 1975) which confirmed that the regular past tense form is learned before the irregular past form. Similar findings were obtained for the sequencing of the definite and indefinite article (for an illuminating discussion see Dabaghi & Tavakoli, 2009). In analyzing their findings, they have attempted a number of explanations for the varying success of EFL learners (Iranians in their case) in learning these grammatical features in English. To them, 'negotiation', 'saliency' and 'individualized attention' were the main driving forces. To get to the level of details they were considering the data, they argue that in addition to the relatively higher frequency of irregular verbs in English, "syntactic complexity, perceptual salience, and the novelty of the form in the target
language" (Dabaghi & Tavakoli, 2009, p. 106) are all possible explanations why the irregular form is learned before the regular form. For article learning, they suggest native "language influence, frequency, stress, and homonymy" (Dabaghi & Tavakoli, 2009, p. 108).

However, a major concern for us here is that an analysis like this one suffers from at least two main analytical flaws. First, it is never crystal clear why different explanations are sought for the acquisition of different grammatical features. A theory of language learning will definitely have more explanatory power when it can pull things together. And this is exactly what a linguistic theory such as OT, we believe, can do. Second, it is also not clear why 'perceptual saliency', for example, is used as a driving and a confounding force for the learning of (and lack thereof) some grammatical feature. They themselves are puzzled by this. Their argument runs like this:

… both regular and irregular forms gain perceptual salience, doing so in different ways. For example, frequency of input for the regular form is higher than it is for the irregular form; this means that the regular forms are also perceptually salient. Therefore, this factor can not be treated as significant. However, a possible consideration to make is, as Gass and Selinker (2001) mention, highly infrequent items can also lead to salience because when an item is infrequent it might get more attention by learners. The number of irregular forms is considerably less than regular forms, thus making them more salient.

(Dabaghi & Tavakoli, 2009, p. 107)

Despite all this, what is especially worth noting is that they make a number of claims which can lead us into OT. First, they pointed out that abstracting out rules is a confounding factor in learning some grammatical features. And this is exactly what an OT analysis does: it dispenses altogether with the notion of rules; hence structural requirements are always violable, and oftentimes impose competing demands in the mind of learners. Learners make better learning when grammatical features are learned as 'items learning', which, to them, turned out to be more effective than abstracting out rules. What this basically means is that structural requirements (past tense forms, articles, etc.) are decomposed into competing violable constraints. In addition, they supported the
claim that language learning is a continuous process of restructuring the learners' interlanguage. In OT terms, what this basically means is that language learning involves relative re-ranking of competing structural requirements simply because whereas some constraints are higher ranking, others are lower ranking. This probably contributes to what Dabaghi & Tavakoli (2009) call the 'perceptual saliency' of some structural features (p. 107). To them, some features are perceptually more salient than others. In OT terms, an irregular past tense form takes precedence over a regular past tense form simply because the set of constraints which militate against the regular past tense form are higher ranking whereas the set of constraints which militate against the irregular past tense form are lower ranking. Violating a lower ranking constraint is not as fatal as violating a higher ranking constraint, and does not therefore result in suboptimal form.

The same rationale applies in the case of providing corrective feedback on local and global errors. To the language learners, not all types of errors are equally salient. Our current theory-driven analysis may, we believe, help shed light on the controversy from a different angle. We firmly believe that a major internal cause of the disparity is the 'unnecessary' separation between both types of corrective feedback – a state of affairs that brings to light the relevance of the first question of the study, namely the relationship between feedback on form and feedback on content. To illustrate, by using the two-way dichotomy of local versus global errors (i.e. feedback on form versus feedback on content), second language learners may be tempted to focus on some error types at the expense of others. Our current investigation has made us confirm the findings of a number of studies (e.g. Cohen, 1987; Radecki & Swales, 1988; Leki, 1991) which have shown that students focus on local errors (a learning that may not last for long) at the expense of more global ones (which are more likely to stand the test of time). In OT terms, *LOCAL is said to be dominated by *GLOBAL, a state of affairs that our suggested tutorial is designed to reverse, i.e. a demotion mechanism in which *GLOBAL is demoted down the hierarchy until it is outranked by *LOCAL. Our current claim is supported by the findings of previous research in that experienced writers tend to revise globally, but less experienced writers tend to revise only locally (see Faigly & Witte, 1981; Sommers, 1980).
To illustrate, despite the fact that students appreciate and expect their instructors’ feedback on their writing (Cohen & Cavalcanti, 1990; McCurdy, 1992; Hedgcock & Lefkowitz, 1994; Hyland & Hyland, 2001), they are most of the time caught up revising for local errors at the expense of global errors (Cf. Radecki & Swales, 1988; Leki, 1991). Students oftentimes report that their instructors focus mainly on grammar and/or that they themselves prefer to receive feedback on grammar, rather than on content (Cohen, 1987; Radecki & Swales, 1988; Leki, 1991). Students, in Ferris’s (1995) study, report receiving and paying their utmost attention to feedback on grammar, content, and organization, in that order, and that their teachers’ comments help them improve their grammar. This has been carried over onto language pedagogy, especially in EFL contexts. That is, although it is reported that instructors’ give comments on ideas and organization (McCurdy, 1992; Hedgcock & Lefkowitz, 1994; Ferris, 1995), they still consider mechanical errors as the most important criterion for responding to students’ writing (Cf. Applebee, 1981; Cohen, 1987; Radecki & Swales, 1988; Leki, 1991; Lee, 2008).

This may partly explain why corrective feedback is only momentarily constructive. For a number of longitudinal studies (e.g. Semke, 1984; Goring-kepner, 1991; Sheppard, 1992; Polio et al., 1998; Fazio, 2001; Chandler, 2003; Truscott & Hsu, 2008; inter alia) have shown that corrective feedback did not stand the test of time. For example, although the experimental group outperformed the control group on the revision task in Truscott and Hsu (2008), both groups were the same on subsequent writing tasks. They came up with conclusion that ”no relation was found between success on the revision task and learning as measured by performance on a new writing task” (see Truscott and Hsu, 2008, p. 298). Similar findings were obtained by Bruton (2009) who believes that ”there is almost no evidence of the effects of correction on subsequent writing” (p. 6). What this basically means is that, as Guenette (2007) rightly points out, “the results of the many experimental studies on written corrective feedback carried out over the last 20 years have been so contradictory.” (p. 40). Extraneous variables such as internal inconsistency of research design, data collection, classroom instruction and maturation of the population could have biased the findings, and therefore interpretation, of previous research findings.
It is also worth noting that An OT analysis pays heed to different EFL contexts by making use of the notion of constraint ranking. To illustrate, one of the basic premises of an OT analysis is that although constraints are universal (available in the interlanguage of each language learner), their ranking is language-dependent (differs from one language learning group to another). What this basically means is that the relative ordering of the regular and irregular past tense forms in Dabaghi and Tavakoli (2009), for example, is not fixed for all language learners. Their findings may then be generalized to that language context, and more specifically to the language learners with the same native language background. This probably explains why conflicting findings were reported for different language learning contexts. For example, upon running a meta analysis for most studies that reported significant effect for error correction on students' writings, Truscott (2009) has shown that those actual gains (if any) are at best very small. For the findings could have "contaminated by an additional, unknown variable" (see Truscott 2009, p. 268). One of the variables that has rarely been investigated is the already existing "interlanguage" of the language learners who have enjoyed hours of instructions on how to write in the target language. This probably adds to why an OT analysis could be more successful in addressing all EFL audiences. For, the already-existing constraint ranking in the minds of one language learning group is not 100% identical with that of another learning group. Corrective feedback can then be maximally effective when it provides the appropriate feedback at the right time for the right group of learners.

Conclusion

Research findings have shown conflicting views regarding the effectiveness (or lack thereof) of corrective feedback (Truscott, 1996, 1999, 2007; Ferris, 1999a, 2004; Chandler, 2003; Bitchener et al., 2005; Bitchener, 2008; Guenette, 2007; Truscott & Hsu 2008; Bruton, 2009), and even attributed this controversy to external factors such as research design and methodology (See the illuminating discussion in Guenette, 2007). In this paper, we have shed light upon one internal factor (namely feedback sequencing) from a constraint-based perspective. In particular, we provide a tutorial to show how the basic mechanics of OT have proved to have an explanatory value in that they, in addition to revealing the simplicity behind systems that initially seem complex, increase our
understanding of how language is best taught/learned in the classroom. The basic assumption of a constraint-based theory, the argument goes, is that learners’ error patterns are due to relative ranking of a set of universal and violable constraints (global and local constraints in our case). Therefore, a different grammar arises because of a unique constraint ranking. What this basically means is that as a viable theory, a constraint-based analysis captures common learners’ errors, allows for individual variation, and, most importantly, sets treatment priorities.

In more concrete terms, one reason why learners collectively commit some particular common errors is that they set out with some initial constraint ranking by which markedness constraints (those which favor forms that are structurally simple) dominate faithfulness constraints (those which favor replication of the input). Despite the fact that most learners are introduced to genuine native language texts which demonstrate the rhetoric of the target language, they still at the production level create texts which markedly deviate from the target language norms. Besides, as each grammar results from a unique ordering of the set of violable constraints, each learner builds up her own grammar uniquely. Collective and individual deviations from the norm can be put on track by re-ranking the relevant set of violable constraints. The basic machinery is constraint demotion; for students, in Guenette’s words, “have to be provided with appropriate feedback, given at the right time and in the proper context” (Guenette, 2007, p. 51). That is, for promoting writing skills in second/foreign language learners, we have shown that the demotion of global constraints below local ones results in substantial improvement of the beginning learner’s grammar in that her grammar would become more faithful to advanced writers’ grammar.

Probably unlike all previous studies on corrective feedback, the present theory-driven analysis calls for making do without the classical distinguishing line between feedback on form and content feedback. At the theoretical front, an OT analysis of error correction assumes that all competing demands (whether content or form feedback) should be provided in a parallel fashion, so that the conflict could be resolved only optimally. This is probably so because when combined, competing requirements, we believe, help each other to adjust.
References


Appendix (1)

First draft

My childhood period was unforgettable one, but it was a combination of happy and sad moments. I remember that happy day when I met my first friend, she moved with her family to our quarter, they were exactly the next door to us, so we visited them and our friendship started. What was also wonderful and exciting in one of my vacations is learning swimming in a nice swimming pool in our quarter, so I was able to float on the water and to dive into it as if I were small shining and beautiful fish. Another nice memory was my last day in grade one at school, my teacher awarded me in front of the whole class because I was the first out of my colleagues, I was really proud of myself. The last good memory I remember was parties especially the birthday's one. I liked so much the gifts I had from my friends and the games we played together at my home celebrating that nice occasion, they really cheered me up! On the other hand there were as all children, many problems and sad experiences I faced. One of the normal problems I hated was to sleep and to wake up early because of my school. So, in somehow, I didn't like to go to school all the days. However, my saddest memory I had was the death of my grandfather whom I loved so much especially I didn't know what the death is in that time, I thought it is a ghostly monster comes to kidnap my friends and my family so I spent horrible nights trying to overcome my pain, it was really a terrible time I had lived. Another sad memory was the war in Kuwait that forced many families to abandon their homes, jobs and life and to immigrate to another place so many of them
became homeless and very poor. I can obviously remember my worst memory, I was afraid of darkness so every night when my parents wanted to go to bed, I cried making a loud noise until they accept me to come to sleep with them. In short, everyone has his own experiments and memories so, I think, it is really worthy to go back, sometimes, to your memories, may be you will discover what kind of person you are.

Second draft
My childhood was an unforgettable one, but it was a combination of happy and sad moments. I remember that happy day when I met my first friend. She moved with her family to our quarter. They were next to us. So we visited them and our friendship started. Joining the kindergarten was a wonderful period. I learned there how to count from one to ten, utter alphabets and a lot of songs too. Another nice memory was my last day in the first grade at school. I was awarded the first prize in front of the whole class, so I was really proud of myself. The last good memory was parties, especially, that of my first birthday. I liked the gifts I had from my friends so much, and the games we played together celebrating that nice occasion. On the other hand, there were many problems and sad experiences I faced. One of these problems was sleeping and waking up early, so I didn't like to go to school all the days. However, my saddest memory was the death of my grandfather whom I loved so much. The problem was that I did not understand the meaning of death. I thought it was a monster was coming to kidnap my friends and my family, so I spent horrible nights trying to overcome my pain. It was really a terrible time I lived. Another sad memory was the war in Kuwait that forced many families to abandon their homes, jobs and immigrate to another place, so many of them became homeless and very poor. I can obviously remember my worst memory. I was afraid of darkness, so every night when my parents wanted to go to bed, I cried making a loud noise until they allowed me to come to sleep with them. In short, I had some sad memories, but I lived many nice moments. I will never forget my wonderful childhood.
Towards a Discerning Image of Learners’ Estimation of Task Difficulty and Motivation: Goal Orientations Highlighted

Mohamed Ridha Ben Maad

University of Carthage, Tunisia

Bio Data
Mohamed Ridha Ben Maad is an Assistant Professor at the Higher Institute for Childhood Education, University of Carthage-Tunis, Tunisia. His main research interests are in applied linguistics and second language acquisition. He also participates in research projects related to early childhood education.

Abstract
This article is based on an experimental study which focused on goal orientation as a researchable learner difference factor in order to explore the nature of interaction between second language learners and language tasks. The study aimed to examine the extent of goal orientation in language learners' profile and how this learner difference variable affects their perception of task difficulty and motivation. Analysis of the findings revealed two unrelated goal orientations which reflected two significantly different response types to task difficulty. As one goal orientation group responded positively to unfamiliar and unplanned tasks, the other goal group did not. In view of these findings, the article accentuates the need to revisit the established task research format that confines task variation to design and sequencing factors and eclipses individual differences.

Keywords: Goal orientation, familiarity, individual differences, motivation, planning, task

Introduction
This article presents an empirical attempt to investigate the reaction patterns of language learners in response to tasks. It seeks to document the role of individual differences, through focus on goal orientations, in shaping language learners’ perceptions of task difficulty and task motivation. The article draws on the long-standing research line attendant to the study of language learning tasks which has long captivated the interest of scholarship in the research field of second language learning (e.g., Ellis, 1987, 2009;
Mehnert, 1998; Skehan, 1998). Tasks in this regard serve as tools to obtain “clinically elicited samples” (Ellis & Barkhuizen, 2005, p.7) for the purposes of categorizing, sampling, and assessing the language learning process.

According to Ellis (2000), the merit of the task-based approach does not only stem from the quality of tasks being clinical elicitation techniques for researchers and useful pedagogical instruments for practitioners, but also from its openness to various theoretical perspectives such as the psycholinguistic perspective. Within this purview, task difficulty predicts the degree of variation in learners’ performance through the experimental manipulation of task variables, such as planning (e.g., Ellis, 2009; Mehnert, 1998; Ortega, 1999, 2005) and familiarity (e.g., Bygate, 2001).

The article also makes reference to research on motivation, particularly to a micro perspective of second language/foreign language (L2/FL) motivation which is anchored in language classroom research (Dörnyei, 1994). Dörnyei (1994) championed a more “education-centered approach to motivation” (p.273) that focuses on what teachers can do to motivate learners. Along with this perspective, Pintrich and Schunk (2002) stated that goal orientation theory is “probably the most active area of research on student motivation in classrooms and it has direct implications for students and teachers” (p. 242). They proposed goal orientation as the construct which does not only bear affective properties, but it also captures the cognitive aspect in human behaviour which may well be relevant to study language tasks from an individual difference (ID) perspective

Research background
The study of individual differences is one of the most active research areas in the field of psychology as featured through the extensive investigation of concepts like personality, intelligence, and motivation. Far from investigating mere idiosyncrasies, this research strand focuses on the kind of individual variation confined to the stability standard. Accordingly, Dörnyei (2005) stated that “ID constructs refer to dimensions of enduring personal characteristics that are assumed to apply to everybody and on which people differ by degree” (p. 4). This intent to capture ID patterns underlies a long-lasting controversy between the collective and the individual: Where the individual seems counterproductive to the accomplishing generalizable results, the collective, or the “grand sweep view”
(Larsen-Freeman, 2006, p.594), overrides the value of individuality in human sciences. This dilemma has drawn a demarcation line within all the research disciplines researching human behavior, including the area of second language acquisition (SLA) which has documented a number of IDs in various educational contexts.

Research findings related to IDs have invariably revolved around the concepts of aptitude and motivation (Dörnyei, 2005). The supremacy of these two concepts has perhaps upset the sense of ecology within this research line, leaving unaddressed, at least disproportionately, a number of learner variables whose exploration might have advanced and refined our understanding of L2/FL learning behavior. In addition to this heavily skewed ID-research picture, the dominating ID concepts of aptitude and variation have yielded little substance to accommodate to the field of SLA. In this regard, Dörnyei (2005) explained this sense of isolation by the fact that “the original product-oriented conception of the two key ID factors, aptitude and motivation, was incompatible with the inherently process-oriented stance of SLA” (p. 6). This product-oriented conceptualization, which is more of a pedagogical necessity than a theoretically-driven choice, is understood in view of the dilemma about the collective and the individual.

Task-based research has been confined to viewing tasks as “neutral devices for testing” (Iwashita et al., 2001, p. 406). Although it has yielded substantial experimental literature, Cumming (2006) observed that this research format seems to have eclipsed the role of IDs. One explanation for the neglect of such learner variation is explained by a reductionist stance within this research line as it operates under the confines of an agenda that seeks to develop universal properties that guide learners towards predictable forms of performance outcomes (see Ellis, 2000). It seems that this reductionist way of defining task difficulty echoes a pedagogical agenda within which task-based research “provides a valid means of packaging language experience and leads to effective learning” (italics mine) (Lynch & MacLean, 2000, p. 224). Task difficulty is indeed a matter a learner perception more than the prerogative of professional raters; and what is demanding for one individual learner is not necessarily so for another. In this regard, some researchers, such as Elder, Iwashita, and McNamara (2002), questioned the real value of such generalizations and demanded that tasks should “be treated with extreme caution and that the findings of SLA research should be revisited with this caveat in mind” (p. 364).
Bachman (2002) also cautioned against the mainstream tendency to build on speculative postulates where difficulty is identified in contexts where the learner factor is completely marginalized. For example, Bachman (2002) criticized Skehan’s (1998) treatment of task demands as detached variables that can be isolated for empirical testing. Bachman claimed that communicative stress and task complexity are fundamentally individual characteristics, and so task demands “are not inherent in tasks themselves, but are functions of the interactions between a given test-taker and a given test task [and so the] empirical estimates of task difficulty are not estimates of separate entity, ‘difficulty’, but are themselves artifacts of the interaction between the test-taker’s ability and the characteristics of the task” (Bachman, 2002, p. 464).

Even with such calls for a central role of IDs in the characterization of task difficulty, task-based literature has documented modest interest among L2/FL researchers in the last decade. For instance, Dörnyei and Kormos (2000) championed a central role for IDs in an “extended task-based paradigm.” This claim was later echoed in the landmark work of Ortega (2005). Ortega (2005) revisited earlier findings (i.e., Ortega, 1995; 1999) about the effect of task planning on L2/FL performance. In her analysis of supplementary interview data, Ortega (2005) was able to discern two types of L2/FL learners: communication-oriented learners and accuracy-oriented learners. Differences in terms of learning behavior (e.g., strategy use) among her informants were inconsistent with the results reported earlier because, as she admitted, her prior use of group averages analysis obscured ID aspects. Ortega therefore called for the need to reconsider the extant task-based findings for a wholesale analysis from an ID perspective, or otherwise statistical interpretations would remain misleading.

Findings in Ortega (2005) aligned with the cautious stance of Ellis (2000) about the consequences of orthodoxy in task-based research as any operationalization of task difficulty will remain impressionistic and reductionist unless the role of the learner is considered. Similarly, Larsen-Freeman (2006) concurred with this line of reasoning since ‘individuals not only determine what aspects of the outside world are relevant to them, but they actively construct a world around themselves and are constantly altering it’ (p. 594). Hence, a realistic description of L2/FL learner behavior should highlight this type of self-regulated learning view which upholds an active role for IDs in the assessment of
task difficulty instead of the exclusive investment on external raters (i.e., professional testing experts or simply teacher practitioners).

The concept of a self-regulated learner has been well attested in other classroom-focused research and presented a consistent and operational accounts of IDs. Achievement goal theory lends itself to this line of research because, according to Midgley (2002), learners’ “goals provide a framework within which individuals interpret and react to events, and result in different patterns of cognition, affect, and behavior” (p. xi). Such patterns may enable individuals, including L2/FL learners, to develop differential perceptions of task difficulty as a function of differences in their goal orientations.

The concept of goal orientation is central to Achievement Goal theory which is concerned with the study of learners’ views and assessment of task achievement. Across the goal-related literature, there is agreement over two types of goal orientation: Mastery goal orientation (MGO) and Performance goal orientation (PGO) (Ames, 1992). Mastery goal orientation refers to individuals who value the learning process and competency growth rather than the learning product. They show more enthusiasm and effort for particularly challenging tasks and a willingness to take risks since a mistake represents a learning opportunity and not sign of failure. Performance goal orientation refers to individuals who develop a product-oriented sense of learning achievement. Driven by constant apprehension of failure, they adopt a maladaptive behavior that is inclined to avoiding challenge. It should be noted that despite being a ubiquitous variable in educational research, the construct of goal orientation has received only little consideration among SLA scholarship (e.g., Botsas & Padeliadu, 2003; He, 2005), let alone in the study of task difficulty.

The classroom-anchored nature of goal orientation enables better researchability for the concept of motivation which can be hardly dissociated from task difficulty. Rather than treating motivation as an isolated variable as in the Gradnerian tradition, goal orientation research provides an incisive account of how one’s perception of difficulty relates to the level of and type of difficulty inherent in a given task. Cumming (2006), concurring with this line of theorizing, states that “research on motivation has mostly involved survey studies that analyze the attitudes of groups of students, not the goals of
specific learners in particular circumstances of language learning” (p. 3). Each goal orientation, however, represents a number of achievement values whereby learners define and judge their success and failure (Ames, 1992). These achievement values transform into cognitive and affective decisions that shape one’s perception of task difficulty. In other words, what seems to be difficult for one individual affiliated with a given goal orientation may not necessarily apply to another individual with a different goal orientation.

The present study was to some extent inspired by the postulate of Ortega (1999) that research “needs to recognize and account for individual differences, which may otherwise obscure the findings” (p. 136). Despite her awareness of learner orientations (i.e., communication oriented versus accuracy-oriented), she did not dichotomize a priori her informants on the basis of such orientations, thus admitting her indecision that she did “not want to reopen the old research agenda regarding (usually unoperationalized) dichotomies of learner type” (Ortega, 1999, p. 136). Building on this methodological concern, the present study operationalized these orientations through their treatment as an ID factor that can predict L2/FL learners’ perceptions of task difficulty. Its research design therefore sought to test the following hypotheses:

1) Goal orientation exists as a fundamental ID variable in L2/FL classrooms.
2) Goal orientation affects L2/FL learners’ perceptions of task difficulty and motivation.

As demonstrated by Hypothesis 1, the nature of this study is exploratory. Indeed, I wanted to examine the extent of the goal affiliation in the L2 classroom experience through an experimental design. Results related to this hypothesis would consolidate the ID variable of goal orientation in a systematic way. They would therefore ensure the possibility to address Hypothesis 2 which seeks to draw an empirical image about the nature of interaction between task and ID variables so as to puzzle out learners’ perceptions of task difficulty and motivation.

The study
A total of 211 full-time students participated in the present study during the university year 2006-2007. The participants were enrolled in their first year of a three-year program
offered by the Department of English at the University of Manouba, Tunisia (a yearly intake of approximately 600 students). Female students outnumbered the male students (females: N = 172; males: N = 39). Their age ranged from 19 to 23 years. The mean length of time they studied English was 6.7 years. As native speakers of Tunisian Arabic, they learned their English exclusively in a classroom environment, thus having little opportunity to use English for communicative purposes outside the classroom setting. This sample represented a fairly homogeneous group in terms of their schooling history and their intermediate English proficiency level. Such selection was intended to limit individual differences in favor of the goal orientation variable. There was common agreement among all the teachers that their students’ proficiency level was intermediate. They reported that the course materials they were using were taken from resources intended for intermediate L2/FL learners.

A Goal-orientation Questionnaire was developed to gain an incisive account of students’ goal orientation (Appendix A). Its main objective was to determine to which goal area the student participants would belong. It comprised two 5-point Likert scales that were expected to document data related to the two levels of goal orientation. Each scale consisted of 10 items. The first scale focused on Mastery goal orientation (e.g., sample item 8: “I like speaking tasks best when they make me learn new things”). The second scale aimed to assess Performance goal orientation (e.g., sample item 14: “I prefer my task performance to be graded only when I do well”). All the 20 items had to be assessed on a Likert-type scale, ranging from strongly agrees (5) to strongly disagree (1).

The Goal-orientation Questionnaire was based on well-cited instruments (e.g., Midgley et al., 1998; Skaalvik, 1997). The distribution of the questionnaire items at one scale observed a sequence that was symmetrical to the allocation of items at the other scale. In other words, Items 1 (i.e., indicating risk-taking behavior) and 11 (i.e., indicating risk-avoiding behavior) shared the concept of risk management as it was the case for Items 2 (i.e., indicating intrinsic evaluation of achievement) and 12 (i.e., indicating extrinsic evaluation of achievement) which shared the concept of achievement evaluation, etc. The expected orthogonal relationship between the two item sequences, one at a time, would make the respondents more/less set on one scale and not the other. Also, it is relevant to note that the questionnaire followed a closed-response design and it
was expected to provide a greater ease of response and reliability than an open-ended response design. However, some open-ended questions were later used in interviews protocols to explain the choices of the respondents.

The raw scores of the questionnaire results were calculated to identify the best scorers at each goal orientation end. More than half of the participants (N = 109) were retained after applying a cut-off score as a screening procedure. These participants completed the same questionnaire they had taken previously to re-examine its reliability. Eventually, the 30 best scorers were chosen for the next experimental step of task performance. The participants performed three narrative tasks under different sequencing conditions (i.e., Task 1 = - familiarity/+ planning; Task 2 = + familiarity/- planning; Task 3 = + familiarity/+ planning).

The narrative tasks used in the present study were composed of stories based on sequenced picture prompts. They were based on one discourse mode where tellers described events in a watch-and-tell style. The 11 picture prompts in each task represented common narrative scripts (e.g., winning a jackpot, a success story of an athlete, and a rock band biography). In order to examine the participants’ familiarity with the experimental tasks, the first three items of the Post-task Questionnaire measured the extent of such familiarity. As to the planning factor, to ensure the unplanned condition (Task 2), the participants were given only one minute of preparation time. They were given ten minutes of preparation time in Tasks 1 and 3 to consolidate the planning condition.

A post-task questionnaire (Appendix B) was administered to these participants after performing the narrative tasks. It measured the performer’s perceptions of task familiarity (Items 1, 2, and 3), task difficulty (Items 4, 5, and 6), and task motivation (Items 7, 8, and 9). By way of illustration, the first three items were meant to document students’ prior experience with similar narrative tasks (e.g., Item 1: “Before working on this task, I had the opportunity to do similar narrative tasks.” The participants’ perceptions of difficulty and motivation were strongly related to the sequencing conditions of both the planning time and familiarity variables. On the whole, 87 questionnaire responses were collected after the three task performance episodes.
Results

Goal-orientation Questionnaire results

In order to check the distribution of the data collected from the questionnaire, a descriptive analysis across the two administration episodes was conducted. Skewness results in Table 1 indicate that the questionnaire responses were normally distributed at both ends. The symmetric distribution consists in the balance between positively skewed MGO figures and negatively skewed PGO figures. However, mean scores from the second administration episode are fairly higher than those from the first administration episode. For instance, the respondents scored higher on Item 9 at the level of first episode ($M = 2.81$) than the second episode ($M = 3.14$), as is respectively the case for Item 20 where the respondents scored higher in the first episode ($M = 3.21$) than in the first episode ($M = 3.01$).

The consistent increase of mean score at the second wave of questionnaire administration indicates that the second sample ($N= 109$) were the best scorers among the initial sample ($N= 211$). Consequently, the affiliation of the former sample with a given goal orientation was reasonably more obvious than that of the latter. To test the reliability of the two questionnaire scales, an Item-reliability analysis reported high coefficient alphas for both goal scales and across data from the two episodes of questionnaire administration. That is, as the MGO scale achieved high and consistent alpha values (i.e., $\alpha = .904$ and $\alpha = .908$) along the two data collection episodes, so did the PGO scale (i.e., $\alpha = .875$ and $\alpha = .897$).

Table 1 Descriptive statistics for and reliability of MGO/PGO subscales

<table>
<thead>
<tr>
<th>Scale</th>
<th>Administration episode 1 (N = 211)</th>
<th>Administration episode 2 (N = 109)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Frequencies estimate</td>
<td>Reliability</td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>SD</td>
</tr>
<tr>
<td>MGO scale</td>
<td>$\alpha = .908$</td>
<td>$\alpha = .904$</td>
</tr>
<tr>
<td>$\alpha = .908$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>2.99</td>
<td>1.41</td>
</tr>
<tr>
<td>2</td>
<td>3.03</td>
<td>1.44</td>
</tr>
</tbody>
</table>
To verify the content validity of the questionnaire, a Principal Components analysis was performed on the 20 items across the two administration episodes. The matrix in Table 2 provides information which supports a two-component solution. The first ten items sort on Component 1 and the last ten items sorted on Component 2.

According to results from Table 2, Component 1 has high and positive loadings on the MGO scale in Administration episode 1 (i.e., range between .60 (Item 1) and .77 (Item 1)) and in Administration episode 2 (i.e., range between .64 (Item 8) and .78 (Item 1)). In contrast, Component 2 had low and negative loadings on the same scale. The range of loadings was between -.06 (Item 10) and -.28 (Item 1) in Administration episode 1 and between -.09 (Item 4) and -.28 (Item 10) in Administration episode 2. The PGO scale shows the reverse loading patterns. Component 1 presents low and negative loadings as low as -.09 (Item 12) and -.10 (Item 15) whereas Component 2 has high and
positive loadings (e.g., Item 14 = .75) in Administration episode 1 and a fairly similar
distribution (e.g., Item 19 = .75) in Administration episode 2.

Table 2 Matrix of Principal Components loadings for 20 MGO/PGO items

|        | 1    | 2    | 3    | 4    | 5    | 6    | 7    | 8    | 9    | 10   | 11   | 12   | 13   | 14   | 15   | 16   | 17   | 18   | 19   | 20   |
|--------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|
| MGO scale |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |
| Admin. 1 | .60  | .63  | .73  | .75  | .73  | .77  | .63  | .70  | .70  | .76  | .13  | .09  | .15  | .15  | .10  | .21  | .14  | .23  | .29  | .13  | .19  | .16  | .19  | .12  | .20  | .22  | .18  | .20  | .06  | .67  | .57  | .65  | .75  | .64  | .71  | .70  | .66  | .62  | .59  | .46  | .21  | .28  | .72  | .68  | .69  | .70  | .60  | .62  | .72  | .64  | .75  | .73  | .53  | .65  |
| Admin. 2 | .28  | .19  | .16  | .19  | .12  | .20  | .22  | .18  | .20  | .06  | .67  | .57  | .65  | .75  | .64  | .71  | .70  | .66  | .62  | .59  | .46  | .21  | .28  | .72  | .68  | .69  | .70  | .60  | .62  | .72  | .64  | .75  | .73  | .53  | .65  | .65  | .69  | .70  | .60  | .62  | .72  | .64  | .75  | .73  | .53  | .65  |

Note: Admin. = Administration episode 1 (N = 211) and 2 (N = 109).

Results following the Principal Components analysis validated the twofold
dimensionality of the goal orientation variable. The content validity of the
Goal-orientation Questionnaire is evidenced by the literally low/negative versus
high/positive loadings between the two extracted components at each scale. The
two-component solution is defined as follows: a) Component 1 stands for the MGO scale
which covers Items 1 to 10 and b) Component 2 stands for PGO scale which covers Items
11 to 20. Overall, results from the Item analysis and Principal Components analysis
respectively confirmed significant reliability and content validity for the Goal-orientation
Questionnaire. The data corroborated Hypothesis 1, consolidating goal orientation as a
research-worthy variable in a L2/FL classroom context.

Post-task Questionnaire results

Table 3 presents strong correlation results across the three areas which the questionnaire
was intended to measure. The highest of these positive correlation coefficients figured in
the ‘difficulty’ scale (r = .93, p < .01) and the lowest in the ‘motivation’ scale (r = .55, p
< .01). It is worth mentioning that all the intra-scale coefficients were bigger than those
between scales. For instance, the familiarity intra-scale correlations of Item 3 (e.g., r
= .75, p < .01) outweighed all the four inter-scale correlations (e.g., the highest being r

51
= .43, p <.05). These distribution patterns suggest that the nine items were largely representative of one area over the other two.

The results also evidenced the internal consistency of the scales was significantly high and that these scales independently measured what they were claimed to measure. However, Table 3 displays 13 instances of significant inter-scale coefficients where the difficulty/motivation concentration captured all of the highest, yet negative correlation values. The data suggested that high task difficulty was strongly related to the lack of motivation and vice versa.

Table 3 Correlations within and between the three Post-task Questionnaire scales

<table>
<thead>
<tr>
<th></th>
<th>Familiarity</th>
<th>Difficultly</th>
<th>Motivation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Item 1</td>
<td>Item 2</td>
<td>Item 3</td>
</tr>
<tr>
<td>Item 1</td>
<td>_</td>
<td>.75**</td>
<td>.66**</td>
</tr>
<tr>
<td>Item 2</td>
<td>_</td>
<td>.75**</td>
<td>.25</td>
</tr>
<tr>
<td>Item 3</td>
<td>_</td>
<td>.43*</td>
<td>.42*</td>
</tr>
<tr>
<td>Item 1</td>
<td>_</td>
<td>.93**</td>
<td>.71**</td>
</tr>
<tr>
<td>Item 2</td>
<td>_</td>
<td>.71**</td>
<td>-.63**</td>
</tr>
<tr>
<td>Item 3</td>
<td>_</td>
<td>- .35</td>
<td>-.44*</td>
</tr>
<tr>
<td>Item 1</td>
<td>_</td>
<td>.65**</td>
<td>.55**</td>
</tr>
<tr>
<td>Item 2</td>
<td>_</td>
<td>_</td>
<td>.70**</td>
</tr>
<tr>
<td>Item 3</td>
<td>_</td>
<td>_</td>
<td></td>
</tr>
</tbody>
</table>

Note: * p < .05 (2-tailed); ** p < .01 (2-tailed); N = 87.

The findings reported in Table 4 refer to the distribution of the participants’ assessment of the difficulty they experienced with the three tasks, regardless of their goal affiliation. The lowest estimates of difficulty spread over the three subscales of Task 3 such as DIFF 9 (M = 2.40, SD = 1.13). Because this task was subjected to both planned and familiar conditions, evidence for difficulty was least associated with the existence of familiarity and planning. Conversely, unplanned tasks (Task 2) and unfamiliar tasks (Task 1) were
found to be particularly associated with difficulty. More specifically, difficulty was strongly related to the lack of planning and less so to the lack of familiarity.

Results suggested that tasks become cognitively demanding (i) when performers have little background knowledge of similar tasks and (ii) even more strenuous when performers are given insufficient pre-task planning time.

Table 4 Distribution of perceived difficulty responses to the tasks

<table>
<thead>
<tr>
<th>Task 1</th>
<th>N</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>DIFF1</td>
<td>29</td>
<td>3.36</td>
<td>1.56</td>
</tr>
<tr>
<td>DIFF2</td>
<td>29</td>
<td>3.48</td>
<td>1.48</td>
</tr>
<tr>
<td>DIFF3</td>
<td>29</td>
<td>3.30</td>
<td>1.45</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Task 2</th>
<th>N</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>DIFF4</td>
<td>28</td>
<td>4.46</td>
<td>0.92</td>
</tr>
<tr>
<td>DIFF5</td>
<td>28</td>
<td>4.54</td>
<td>0.69</td>
</tr>
<tr>
<td>DIFF6</td>
<td>28</td>
<td>4.43</td>
<td>0.96</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Task 3</th>
<th>N</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>DIFF7</td>
<td>30</td>
<td>2.40</td>
<td>1.35</td>
</tr>
<tr>
<td>DIFF8</td>
<td>30</td>
<td>2.63</td>
<td>1.38</td>
</tr>
<tr>
<td>DIFF9</td>
<td>30</td>
<td>2.40</td>
<td>1.13</td>
</tr>
</tbody>
</table>

Note: N=87

ANOVA results in Table 5 showed a strong statistical significance for the effect of both goal orientation and task sequencing conditions on the participants’ motivation evidenced by the significant difference found in the effect of goal orientation on motivation ($F(1, 86) = 16.39, p < .05$). Such difference appeared to be even more significant than that of the task-originated conditions of familiarity and planning in their effect on motivation ($F(2, 87) = 4.84, p = .010$). More interestingly, the statistical difference in the effect of goal orientation and task conditions on motivation is equally significant ($F(2, 87) = 6.32, p = .003$). This statistical significance implies that one’s goal orientation interacts with one’s response to task demands, and so moderating one’s motivation for a given task. However, the analysis of variance did not determine which of the two goal orientation levels was more/less impervious to task conditions nor did it establish how motivation appears under each of these goal levels.
Table 5 Between-subjects effects of goal orientation and tasks on motivation

<table>
<thead>
<tr>
<th>Source</th>
<th>Mean Square</th>
<th>Df</th>
<th>F</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Goal orientation</td>
<td>19.68</td>
<td>1</td>
<td>16.39</td>
<td>.000</td>
</tr>
<tr>
<td>Task</td>
<td>5.80</td>
<td>2</td>
<td>4.84</td>
<td>.010</td>
</tr>
<tr>
<td>Goal x Task</td>
<td>7.58</td>
<td>2</td>
<td>6.32</td>
<td>.003</td>
</tr>
</tbody>
</table>

Note: The mean difference is significant at the .05 level

Follow-up descriptive statistics reported in Table 6 provide a more detailed picture of the relationship between goal orientation and motivation across the three experimental task conditions. Two patterns of difference were detected between the scores related to the highly demanding tasks (i.e., Task 1 and Task 2) and the scores related to less demanding tasks (i.e., Task 3). When it comes to the demanding task conditions, the MGO goal group showed more motivation during task engagement than the PGO goal group, especially when the narrative tasks were unfamiliar to the participants (i.e., MGO = 4.07 > PGO = 2.36 in Task 1). The reverse was the case for these participants when they performed Task 3 under the relaxed conditions of +familiarity/+ planning. This time, the PGO goal group showed more motivation than the MGO goal group (i.e., MGO = 3.87 < PGO = 4.12).

Table 6 Distribution of motivation scores among goal orientation groups

<table>
<thead>
<tr>
<th>Goal orientation</th>
<th>Task</th>
<th>Mean</th>
<th>SD</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>MGO</td>
<td>Task 1</td>
<td>4.07</td>
<td>.99</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>Task 2</td>
<td>3.86</td>
<td>1.29</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>Task 3</td>
<td>3.87</td>
<td>.83</td>
<td>15</td>
</tr>
<tr>
<td>PGO</td>
<td>Task 1</td>
<td>2.36</td>
<td>1.22</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>Task 2</td>
<td>2.50</td>
<td>1.23</td>
<td>15</td>
</tr>
<tr>
<td></td>
<td>Task 3</td>
<td>4.12</td>
<td>.96</td>
<td>15</td>
</tr>
</tbody>
</table>

Note: N =87

The patterns of difference between the two goal groups concerning their motivation echoed their changeable perceptions of task difficulty. Figure 1 displays evidence for how the operationalization of IDs, such as goal orientation, may offer a more complex
picture of the relationship between task difficulty and motivation. More specifically, the distribution of motivation in the first line plot (left) seems to draw a commonplace picture of motivation (i.e., high difficulty = low motivation). However, the second line plot decomposes this picture, intimating that the MGO goal group was reversing the initial difficulty-motivation pattern.

![Figure 1 Changes in motivation perceptions across task sequencing conditions](image)

These individuals showed a positive attitude towards the demanding nature of Tasks 1 (i.e., due to unfamiliarity) and 2 (i.e., due to lack of planning). However, their extent of motivation dropped conspicuously at the level of Task 3, the least demanding of the three tasks. In sum, their perception of difficulty was not a debilitating factor that hampered their task engagement, which was contrary to the PGO group whose motivation was affected negatively by the amount of difficulty the tasks. These results confirmed Hypothesis 2 that Goal orientation affects L2/FL learners’ motivation for and perceptions of task difficulty.

**Discussion**

Results from the 211 completed questionnaires confirmed the reliability and the consistency of the Goal-orientation Questionnaire as a sampling instrument. Equally verified over two administration episodes (Table 2) was the content validity of the questionnaire questions which proved their factorability into a two-component solution
yielded by factor analysis, building on a set of empirically verified instruments (e.g., Midgley et al., 1998; Skaalvik, 1997). The configuration of goal orientation aligned with the two-dimensional paradigm prevailing in goal literature, as opposed to competing models such as the multiple-goal model that subdivides PGO into *Performance-approach* orientation and *Performance-avoid* orientation (see Elliott, 1999).

The two-dimensional goal orientation profile supported by the present study somehow echoed Ortega’s categorization of L2/FL learners where the personality traits of the “communication-oriented” learners as risk-takers and process-focused are typical of MGO individuals and the personality-traits of “accuracy-oriented” learners are commonplace in the description of PGO individuals. Hence, delimiting goal orientation into two researchable units of analysis may add regularity and a firm footing for future L2/FL research attendant to the study of goal orientation as a central ID factor in the equation.

The psychometric properties of the Post-task questionnaire reliability and content validity were substantiated. These results confirmed the distribution of task conditions along the three tasks during the experimental phase (Task 1= - familiarity, Task 2 = - planning, and Task 3= +familiarity/+planning). The purpose of these results was to operationalize difficulty from an ID perspective. Instead of considering participant factors as an anomalous component in the process of defining task difficulty features, Elder et al. (2002) argued that “there may be some value in canvassing test-takers’ perceptions of task difficulty to determine how influential these are in test performance” (p. 350). The findings elicited by the Post-task Questionnaire in this study meshed with an “interactive approach” to task difficulty (Iwashita et al., 2001, p. 411) because the inherent features of difficulty cannot be literally dissected from individual differences. So, any changes in the effect of difficulty on performance were accounted for by comparing the performance of the two goal groups across the experimental task conditions.

The two line plots presented in Figure 1 illustrated how such variation figured along the two goal-orientation groups. This scope of variability has been further documented by interview data collected after the experimental procedure. The PGO interviewees reported their profound anxiety about and hypersensitivity towards errors due to task difficulty, and so they were overwhelmed by a self-defeating feeling that they
failed their tasks. Concurring with the Post-task Questionnaire data, they found task unfamiliarity as an indicator of difficulty. Three among the five PGO interviewees attributed their feeling of anxiety to their unfamiliarity with the task(s). However, the more familiar with task conditions, the less inhibited they felt. Where fear from making mistakes due to task difficulty was the most recurrent theme among the PGO participants, the MGO interviewees expressed a positive position towards errors being a natural feature of task difficulty. They demonstrated a relaxed attitude towards errors as a necessary ingredient in the language learning process. This stance has been substantially evidenced in goal research (e.g., Ames, 1992; Midgley et al., 1998).

The two goal orientation groups also reported completely disparate views of the tasks they performed. The PGO goal group systematically associated difficulty with anxiety whereas the MGO goal group considered difficulty necessary for learning. Therefore, what was motivating for one goal group was squarely disheartening for the other goal group. In this vein, some MGO participants voiced their lack of interest in performing unchallenging tasks, especially towards the end of the experiment (Task 3) while the PGO group showed more eagerness to perform the very tasks as they grew more familiar with them. This assertion accords with the account of Plough and Gass (1993) that task familiarity does not necessarily have a positive effect on task completion as because, as Skehan (1998) once posited, familiarity may make some type of learners feel the “staleness of doing something they might find unchallenging” (p. 113).

In light of the lack of a standard scale of task difficulty, the results of the Post-task Questionnaire did not only identify difficulty features post hoc, they also confirmed the a priori sequencing effect of task conditions such as familiarity and planning, as has been the standard in mainstream task-based research. In view of that, reporting the participants’ estimates of difficulty concurred with the study’s effort to give more substance to the scope of the learner differences in researching pedagogic tasks. The empirical support for the speakers’ perceptions of task difficulty has provided a posteriori estimate of difficulty to make sure that the operationalization of task conditions was not simplistically confined to preset estimates of difficulty (e.g., Skehan, 1998). Understanding that what seems to be highly difficult for one learner does not necessarily yield the same response pattern for another learner. In this study, the integration of goal
orientation in the statistical treatment of the data enabled us to capture the extent of ID variation in a systematic way.

**Conclusion**

The present study empirically attested to the active role of IDs in defining variable perceptions of difficulty and motivation among task takers. Building on an “extended task-based paradigm” (Dörnyei and Kormos, 2000, p. 297), the results established that goal orientation is a noteworthy ID factor in L2/FL classrooms. This psycholinguistic reality has been long obscured in mainstream task-based research which ironically has a tradition to lean on feeder disciplines like psychology. The results also suggested that one’s goal affiliation accounts for his/her perception of the difficulty of and motivation for a given task.

Indeed, some participants (MGO) developed a positive response to highly demanding tasks and their motivation exacerbated when they lost the sense of challenge with undemanding tasks. The converse was the case for another type of learners (PGO) whose sense of achievement depreciated in the face of demanding tasks, yet their motivation was reinstated when tasks were less taxing. Also the PGO participants showed more vulnerability to unfamiliar tasks than their MGO counterparts. In light of these results, the picture of task-based engagement seems to be more representative of the reality of L2/FL classroom although it remains to prove whether this type of ID variation has a significant bearing on the learners’ task output and their course of L2/FL development.

**References**


**Appendix A**

**Goal-orientation Questionnaire**

Circle one number for each statement to mark your level of (dis)agreement (5=Very true; 1=Very untrue).

| Section 1 |
|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|
| **SA** | **A** | **N** | **D** | **SD** |
| 1 | Challenging tasks that arouse my curiosity are important to me | 1 | 2 | 3 | 4 | 5 |
| 2 | I am confident I will do well in the speaking task no matter how difficult it is | 1 | 2 | 3 | 4 | 5 |
| 3 | It does not matter for me if the speaking task is graded | 1 | 2 | 3 | 4 | 5 |
| 4 | I prefer connecting task content to my personal experience | 1 | 2 | 3 | 4 | 5 |
| 5 | When I face difficulty in performing a speaking task, I always try different ways until it is finished | 1 | 2 | 3 | 4 | 5 |
| 6 | I do not mind making many mistakes if I learn from them | 1 | 2 | 3 | 4 | 5 |
| 7 | I feel more successful when I see my speaking skill improving | 1 | 2 | 3 | 4 | 5 |
| 8 | I like speaking tasks best when they make me learn new things | 1 | 2 | 3 | 4 | 5 |
| 9 | I prefer using notes rather than memorizing parts of the task content | 1 | 2 | 3 | 4 | 5 |
| 10 | I cannot be satisfied with my performance just because I receive a positive reaction from my teacher | 1 | 2 | 3 | 4 | 5 |

<p>| Section 2 |
|-----------------|-----------------|-----------------|-----------------|-----------------|
| <strong>SA</strong> | <strong>A</strong> | <strong>N</strong> | <strong>D</strong> | <strong>SD</strong> |
| 11 | Performing better than the other students makes me confident | 1 | 2 | 3 | 4 | 5 |
| 12 | The opinions my classmates hold about my speaking performance are so important to me | 1 | 2 | 3 | 4 | 5 |
| 13 | Demonstrating my speaking skills to others is always important | 1 | 2 | 3 | 4 | 5 |</p>
<table>
<thead>
<tr>
<th></th>
<th>I prefer my task performance to be graded only when I do well</th>
<th>1 2 3 4 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>15</td>
<td>I feel successful in my speaking task when I avoid many mistakes</td>
<td>1 2 3 4 5</td>
</tr>
<tr>
<td>16</td>
<td>I prefer memorizing to improvising in order to handle the difficult parts of the task</td>
<td>1 2 3 4 5</td>
</tr>
<tr>
<td>17</td>
<td>I feel more comfortable with familiar tasks rather than new ones</td>
<td>1 2 3 4 5</td>
</tr>
<tr>
<td>18</td>
<td>My constant fear of failure always motivates me to be successful</td>
<td>1 2 3 4 5</td>
</tr>
<tr>
<td>19</td>
<td>I do not want to take risks when I feel unable to complete the task</td>
<td>1 2 3 4 5</td>
</tr>
<tr>
<td>20</td>
<td>I prefer waiting to see how others perform the task so that I will not make the same mistakes</td>
<td>1 2 3 4 5</td>
</tr>
</tbody>
</table>

**Appendix B**

*Post-task Questionnaire*

Circle one number for each statement to mark your level of (dis)agreement (5 = Very true; 1 = Very untrue).

1) Before working on this task, I had the opportunity to do similar narrative tasks
   True 5 4 3 2 1 Untrue

2) Before working on this task, I was sure about what to do with the pictures
   True 5 4 3 2 1 Untrue

3) After working on this task, I expect to do more similar narrative tasks
   True 5 4 3 2 1 Untrue

4) Before working on this task, I expected it to be hard to complete
   True 5 4 3 2 1 Untrue

5) While working on this task, I felt it was hard to complete it
   True 5 4 3 2 1 Untrue

6) After working on this task, I felt relieved to complete it
   True 5 4 3 2 1 Untrue

7) Before working on this task, I felt it would be an interesting experience
   True 5 4 3 2 1 Untrue

8) While doing this task, I felt it was enjoyable to have some moments of pressure
   True 5 4 3 2 1 Untrue

9) After working on this task, I felt I was eager for more tasks like this
   True 5 4 3 2 1 Untrue
Designing a Computer-mediated, Task-based Syllabus: A Case Study in a Taiwanese EFL Tertiary Class

Julian Cheng Chiang Chen

University of Maryland, USA

Bio Data
Julian Cheng Chiang Chen recently earned his PHD in the Second Language Education and Culture program at the University of Maryland College Park, USA. He used to teach EFL in Taiwan at the college level for years. His research areas include computer-mediated communication, task-based language teaching and blended learning. He is also interested in exploring the effect of Second Life, a 3-D virtual learning environment, on ESL/EFL learners’ second language acquisition.

Abstract
Despite years of formal instruction, most Taiwanese tertiary students still cannot function in English spontaneously and are demotivated by commercially ready-made textbooks that fall short of addressing students’ real-life needs and interests. This case study was conducted in response to the concerns raised above with hopes that EFL students could use English in authentic contexts but not for artificial purposes imposed by forms-focused instruction. Operationalized under the 10 methodological principles (MPs) for the task-based language teaching (TBLT) syllabus design (Doughty & Long, 2003), this case study mirrored the six key components of a TBLT design: (a) needs and means analysis, (b) syllabus design, (c) materials design, (d) methodology and pedagogy, (e) testing, and (f) evaluation. In particular, it capitalized on Taiwanese students’ cultural competence in L1 as a springboard for channeling their L1 knowledge into L2 production. Due to the booming popularity of computer-mediated communication (CMC) in the digital age, this study incorporated blogs as a platform for students to jointly construct their sites, interact with peers, transfer their background knowledge from L1 to L2 in the process of task completion, and to develop their communicative and cross-cultural competence in a collaborative virtual community. Content analysis was employed to illustrate the process of how this case study was carried out under the TBLT framework. Overall evaluation of this TBLT syllabus was highlighted by students’ blogging vignettes, followed by pedagogical implications for English teachers who are interested in adopting CMC task-based instruction in EFL contexts.
Keywords
Task-based language teaching, computer-mediated communication, blog, project-based learning, EFL context

Introduction
Despite the fact that the pendulum of teaching methods in SLA history has shifted from a behavioral drill-and-practice approach (e.g., grammar-translation and audio-lingual methods) to a communicative and learner-centered approach (e.g., communicative language teaching and task-based instruction), the former still dominates how English should be taught and learned in most EFL contexts (S. Chen, 2002; Li, 1998). Even though most Asian EFL countries (e.g., Taiwan, China, Korea and Japan) have started to enact English education at the elementary school level with a focus on oral communication skills (Butler, 2004), conventional teacher-led instruction, centering on grammar-translation and pattern-drilled activities, still drives English education at the high school level (Q. Chen, 1999; Silver, Hu, & Iino, 2002). The flip side of this phenomenon is also epitomized in Seargeant’s (2009) critical analysis of English education in Japan from the perspectives of globalization and language ideology in that most Japanese students still have little success in mastering English or do not find it necessary to use English on a daily basis, despite the Japanese Ministry of Education’s effort to stress the importance that “…in the 21st century, it is essential for them to acquire communication abilities in English as a common international language” (pp. 6-7). Among those complex ideological factors that impact the success of English education with a communicative approach in Asian EFL contexts (e.g., cultural, political, linguistic, etc.), the socio-political factor, such as examination-oriented educational system, nevertheless, still leads the EFL instruction back to the form-focused continuum.

A case in point is that English textbooks used at the secondary school level in Taiwan, despite the communicative components partially incorporated, still follow the linguistically pre-set recipes in order to prepare students to pass the college entrance exam or in F. Su’s (2000) term, “teach to test” (p. 110). Consequently, students are demotivated by those commercially ready-made textbooks that incorporate “oversimplified…inauthentic communicative structure and unrealistic situational content” (Bartlett, 2005). The phenomenon has greatly impacted Taiwanese EFL
students’ motivations and attitudes toward English learning, which unfortunately are dampened by the grammar-oriented and test-driven teaching pedagogy (S. Chen, 2002; Zhang, 2004). Taken together, the socio-historical, political and cultural constraints of traditional teaching method, big class size and English as not a major means of communication (Savignon & Wang, 2003) fall short of addressing the real needs and interests of Taiwanese EFL students, thereby failing them to function in English spontaneously for communicative purposes despite years of formal instruction.

Today’s EFL teaching and learning at the tertiary level in Taiwan, however, does not deviate away from the form-focused instruction. Even though current Taiwanese college students have learned English since elementary school (Y. Su, 2006) and are free from the pressure of the entrance examination, Freshman English mandated in most colleges still emphasizes the receptive skill, reading, with a focus on rote drills and grammar translation (Chern, 2003). Since students are only required to take English for credits in their freshman year, coupled with the fact that English is not the official language spoken in Taiwan, this status quo may explain why most Taiwanese EFL tertiary students are less motivated to learn English and switch back to their mother tongue after they walk out the English class door. Speaking with their peers in English on the street may also seem awkward to them since Chinese is still the medium for daily communication.

The anxiety of speaking English in real-life situations can also be exemplified in an awkward scenario where Taiwanese EFL students are embarrassingly lost for words when approached by foreigners asking for directions on the street. As Horwitz, Horwitz, and Cope (1986) stressed, “…any performance in the L2 is likely to challenge an individual's self-concept as a competent communicator and lead to reticence, self-consciousness, fear, or even panic” (p. 128). This phenomenon points to the fact that the grammatical rules and pattern drills to which they have been taught since high school are unfortunately not “translated” into spontaneous oral communication in real-life situations (Cheng, 2000). As a result, opportunities provided for Taiwanese EFL students to use English for communicative purposes in real-life settings are still lacking. In order to mitigate the aforementioned discrepancies, a call for a more learner-centered and
task-based course design that suits Taiwanese EFL students’ current and future needs of using English in the real world is desired.

This case study, therefore, was conducted in response to both the concerns raised above and my ambition to design a task-based instruction with hopes that my Taiwanese tertiary students could learn/use English for their real needs in authentic contexts but not for artificial purposes, imposed by forms-focused instruction. In particular, it was carried out under the 10 methodological principles (MPs) for a task-based language teaching (TBLT) syllabus design proposed by Doughty and Long (2003). Additionally, this study mirrored the six key components of a TBLT syllabus design: (a) needs and means analysis, (b) syllabus design, (c) materials design, (d) methodology and pedagogy, (e) testing, and (f) evaluation (Doughty & Long, 2003, p. 50). It also capitalized on Taiwanese students’ native culture as a springboard for channeling their L1 knowledge into L2 production since native culture-related themes were more meaningful and familiar to them. The task type (i.e., introducing Taiwanese culture to foreigners) also triggered students’ motivation and matched their desirable goal to have opportunities to accomplish target tasks in the real world around them.

Also worth noting in current EFL contexts is the burgeoning popularity of pedagogical application of teaching English in computer-mediated communication (CMC) environments (C. Chen, 2005; Chun, 2003; Cummings, 2004; Dunbar, Linklater, & Oakey, 2000). The CMC settings, be they email exchanges, blogs or instant messengers (IMs), open up a more interactive and dynamic arena for learners to connect to the world, regardless of the time and distance (Warschauer, 1997). CMC is also closely in tune with students’ day-to-day practices, evidenced in their frequent use of emails, IMs, Skype, blogs, and chat rooms, for example. The implementation of TBLT in CMC environments, therefore, sheds pedagogical light on the promising effects of marrying these two approaches (Blake, 2000; Gonzalez-Lloret, 2003). Motivated by the computer-mediated, TBLT approach (Doughty & Long, 2003), this project incorporated blogs as a platform to provide an optimal environment for students to jointly construct their sites, interact with peers, transfer their background knowledge from L1 to L2 in the process of task completion, and develop their communicative and cross-cultural competence in a collaborative virtual community.
Purpose

Based on the rationale specified above, the purpose of this case study was twofold:

1. Examined how a TBLT course embedded in Taiwanese EFL students’ native culture could boost their motivation and raise their cross-cultural awareness with the content bearing on their background knowledge.

2. Explored how the 10 MPs under the TBLT framework could inform this case study, optimize students’ oral proficiency and communicative competence and be realized in students’ blogging and project-based learning.

Theoretical Framework

Task-based Language Teaching (TBLT)

Seeing the pedagogical potential that a task-based, communicative approach can prepare English language learners to “use” English for authentic purposes through interactive and real-world tasks, a growing number of English teachers have started to gear their instruction toward TBLT (Skehan, 2003). Drawing on the CMC nature of interaction and TBLT course design, this study was grounded in 10 TBLT methodological principles (Doughty & Long, 2003) informed by Interaction Hypothesis (Long, 1981, 1983, 1985; Pica, 1987; Pica & Doughty, 1985; Swain, 1995). To begin with, the pedagogical implication of Interaction Hypothesis for EFL instruction is that it opens the door for EFL learners to consolidate their linguistic repertoire and fine-tune their current interlanguage competence through the two-way negotiation for meaning with their peers or native speakers (Long, 1981, 1983, 1985). However, this ideal language acquisition will not take place without EFL teachers providing students with the opportunity to try out the target language in the classroom, let alone using it outside the classroom (Willis, 1996). Well-designed tasks that incorporate the key elements of Interaction Hypothesis, in this case, can facilitate the whole interactional process to realize language acquisition of EFL learners.

TBLT, hence, has the potential to bring language learning to the real life, if effectively implemented into the teaching pedagogy. In addition, the CMC environment offers a fertile ground for interaction, which maximizes TBLT and erases the in/out-of-class boundary (Jarvis, 2005; Warschauer, Turbee, & Roberts, 1996). Despite
the advantages that task-based CMC can endow EFL learners, it is a pity that this innovative approach has either not been fully applied to most Taiwanese EFL classrooms, or has been inappropriately adopted. For example, some Taiwanese EFL teachers may misconstrue “tasks” as “exercises,” or design so-called “task-oriented” activities to disguise the intention to inherently practice linguistic elements in a “fun” way (Ellis, 2000). As a result, the implementation of TBLT in CMC within the interactionist SLA framework is still lacking in Taiwanese EFL classrooms.

Aligning with the interactionist model, TBLT seems to be a promising pedagogy for L2 acquisition to take place. Nevertheless, how can it be implemented into the internet-based environment to enhance learning and facilitate teaching in and outside the classroom? Also, under what conditions can a task-based approach mediated by computer-assisted language learning (CALL) be more effective for L2 acquisition? With respect to these concerns regarding the pedagogical application of TBLT supported by CALL, Doughty and Long (2003) proposed 10 teaching methodological principles (MPs) for CALL as illustrated in the screenshot table below (Doughty & Long, 2003, p. 52):

<table>
<thead>
<tr>
<th>Table 1. Language Teaching Methodological Principles for CALL</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>ACTIVITIES</strong></td>
</tr>
<tr>
<td>MP1 Use tasks, not texts, as the unit of analysis.</td>
</tr>
<tr>
<td>MP2 Promote learning by doing.</td>
</tr>
<tr>
<td><strong>INPUT</strong></td>
</tr>
<tr>
<td>MP4 Provide rich (not impoverished) input.</td>
</tr>
<tr>
<td><strong>LEARNING PROCESSES</strong></td>
</tr>
<tr>
<td>MP6 Focus on form.</td>
</tr>
<tr>
<td>MP7 Provide negative feedback.</td>
</tr>
<tr>
<td>MP8 Respect “learner syllabuses”/develop-mental processes.</td>
</tr>
<tr>
<td>MP9 Promote cooperative/ collaborative learning.</td>
</tr>
<tr>
<td><strong>LEARNERS</strong></td>
</tr>
</tbody>
</table>

(This table is adapted from Doughty 2000b, 2001b.)
The 10 MPs provide an insightfully pedagogical framework for language teachers when they design task-based oriented syllabi in CALL environments. For example, the principles of “use tasks, not texts” (MP1), “learning by doing” (MP2), “cooperative learning” (MP9) and “individualized instruction” (MP10) foreground the importance of creating opportunities for learners to actually use English for meaningful purposes and communicative needs in a language class. The 10 MPs also resonate with the interactionist model to further operationalize how second language can be acquired through interaction in CALL environments. The principles of providing rich input (MP4) without simplifying it (MP3), focusing on form, not forms (MP6) and fine-tuning language output provided by negative feedback (MP7) also align TBLT design with the interactionist SLA theory. Within the CALL paradigm, the nature of CMC stands apart from the other CALL tools in that it offers a more ideal interface for interpersonal interaction, and the incorporation of TBLT in CMC can further enhance L2 acquisition as well (Chapelle, 2005). In order to address the aforementioned issues and get a better understanding of the underlying constructs of TBLT and CMC as well as how this blended approach can foster EFL learners’ engagement and language development, this case study intended to test this hypothesis, situated in 10 TBLT MPs mediated by blogging in the Taiwanese EFL tertiary setting.

Research Questions
This case study aims to answer two research questions:
1. To what extent could oral proficiency, communicative and cross-cultural competence of Taiwanese tertiary EFL students be enhanced by a TBLT syllabus tapping into their L1 cultural knowledge?
2. How could blogging in project-based learning be orchestrated to support TBLT syllabus design and collaborative learning as well as serve as a criterion-referenced assessment?

Setting and Participants
There were overall 25 Taiwanese EFL students (around 12-13 each term) enrolled in the course “Practical English II” with a focus on oral proficiency, presentation skills, and
communicative competence in two consecutive semesters. Their majors were across departments (e.g., business administration, civil engineering, electrical engineering and architecture) in a technical university in Northern Taiwan. Student ages ranged from 19 to 22, at the intermediate level of English proficiency. Class met once a week, two hours per session. In the initial stage of the course, nearly half of the students showed difficulty expressing ideas in English due to either insufficient chances to speak in their former English instruction, or limited English oral proficiency. Students were informed at the outset of the course that this course was task-oriented and culture-embedded. It aimed to suit their needs of using English to introduce Taiwanese culture to foreigners visiting Taiwan and to provide them with more opportunities to use English in the real world. Despite the fact that a textbook was still used in class, the content of the textbook was closely related to the theme of this task-based course with the focus on how students presented and fulfilled various interactive real-life tasks assigned to them instead of mechanically following the scripted text and rote drilling.

In addition, they were told that they would create their own blog sites to accomplish the project task. They were also informed that the course was part of this research study and their blog entries and materials used in the class would be analyzed only for the purpose of this research and for the possibility of publication. Moreover, they were told that one of the assigned tasks, a street interview project, was designed not only for the sake of this inquiry, but also for the benefits of getting the chance to practice English with foreigners in a real-life situation. To add, they were made aware that the oral interview data they presented in class would only be shared within the class. As such, no oral recording data in which students interacted with their foreign interviewees is used in this study. For the sake of confidentiality, the identities of the foreigners who agreed to do street interview with the students are not revealed in this study since they were not included in the data analysis process to begin with. Only students' personal reflections on the real-life task-based experience were analyzed from their blog entries. In addition, all copies of the oral recording data were destroyed and students' blog links were removed by the teacher/researcher after the data analysis was completed. They were also assured that their real Chinese names would not be revealed in order to protect their identities. Informed consent was received from students before each class started.
Methods
Following Doughty and Long’s (2003) 10 TBLT principles of the design, implementation and evaluation operationalized in the six components mentioned above, I will exemplify how this case study can be carried out under this framework, how each stage mirrors the 10 MPs, and present overall evaluation of this TBLT syllabus highlighted by students’ blogging vignettes.

Design
1. Needs Analysis
Different from the other mandatory English courses with a pre-set textbook and syllabus regulated by the university, this elective course “Practical English II” was not obliged to the same requirements. As such, I, as the instructor, had the flexibility to design my own task-based syllabus and materials. The class in the first semester was a pilot study to seek information about 1) the extent to which a TBLT syllabus could match students’ needs of English learning in Taiwan, 2) their attitudes toward assigned tasks, 3) their self-perceptions about the progress of communicative and cross-cultural competence through task completion, and 4) the effect of using blogs to supplement TBLT design. In order to more appropriately address learners’ current and future needs, triangulation of various sources and methods was employed: a) informal interviews with students about their learning needs, b) instructor’s end-of-course evaluation with students, and c) course documents and students’ project work. A case in point is that students were required to post their thoughts about whether their expectation of this course and their needs had been met. Data collected in the first semester also served as a baseline for the needs analysis for the class in the second semester.

2. Syllabus and Materials
The main theme of this course was to introduce various aspects of Taiwanese culture and to promote Taiwanese hospitality to foreign friends so that they could not only enjoy their stay in Taiwan, but get to know the culture and locals better. To illustrate this, the real-life tasks students needed to fulfill ranged from helping foreigners with traveling to different tourist spots, informing them of where to have fun and enjoy authentic local Taiwanese food, and answering questions raised by foreigners about the origins and
celebrations of holidays and festivals. Aligned with the needs analysis, the task-based syllabus was categorized into three major task types: 1) introducing local Taiwanese food, 2) showing tourist attractions in Taiwan, and 3) explaining origins of Taiwanese holidays and festivals (see Table 2 in Appendix A). Inspired by the module of TBLT material development in a Korean TBLT project at the University of Hawaii (Chaudron et al., 2005), the corresponding pedagogic tasks in Table 2 were also sequenced in accordance with task complexity as the levels of task difficulty gradually increased toward the completion of the target task. The interaction of task complexity, difficulty and complexity was carefully manipulated in order to scaffold students with sufficient input and fine-tune their modified output before moving them toward more cognitive demanding tasks (Robinson, 2007).

In the initial stage of sequenced pedagogic tasks (see examples of pedagogic tasks 1 and 2 [denoted as PT1, PT2] in Table 2), for example, the goal was to provide students with rich exposure to genuine and elaborated input, registered in target discourse samples gleaned from students’ street interview data in the first semester. They were informed to focus on listening as a warm-up and not to worry about new vocabulary or terms, which they would pick up through each task completion. There was also more visual support (i.e., pictures, maps) to direct students’ attention to resources in elements of a pedagogic task. They were simply required to identify or label the distributed authentic pictures or maps with objects to which they listened in the recordings of discourse samples. The purpose was to build up their confidence and acquaint them with the pedagogic tasks they were to deal with before moving on to more complex tasks. With the degree of task complexity and difficulty increasing across task conditions, the task demands on them were accumulating in PT 3 and PT4. Despite the fact that they still collaborated in pairs or groups, each pair and group held different pieces of information and they needed to elicit the other required elements and contribute their parts in order to fulfill the intended task. Productive oral skill was more emphasized and orchestrated in a two-way, close-ended jigsaw activity (Pica & Doughty, 1985).

In the target task 1 (i.e., introducing local Taiwanese food), each group watched different video scenarios of how to make local food and contributed their new knowledge as to the ingredients and procedure and jointly filled out the worksheet on the missing
information elicited from the other groups. In the target task 2 (i.e., showing tourist attractions in Taiwan), students first gathered information in their expert group based on the authentic tourist brochure and guidebook excerpt along with a metro transit map and train schedule. After that, each person from the expert group had to tell the other members newly formed in the home group about the means of transportation and interesting facts about the tourist spot in order to complete the worksheet. In the target task 3 (i.e., explaining origins of Taiwanese holidays and festivals), a Find Someone Who activity was employed in that a piece of elaborated excerpt about Taiwanese festivals (modified from the Taiwanese Tourist Bureau site) and a matrix of Taiwanese festivals were given to each student where some information was missing so students needed to go around the class to find other students who had the information he/she did not have in order to fill out the chart categorized in date, festive food, origins and customs.

The final task PT 5/PT 6 (for target task 1 only) was designed to approximate the real target task and served as the exit test. In this final stage, the demands of task complexity and difficulty reached to the fullest where visual and contextual supports were absent and students needed to work individually instead of seeking assistance from pairs or group members as they did in previous pedagogic tasks. Building on the new language knowledge constructed in previous tasks, they simulated scenes involving foreigners sampling local food versus food stand owners selling food in night markets (target task 1), planning and reporting a one-day itinerary for foreigners visiting different tourist attractions in Taiwan (target task 2), and show-and-telling a cultural artifact or souvenir that represented the custom or tradition of a Taiwanese festival (target task 3).

**Implementation**

1. Testing

In addition to the aforementioned PT5 and PT6 as the exit test, there were also three types of assessment employed to gauge students’ actual performance in the course of implementing the task-based syllabus:

A. Simulated interview

In the midterm period of the course, a one-on-one interview between each individual student and me was administered as a summative assessment. S/he had to randomly select
a paper strip to determine which target task was to perform and then was directed to the PowerPoint slides projected on a big screen indicating the target discourse situation s/he just picked. All the situations related to the three target tasks were stored in a PowerPoint folder and presented to students in random order. For example, the interviewee who was allocated to the target task 2 (showing tourist attractions in Taiwan) would see a slide of a foreigner asking a passerby how to get to different tourist spots in one day. The next slide would show the components of completing the task (e.g., transportation and itinerary). S/he was asked to play the role of the passerby whereas I would play the role of the foreigner. Follow-up questions would be posed if necessary. They had 45 seconds to organize their thoughts and five minutes to simulate the scenario with me.

B. Street Interview

Another exit test that attempted to replicate genuine conditions to assess students’ oral proficiency and communicative competence was the project of interviewing foreigners on the street, which was assigned in the second half of the course. Students were required to work in pairs and integrate what they had learned in class to design questions surrounding the three target tasks of Taiwanese culture that were of interest to them. It also served as a needs analysis for them before they worked on their blog sites and incorporated the data they collected in their street interview. They were required to tape record all the interview process and to report the findings in a PowerPoint format regarding their interview questions, cross-cultural differences and similarities, and lessons learned from the interview.

Also worth noting is that the “street interview project” where students collected data about how much foreigners knew about Taiwan before designing their blog sites also benefited the needs analysis of this course. That is, their interview data also became the target discourse samples of what the real-life scenarios looked like when they conversed with foreign friends in English across culture-specific topics. These authentic samples also foregrounded the syllabus design for various pedagogical tasks gradually leading to the completion of authentic target tasks (see Chaudron et al., 2005, for a full version of the process of collecting discourse samples).
C. Taiwan Blogger

As addressed in the second research question, blogs were blended into this case study in order to maximize the effect of the TBLT syllabus. It also served as a criterion-referenced assessment, which not only tapped into the repertoire of students’ cultural knowledge and new language knowledge acquired through completion of pedagogic tasks, but tuned into the CMC tool that further broadened their language use. Students in pairs were required to set up their blog sites to document their learning progress throughout the whole semester. They were also informed that their English blog sites were to help foreigners better understand Taiwan and provide their foreign friends with hands-on information regarding different aspects of Taiwanese culture. In this sense, they needed to select a Taiwanese culture theme related to any of the three major target tasks that each group was interested in presenting and researched relevant information on government-owned English sites targeting foreign tourists.

Students were asked to design their blogs by composing the content and posting personal thoughts and summaries about their findings supported by multimedia, such as pictures, audio and video files (some pairs had videotaped their street interviews and uploaded them to their sites). Each pair was also assigned to visit their peers’ blog sites and to comment on how foreigners could benefit from their sites. For the final evaluation, each pair presented their blog sites projected on the big screen and orally reported to the class the process of doing the blog project as well as their reflections on what they had learned from this experience.

2. Methodology and Pedagogy

In response to the 10 TBLT methodological principles, the implementation of the task-based syllabus in this case study mirrors the 10 MPs to certain extent. Aligning with MP1, the core of this syllabus design was to develop and sequence meaningful authentic tasks that Taiwanese EFL learners needed to perform in real-life situations (Ellis, 2000; Skehan, 2003). Different from a conventional EFL syllabus focusing on compartmentalized language elements and tedious rote drilling, this task-based syllabus viewed language learning “… as a living entity through using it and experiencing its use during task completion” (Doughty and Long, 2003, p. 56). The aim of this TBLT syllabus
was to allow students to capitalize on their cultural knowledge and to learn the target language step by step through accomplishing the three target tasks with the increased demands of complexity and difficulty (Robinson, 2007). The spirit of “learning by doing” (MP2) also manifests itself in this case study where students were required to accomplish each problem-solving pedagogic task and project-based task (e.g., street interview and Taiwan blogger) in pairs or groups (Long, 2009). It further triggered their intrinsic motivation in that all the tasks were interesting but also challenging to them as opposed to mechanically practicing those linguistic items bit by bit. On a positive note to MP3, all pedagogic tasks were grounded in modified, sequenced and elaborated input through target discourse sample collection (i.e., street interview data and authentic artifacts) rather than artificially simplified texts mostly used in commercial textbooks (Long, 1996). Most pedagogic tasks and project work required students to negotiate meaning through task completion (e.g., jigsaw activity) as well.

Another positive note is that the class was conducted in the target language, meaning that I only used English in instruction and all the materials were also in English rather than Chinese-English translation. Even though students at the outset of the course were not used to the whole English instruction (since they had been taught in grammar-translation method for years and few EFL teachers would use English throughout the course), they eventually adapted themselves to the intensive exposure to rich input (i.e., teacher talk and elaborated texts from authentic documents) as opposed to impoverished input (MP4) (Long, 1985; Pica & Doughty, 1985). Given the fact that this syllabus was task-oriented, not form-focused, students felt for the first time that they could finally learn English for a meaningful purpose, not just for grammar per se. Inevitably, students still made certain grammatical errors during their oral production. However, I did not interrupt their “trial-and-error period” where they explored and internalized the chunks of new language, such as cultural terms and expressions they learned through task completion (MP5) (Wray, 2000). Instead, I would fine-tune their oral production by providing recasts to keep the conversation going without correcting them on the spot that might “traumatize” their risk-taking potential to produce oral output in the future (Long, 2007). After hearing the same non-target like productions occurring among most learners, I would draw their attention to the common errors they had made
by asking them to find out the mistakes in sentences on the board with global correction if necessary (MP7). In addition, the discourse samples in pedagogic tasks also raised their awareness of those “trouble spots” when they heard the non-target like output produced by their peers in the street interview with foreigners (Swain, 1995). Once again, the whole form-focused instruction would not last long but only be activated in response to their current interlanguage system (MP6) (Long, 1981, 1983).

The learner-centered task-based syllabus that accommodated their cultural knowledge and took into account their internal syllabus without forcibly feeding them on undigested linguistic recipe also echoed the MP8 (Lightbown, 1983; Pienemann, 1984). As mentioned above, most problem-solving pedagogic tasks required pair or group work to facilitate the process of negotiation for meaning where language acquisition could take place. The project tasks (i.e., street interview and Taiwan blogger) also promoted collaborative learning in that less proficient learners could benefit from the feedback provided by proficient learners through interaction whereas the latter could strengthen their language output by elaborating on it for their counterparts (MP9) (Oxford, 1997; Pica & Doughty, 1985). Last but not least, this task-based syllabus was designed to suit the real needs of Taiwanese EFL learners through initial needs analysis. It also aimed to make instruction more individualized and language learning more meaningful by inviting learners to accomplish each motivating, learner-centered and communicative task (MP10) (Sawyer & Ranta, 2001).

**Evaluation**

This TBLT case study embedded in Taiwanese culture was piloted and evaluated in the first semester through informal interviews with students by the end of the course, coupled with students’ self-reports on their progress and opinions about the task-based syllabus design. They were required to post their thoughts on their blog sites and comment on the work done by their peers as well. The feedback gleaned from the class in the first half of the year-long project served as a baseline to modify the content, materials and tasks for the second half of the project. As an EFL instructor, I was also eager to know the extent to which the innovative task-oriented syllabus design could really benefit my students’ learning progress, which further informed my teaching. I was pleasantly surprised to
learn that most students in the first semester positively commented on how much they had learned from this TBLT course and could finally use English in the real world. As Maria, a student in the first class, vividly illustrated how a task-based course could not only improve her English oral proficiency, but foster her cultural competence as well:

This is my first time [to] use English [to] write article[s], and in our life, we can’t speak English all the time...[Our team's final project] is “betel nut girls,” I think we have a lot of fun. We can know about things [we didn't know before], and start to know what’s Taiwan[ese] culture...especially [if] it happens in our [daily life]. I found I [grew] up in this class, because I try to talk to foreigners, I try to speak English with assertiveness. I really like this English class.

Maria’s voice also echoed those students in the second class in a sense that grounding a task-based syllabus in EFL learners’ cultural knowledge could facilitate the speed of language acquisition since they could use their background knowledge to compensate for their limited English proficiency (Lantolf & Pavelenko, 1995; Lantolf, 2006). It also respected their internal syllabus as opposed to cramming them with undigested, cold and isolated linguistic components as is done in most traditional EFL classes (Li, 1998). Furthermore, incorporating students’ project work on blog sites further optimized the TBLT syllabus. Not only could they see their learning progress and achievement by documenting each task completion through blogging, but they also received member support and encouragement from their peers and me in the virtual community. The evidence of cooperative/collaborative learning in students’ project work, highlighted by peer support, further verifies the importance of scaffolding, from the sociocultural perspective that also informs task-based instruction (Ellis, 2000; Skehan, 2003).

When presenting their blog sites in the final group report, both classes were also amazed at how many tasks they had accomplished and how much progress they had made when compared with the first time they entered the class. As Maggie, a student in the second class, delineated,

This semester, I can feel my English better than before. In the practical English, I have more [first-hand] experiences, such as [doing a] presentation in English, shar[ing] my thought[s] in English, and mak[ing] an English blog. In class, we have to talk with classmates and Julian in English, so it pushes me [to] try to speak in English. I also know many Taiwanese things. I can talk about holidays, foods, and tourist places in English. As for this final project, mak[ing] an
English blog about Taiwan, I think it is a good idea. We can surf classmates’ blogs to know more Taiwanese information.

Maggie’s vivid account of her being “pushed” to use English in order to get her meaning across also validates the notion of pushed output, a vital vehicle for language acquisition to take place as claimed by most cognitive interactionists (Long, 1981, 1983, 1985; Pica, 1987; Pica & Doughty, 1985; Swain, 1995). When reflecting on which task benefited them the most, both classes unanimously singled out the “street interview,” the criterion-referenced assessment administered in the second half of the course. Due to the previously mentioned Asian EFL phenomenon where EFL students seldom have the chance to speak English with foreigners, they found it relatively challenging but felt highly motivated and “fresh” to embark on this adventurous task. They also gained a sense of achievement of being able to build on their cultural knowledge and practices in pedagogic tasks to help them interact with a foreign friend while introducing Taiwanese culture at the same time. As Jacky in the second class stressed,

[I]'s very practical by doing some difficult but interesting assignment…especially [the] street interview…I used to feel nervous to chat with foreigners. But through this I become natural in front of them. No more pressure, and talk[ing] to them [is] like [having a] normal conversation. Julian also taught us how to introduce Taiwan. When foreigner[s] ask "what is Taiwan?" I usually stammer[ed] and [gave] some inconsequential answers. But now I can [finally] say something about it.

Jacky’s reflective thoughts also mirror vital methodological principles in TBLT that can benefit EFL students’ language development through the real-life task of actually interviewing foreigners on the street (MP1: Use task, not text and MP2: Learning by doing) (Ellis, 2000; Long, 2009; Robinson, 2007; Skehan, 2003). Also, this TBLT syllabus was designed to tailor to the needs and interests of my EFL students (Needs analysis, see Doughty & Long, 2003, p. 55) which not only boosted their motivation, but fostered their sense of achievement during the task completion (MP10: Individualize instruction) (Sawyer & Ranta, 2001). Additionally, Jacky, like most of his peers, was able to capitalize on his cultural knowledge about Taiwan, which served as a springboard to
leverage his interlanguage repertoire when interacting with his foreign friends in English (Lantolf & Pavelenko, 1995).

Although most students seemed to express positive attitudes toward the overall task-based learning experience, their initial encounter with foreigners, nevertheless, did not turn out as positive and smooth as expected. In their blog entries, they reported the issue of approaching their “potential interviewees,” coupled with “non-understanding/gap” noticed in their language output when reflecting on using English to interview foreigners for the first time. Two students, Daniel and Benson, commented on the difficulty they came across when trying to approach foreigners for their street interview assignment:

We started to find a foreigner to help us in NTNU. Many foreigners study Chinese in [the] language center [there]. In the beginning we met three foreigners, but they refused us. We felt so frustrated.

Their frustration of failing to reach out to foreigners not only increased their anxiety about accomplishing the assigned task, but further exposed them to a risk-taking, real-world communication context outside of class. Their self-consciousness of performing L2 vis-à-vis a native English speaker was heightened by the affective factor (e.g., fear of speaking English), unfamiliar linguistic and sociocultural register in L2 (e.g., uncertainty about the appropriate context of utterance), and vulnerable self-esteem (e.g., low confidence of handling the spontaneity in authentic communication in L2) (Horwitz, Horwitz, & Cope, 1986, p. 128). Additionally, the communication breakdown occurring during their interview with their interviewees in English also raised a red flag that the “language gap” still existed in their current interlanguage system (Long, 1983; Swain, 1995). They were appalled to realize that their current English speaking proficiency did not prepare them well to tackle the real-life task, despite years of practice in conventional English classes. For example, Daniel and Benson, the two aforementioned students, raised their metalinguistic awareness when listening to their audio recording of their street interview and reflecting on the whole experience. They noticed that their English pronunciation had become the “trouble spot” that caused the communication breakdown when they finally got the chance to test it out with their foreign friend in real life:
After this project of interview, I really have some deep feeling[s]. That is, correct pronunciation is really important. Because when I review my recording again, even I can’t understand what I was talking about, let alone that foreigner[s] can understand me as well. Furthermore, when we had small talk with Robert, he taught us how to use exact words to describe something. For example, beard, goatee and whiskers are all different... Therefore, to pronounce correctly when you communicate with foreigner[s]... is very important.

Despite the fact that most students initially felt that the street interview task was daunting due to the aforementioned language anxiety in speaking and interacting with foreigners, they nevertheless voiced that they found the effort they put into the task rewarding and meaningful. If we keep following Daniel and Benson’s case, their reflective report epitomizes the effect TBLT (e.g., street interview in their case) can have on bringing language learning to life by transferring what they have learned in class to real-life contexts. To illustrate, after finishing their street interview with the native speaker, Robert, Daniel and Benson befriended him and planned a one-day trip for their foreign friend. They invited him to a local night market and accompanied him to several tourist attractions in Taipei. To their pleasant surprise, they did not expect that they could make friends with a native speaker simply by doing an assigned task. They took on the role of professional local tour guides to show him around Taipei, take him to sample a variety of local food, and explain to him the history and customs of Taiwanese festivals and customs—the exact three target tasks required in this TBLT course that were realized in the local Taiwanese context. As Daniel vividly stated (also see Appendix B3 for a full version of the street interview and itinerary originally posted on their blog site),

We have planned this visit for our exchange student friend and invite[d] Robert to join us. That day, I was the leader of this visit and also the tour guide. I was really nervous, because I must use whole English to explain and introduce to foreigner friends about Lung Shan Temple... Even if it’s a little challenging to transfer the Chinese history to English, I still think it’s very interesting. Because this was my first time to be a tour guide, I never thought I would use English all the way. Haha... When we finish this tour, Robert really enjoyed this visit to Lung Shan Temple and thanked us for our treat. We knew each other more, and made friend[s] with him. I think we really made it!!
Concluding Remarks

Implementing a task-based syllabus in a Taiwanese EFL tertiary class has taught me a valuable lesson as much as my students have learned. Despite the overall positive results from course evaluation and class observation, there is still room to improve. For example, I could have administered a questionnaire to elicit more specific information about how this course had fulfilled students’ needs in terms of the assigned pedagogic tasks and project work. I could also have done follow-up semi-structured interviews with students in both classes, inquiring about their experience of interacting with foreigners after finishing the course and what other suggestions that they thought would have made the syllabus more complete and tasks more effective. The information collected from both sources could have made the needs analysis more precisely aimed at how pedagogic tasks could be designed to approximate the real target tasks.

Another lesson I have learned from this case study is the transferability of task-based abilities. When compared with the effect of task-based assessment, traditional linguistic and forms-focused testing that focuses on compartmentalized aspects of linguistic elements, unfortunately, does not accurately predict the learners’ ability to successfully perform language tasks in the real world. Given the fact that the spirit of task-based evaluation is the performance and completion of the task itself, tasks should be motivating and tailored to the learner’s real performance of tasks required in real life (Long & Crookes, 1992). As such, my students came to realize that they could finally apply what they had learned and practiced in class to the real-life scenarios of interacting with foreign friends in English on the street. All the tasks became meaningful and engaging to them and they could “kill two birds with one stone” since they practiced English and accomplished real-life tasks required of this course all at once (Ellis, 2000; Long, 2009; Sawyer & Ranta, 2001; Skehan, 2003).

Following the same thread of thought of using English for communicative purposes in the real world, it also draws my attention to the concept of using English as a Lingua Franca (ELF) realized in contexts where English is spoken as the common language between speakers whose mother tongue is not English (Jenkins, 2007). Our students are very likely in the future to interact with other speakers whose native language may or may not be English but use the target language as a means for
communication. Jenkins’s (2003) statistical survey on “Who speaks English today” also reveals that there are more nonnative English speakers who use ELF for various communication purposes (e.g., business, study, travel) than native English speakers in the world. As such, each interlocutor brings his/her sociolinguistic repertoire into the two-way interactional channel, dictated by regional varieties in L2 (e.g., accents, intonations, expressions, pronunciations and sentence structures; see Jenkins, 2003, p. 23). In this ELF vein, we, as EFL teachers, should raise our students’ awareness that the “localized” variety in socio-cultural/linguistic forms should be valued and legitimized when they interact with other speakers of ELF or even native English speakers (Jenkins, 2007, p. 19). By the TBLT token, since the ultimate goal of our students is whether or not they can eventually accomplish a real-life task, it is the “intelligibility” in communication that facilitates the task completion that really matters, rather than being constrained by normalizing English to the so-called “standard English” or “native accent” (ibid.). Therefore, how to incorporate the socioculturally/linguistically localized elements into TBLT design in the context of the global uses of Englishes should be the vital task for our EFL teachers, as opposed to only viewing communication as “universal” under the native English norms without taking “culturally specific values” into account (Seargeant, 2009, p. 59).

Apparently, designing a TBLT syllabus places more demands on some EFL teachers who have been too used to following a teaching recipe that is more scripted and rote-drilling. Maybe we EFL teachers should ask ourselves the question, “Is what I teach actually what my students desire to learn?” The sense of achievement flashing in my students’ eyes after they completed each task, strengthened by witnessing their real progress in communicative and cultural competence, has made me believe that a task-based syllabus incorporating students’ L1 cultural repertoire has the potential to empower both EFL teachers and students, if we do try to bring it to life.

Notes

1 For the sake of reader-friendliness, corrections on typos and major grammatical errors were made on students’ verbatim comments.
In order to facilitate their research process for the blog site project and provide more authentic target language input, students were referred to different English sites published by Taiwanese government, such as the Government information office (http://www.gio.gov.tw), the Tourism Bureau (http://admin.taiwan.net.tw/english), and the Council for Economic Planning and Development (http://cepd.goc.tw/encontent).

Again, to protect the students’ identities, their Chinese names were not used, the blog address link was deleted, and all the photos were blurred. For the sake of confidentiality, a pseudonym was chosen for the foreigner whom they interviewed and his photographic image was also altered to protect his identity. The typos and grammatical errors were kept as they appeared in their original blog site to respect authorship.

References


Butler, Y. G. (2004). What level of English proficiency do elementary school teachers need to attain in order to teach EFL? Case studies from Korea, Taiwan, and Japan. TESOL Quarterly, 38, 245-278.


85


UK: Wiley-Blackwell.


### Appendix A: Table 2 Scope and Sequence of Task-based Syllabus

<table>
<thead>
<tr>
<th>TASK TYPE</th>
<th>Showcase Taiwanese Culture to Foreigners</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>TARGET TASK</strong></td>
<td><strong>Introducing local Taiwanese food</strong></td>
</tr>
<tr>
<td><strong>PEDAGOGIC TASK (PT)</strong></td>
<td>1. Listen to different genuine discourse samples of foreigners talking about local food they like to eat in Taiwan.</td>
</tr>
<tr>
<td></td>
<td>2. Divide students in groups and distribute to each group pictures of local Taiwanese food. Listen again to fragments of deliberately chosen elaborated scenarios where students need to identify the food mentioned in the vignettes with the correct picture.</td>
</tr>
<tr>
<td></td>
<td>3. Listen to several examples of foreigners’ opinions (recycled samples from PT1) about why they like and do not like Taiwanese food. Students in pairs rematch the pictures into two groups (i.e., food foreigners like vs. do not like). Compare answers with the other pairs and discuss the results based on the vignettes.</td>
</tr>
<tr>
<td></td>
<td>4. Play videos of how to make Taiwanese food (e.g., rice dumplings, moon cakes). One group (expert group) stays in class to watch the video and the other groups stay outside. The group in class writes down the ingredients and procedure and shares their information with each other. By the same token, another group outside the class can come in to watch.</td>
</tr>
<tr>
<td></td>
<td>5. Play videos of how to make Taiwanese food (e.g., rice dumplings, moon cakes). One group (expert group) stays in class to watch the video and the other groups stay outside. The group in class writes down the ingredients and procedure and shares their information with each other. By the same token, another group outside the class can come in to watch.</td>
</tr>
</tbody>
</table>
the video while other groups wait outside for their turn. After each expert group is finished, reform new groups (home groups) with different expert group members and share with the new home group members what they saw in the video.

5. Each pair chooses one local Taiwanese food and comes up with a recipe as to how to make this local food.

6. Each student brings a local Taiwanese food. Half of the students will play the role of foreigners trying to sample the food in the night market whereas the other half plays owners selling food in the local food stands. They are required to attract customers by introducing why the food is tasty, the ingredients and how to make the food. Students take turns playing the food stand owners.

schedule and train schedule will be given as well. Each group (expert group) is assigned to focus on one tourist spot and read the information about the place and how to get there by taking public transportation. A new group (home group) consisting different members from other expert groups will be formed and each member will tell their home group members about the sightseeing spot they were earlier assigned to regarding the transportation and why it is worth visiting.

5. Each student pair plans a one-day itinerary for foreigners who are visiting different tourist attractions in Taiwan. The itinerary will specify the schedule, transportation, destination, activities and dining. Each pair reports their itinerary to the whole class.

Taiwanese Tourist Bureau. A matrix of Taiwanese holidays/festivals is given to each student where some information is missing so students need to go around the class to find other students who have the information h/she does not have in order to fill out the chart. Whoever gets all chart blanks filled wins.

5. Each student is assigned to bring a cultural artifact or souvenir that represents the custom or tradition of the festival. S/he has to tell the story about the origin and celebration of the festival together with some interesting facts about it.
Appendix B: Sample of students’ street interview and follow-up itinerary

us that our questions have to be logical and organized. We design a questionnaire before we interview with a foreigner. Our main idea is about festival. We conclude five points to compare the most important festival between Taiwan and America.
We start to find a foreigner to help us in NTNU. Many foreigners study Chinese in language center. In the beginning we meet three foreigners, but they refuse us. We feel so frustrated. So we have to find another way to solve it. Our strategy is to meet foreigner inside language center.
Our applied English class teacher asks for us must have a street interview with native speaker, we go to NTNUs’ language center and know our dear friend, Robert, he comes from in Los Angeles, has been to Taiwan for three months. His major is philosophy and come to Taiwan for leaning Chinese. He also is a vegetarian.

When we ask Robert, our street interview friends, what’s the most impressive festival in Taiwan? His first direct reaction is Taiwan’s election. Wow… Taiwan election phenomenon is so famous to foreigners. The election and politics in Taiwan is really special compare with America, maybe it also represents Taiwan’s important festival sometimes. People are enthusiastic at politics, gather together to announce something, use speaker, picture, any kind of equipment to help them. It really makes our environment more prosperous, but noisy.
Main Idea

- Festival
- Chinese New Year & Christmas
- Happy Holidays
- From Thanksgiving to Christmas.

Our main idea is to compare Chinese New Year with Christmas. After listen to our recording we know Happy Holidays is from Thanksgiving to Christmas. Its purpose is to include more people to celebrate something.

Main Idea – Compare I

<table>
<thead>
<tr>
<th></th>
<th>Chinese</th>
<th>America</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Purpose</strong></td>
<td>A new beginning</td>
<td>Celebrate Jesus’ birth</td>
</tr>
<tr>
<td></td>
<td>Family</td>
<td>Family</td>
</tr>
<tr>
<td><strong>Food</strong></td>
<td>Boiled-dumpling</td>
<td>Goose</td>
</tr>
<tr>
<td></td>
<td>New Year’s rice cake</td>
<td>Pudding/ Custard/ Christmas cake</td>
</tr>
<tr>
<td></td>
<td>Fruit cake</td>
<td></td>
</tr>
<tr>
<td><strong>Activities</strong></td>
<td>Family reunion</td>
<td>Christmas party</td>
</tr>
<tr>
<td></td>
<td>Return natal home</td>
<td>Christmas trees/lights/ decoration</td>
</tr>
<tr>
<td></td>
<td>Set off fireworks</td>
<td>Decoration: Jesus’ model</td>
</tr>
<tr>
<td></td>
<td>New clothes</td>
<td>Christmas carols &amp; Go to church</td>
</tr>
<tr>
<td></td>
<td>Clean house</td>
<td></td>
</tr>
</tbody>
</table>

Purpose Connection:
As Benson mention, we want to compare with the two biggest festival in America and Taiwan. These two festivals are the most important in these two countries. We want to compare the difference through five parts: purpose, food, activities, gift and market. The red color word is the difference between our original thoughts before
After this interview, Robert gives us his experience and answer. It has some difference with our original thought, so we headline these points and focus them.

The first part purpose, in Taiwan, it is symbolic of a new begging. In America, it’s for Celebrate Jesus’ birth. But these two festivals’ main ideas are for family. Family can get together in this time and have a family reunion.

Food connection:
In Taiwan, boiled dumpling is always the main dish, and people will go to market buy rice cake. For America, in the first, we think turkey is also for Christmas, but Robert tells us that goose is more for Christmas and turkey is more for Thanksgiving. Food cake is very popular in Christmas, people will make cake by themselves in their house and send it to neighbor or others.

Activities connection:

<table>
<thead>
<tr>
<th>Main Idea – Compare II</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Chinese</strong></td>
</tr>
<tr>
<td>Gift</td>
</tr>
<tr>
<td>Put money into red envelope</td>
</tr>
<tr>
<td>Market</td>
</tr>
<tr>
<td>Market is noisy and thriving</td>
</tr>
<tr>
<td>Spring couples’ decorations</td>
</tr>
<tr>
<td>New Year’s merchandise</td>
</tr>
<tr>
<td>Watermelon seeds/ Clothes/ Fielded/ Snacks</td>
</tr>
</tbody>
</table>

Market:
In Taiwan, the market always very thriving and noisy compare with America, stores sell spring couplets, decorations and New Year’s merchandise. Such as watermelon seeds, clothes, pickles and snacks. But in America is really different with Taiwan, we can see many kinds of toys on sell. Parent will buy toys be Christmas gift for their children. We also can see many Santa Claus in front of stores, it’s not only for sell but also help them to sell product.

Gift connection:
Robert teaches us that kind of sock call stocking to put gift into it. It’s also very traditional. They will write their name on stocking. We really appreciate Robert’s help, we invite him for a dinner together and treat him. We go to a vegetarian restaurant and have a small talk. We really make friends with each other. He is really a very kind person. So we also invite him to join our foreigner program, we plan a culture travel to Lung Shan Temple on 12/18 and leave the phone number and mail. Keep contact.

We have planed this visit for our exchange student friend and invite Robert join us. That day, I am the leader of this visit and am also the tour guide. I really a little
nervous, because I must use whole English explain and introduction to foreigner friends about Lung Shan Temple. The first picture is that I am introducing history for them. Even it’s a little challenge to transfer the Chinese history to English, but I still think it’s very interesting. Because this is my first time to be a tour guide and never think will use English. Haha…When we finish this visit, Robert really enjoys this visit and thanks for our treat. We know each other more, and make friend with him. I think we really do it!!

After this project of interview, I really have some feeling deeply. That’s pronunciation correctly really important, because when I review my record again. Even I can’t understand what I talking about, foreigner can’t understand you as well. Furthermore, when we have small talk with Robert, he taught us how to use exactly words to describe something. Such as bear, goatee and whisker are all different. I think this time is not only a street interview homework, but also a culture exchange. Through the interview, we can know more about local foreign culture.
To pronounce correctly when you communicate with foreigner. Oral expression is very important. I did street interview twice, I also take another English course this semester. I think "street interview" is a very useful tool to make friend with foreigner. Next time when I have free, I don't take English courses anymore. Maybe I'll go to NTNU to make some foreigner friends.

There are many people study Chinese in NTNU. If I can go to other countries, I can see many fresh things. It could help me to expand my vision and to experience different lifestyles.
L2 Phone-based Interaction (PBI) and Development of Communicative Competence: A Case Study of an Adult’s English Learning in EFL Context

Katie Kim

**Georgetown University, USA**

**University of New Hampshire, USA**

**Bio Data**

Katie Kim is a Ph.D candidate of the Department of Linguistics (Concentration: Applied Linguistics) at Georgetown University and lecturer of Department of English at University of New Hampshire (Linguistics/TESOL program), starting in fall 2012. Her research interest includes, but not limited to, second language acquisition, bilingualism, psycholinguistics, and teaching English to speakers of other languages.

**Abstract**

While L2 interaction over the telephone is growing method (Hong, 2008) for L2 learning in some EFL countries (e.g., Korea, Japan, China), there has been no research investigating whether and how phone-based speech communication contributes to a learner’s L2 speaking ability. The purpose of this case study is to investigate the effectiveness of phone-based L2 interaction (PBI) by examining development of an EFL learner’s L2 communicative competence and increase of the learners’ willingness to communicate (WTC). In the study, a Korean EFL learner participated in 10 minutes of daily telephone conversations for two months exclusively in English. The learner’s development of communicative competence was assessed for the two-month period. To measure the development of communicative competence, the learner’s daily conversations were analyzed based on the Canale's (1983) theoretical framework, which includes four criteria assessing communicative competence: grammatical competence, sociolinguistic competence, discourse competence, and strategic competence. Also, to measure the increase of WTC (McCrosky, 1977; 1992; 1997), the
learner’s changes in affective variables, such as motivation, L2 anxiety and self-confidence, were taken into consideration. The result of this study suggests that PBI can be effective for developing low-WTC EFL learners’ communicative competence and for increasing those learners’ WTC.

**Keywords:** phone-based interaction, case study, communicative competence, willingness to communicate, qualitative study, EFL context, adult second language learning

**Introduction**

For the last two decades, phone-based English interaction (PBI) has been a highly popular way to improve English oral proficiency in Korea. According to news reports (Hong, 2008), more than 160 PBI companies have been established, and they currently invest approximately 200 million dollars a year for that business. However, no empirical research on its effectiveness to develop learners’ L2 communicative ability has been conducted, which is precisely the goal of this paper. Specifically, the current study attempts to investigate the effectiveness of PBI to develop EFL students’ communicative competence.

Since communicative competence was recognized as a primary goal of L2 learning (Brown, 2001; Hadley, 2001; Campbell & Wales, 1970; Canale & Swain, 1980), authentic interaction involving genuine use of L2 (Finoccharo & Brumfit, 1983) has been regarded as a fundamental element facilitating learners’ L2 oral proficiency. However, unlike ESL contexts, EFL contexts do not have ready-made authentic L2 beyond language classrooms (Brown, 2001). Also, even in the classroom, L2 students receive limited number of speaking turns (Cheon, 2003). Moreover, when the learners do not have a fair degree of willingness to communicate (McCrosky, 1977, 1992, 1997), those learners might have even more reduced chances for L2 interaction.

The current study investigated the telephone as a pedagogical tool to provide genuine L2 interaction for any individual EFL learners at any time. The study tried to
determine if PBI can bring any positive effects on communicative ability, especially for those learners who have less significant success in the L2 classroom due to their lower willingness to communicate (WTC). Thus, the purpose of this study is to investigate the influence of PBI on EFL learners’ development of communicative competence and the changes in WTC that trigger learners’ voluntary participation in L2 interaction.

**Communicative competence as a goal of language learning**

From the failure of the teaching approaches of the 1960s and 1970s (i.e., Audiolingualism, Grammar Translation) in developing learners’ communicative ability in real-life situations, the early educationalists noticed the limitation of form focusing approach (Chen, 2005) in positively affecting language learning. As a reaction to previous methods, Communicative Language Teaching (CLT) (Brown, 2001; Hadley, 2001) appeared in the 1980s with an emphasis on communicative properties of second or foreign language (L2). Unlike the previous approaches, CLT stressed the development of “communicative competence,” and since then, the notion of L2 proficiency moved from structural accuracy (Hadley, 2001, p. 8) to “the ability to function in a truly communicative setting—that is, in a dynamic exchange in linguistics competence” (Savignon, 1972, p. 8).

CLT emphasized meaningful interaction and communication with other people in the target language (TL), and this oral proficiency was represented by “communicative competence” (Campbell & Wales, 1970; Canale & Swain, 1980; Hymes, 1972; Munby, 1978; Charolles, 1978; Savignon, 1972; Widdowson, 1978) as an ultimate goal. Communicative competence consists of some combination of organizational competence (grammatical and discourse), pragmatic competence (functional and sociolinguistic), strategic competence, and psychomotor skills.
It is important to note that the components of communicative competence not only include L2 accuracy as a product of grammatical competence and psychomotor skill (pronunciation and intonation), but also involves L2 fluency in authentic language context represented by discourse and pragmatic competence. Also, strategic competence in communicative competence indicates the importance of unrehearsed L2 use to express the learners’ communicative needs in the real world (Brown, 2001, p. 69).

Even though CLT was introduced almost three decades ago, Brown (2001, p.42) stated that CLT is still one of the currently recognized approaches as a generally accepted norm to develop real-life communication skill in L2 learning and teaching field. In other words, communicative competence has become as a central notion of oral proficiency in language teaching and learning.

**Conversational interaction for communicative competence development**

According to Canale (1983, p.5), communicative competence refers to “the underlying systems of knowledge and skill required for communication,” and his theoretical framework of communicative competence involves learners’ grammatical competence (the degree of the learner’s mastery of the linguistic code), sociolinguistic competence (the degree of the learner’s understanding of appropriate language use in various contexts), discourse competence (the ability to combine ideas cohesively and coherently), and strategic competence (the linguistic ability to compensate the breakdown in communication). As these four components show, communicative competence puts more weight on the learners’ L2 production, or how they actually produce the TL in different contexts, rather than on the learners’ L2 knowledge, or what they know about it.
Finoccharo & Brumfit (1983) demonstrated that TL is acquired effectively when the learners are involved in L2 interaction that encourages them to have negotiation of meaning. Regarding L2 learning, interaction and negotiation of meaning, the interactionist perspective in SLA explains how those elements are related, and how they contribute to the learners’ L2 acquisition.

Pica (1994, p. 494) defines negotiation as “the modification and restructuring of interaction that occurs when learners and interlocutors anticipate, perceive, or experience difficulties in message comprehensibility.” Also, according to Gass (1997, p. 107), “negotiation refers to communication in which participants’ attention is focused on resolving a communication problem as opposed to communication in which there is a free flowing exchange of information.” In this sense, negotiation of meaning takes place when a learner’s interlangauge (IL) knowledge does not match with that of TL (Gass & Selinker, 2001, p.291), and the process of communication seems to involve the negotiation of meaning.

In terms of the significance of negotiation of meaning for the learners’ communication ability, input (Krashen, 1985) and interaction (Long, 1980, 1981, 1983) seem to play crucial roles. Varonis & Gass (1985) observed that negotiation of meaning in conversation happens when the interlocutor interrupts the flow of the conversation to clarify what the conversation is about, and those negotiations are more frequent when the conversations involve more non-native speakers of the TL. Therefore, negotiation of meaning plays a role in making received input comprehensible, and it is realized through the interaction among the interlocutors. Moreover, corrective feedback (Mackey & Oliver, 2002; McDonough, 2005; Philp, 2003; Lyster, 1998) during interaction substantially contributes to the learners’ IL development in terms of noticing the gap between the IL and the TL (Schmidt & Frota, 1986), and bringing their IL system closer to the TL system (Swain, 1985).
However, according to Swain (1985, 1995), input, interaction, and interaction feedback are insufficient to develop an L2, and comprehensible output is important to developing their L2 competence. Swain stated that (1985, p. 252) output provides “opportunities for contextualized, meaningful use, to test out hypothesis about the target language” for the learners. Through comprehensible output (Swain, 1985, p. 249) learners can be “pushed toward the delivery of a message that is not only conveyed, but that is conveyed precisely, coherently, and appropriately.” Therefore, the new L2 knowledge that the learners acquired through the negotiation of meaning seems to be internalized and restructured more effectively when the learners actually produce it.

In sum, comprehensible input, interaction, and comprehensible output all appear to benefit the learner’s L2 oral proficiency. However, as Canale’s four components of communicative competence show, input and interaction should be provided with the consideration of natural and genuine use of L2, and the learners’ output should be expected to be appropriate in its use. As a result, input, interaction, and output will be the most effectively applied in CLT context, when the genuine use of L2 is reflected in them.

**Developing communicative competence in EFL context**

**The issue of authentic language input and interaction.** Unlike ESL context, EFL context does not have opportunities to use the target language readily available outside of the classroom. Therefore, the countries in EFL context, such as Japan, Korea, Thailand and China, have extremely limited opportunities for authentic language input, interaction and output. In other words, EFL learners do not seem to experience “genuine uses for English in their own lives” (Brown, 2001, p. 117), and it hinders meaningful language communication. Due to the lack of communicative use of
language, EFL learners possibly have severe limitation to develop their L2 fluency when it is compared to their ESL counterparts, who have greater amounts of natural L2 exposure and opportunities to practice in their actual lives.

According to Cheon (2003), EFL learners receive a limited number of speaking turns in the traditional L2 classroom because the speaking turns are shared by a large number of language learners. That is, EFL learners might have minimal opportunities for active interaction even in the L2 classroom context due to the limited chances of speaking turns. Consequently, EFL learners seem to severely suffer from the lack of opportunities for authentic language interaction (Cheon, 2003), which results in very little success in L2 oral fluency.

The issue of peripheral participation in smaller class. Even in smaller EFL classes with more speaking turns with language teachers and peers, some learners with lower willingness to communicate (WTC) might experience difficulties in participating in the L2 interaction in the classroom. WTC (McCrosky, 1977, 1992, 1997) refers to the probability of engaging in communication when the opportunity is given to the interlocutors, and is first developed in L1 communication. In WTC, personal traits such as introversion, self-esteem, communication competence, communication apprehension and cultural diversity (McCroskey & Richmond, 1990) are regarded as the antecedents, which determine an individual’s “predispositions toward verbal behavior” (Mortensen et al., 1977).

The influence of WTC in L2 learning contexts was also studied by some communication experts (e.g., McCroskey et al., 1985a, 1985b), and the path model of L2 WTC was developed by MacIntyre (1994). As MacIntyre and Charos’ (1996) WTC path model in Figure 1 indicates, learners’ affective variables, such as negative perception of their L2 competence and high L2 anxiety and low motivation, can contribute to lower L2 WTC. Also, MacIntyre and Charos (1996, p. 18) noted that
introversion is also another significant factor raising the learners’ L2 anxiety; in other words, the learners with extrovert personalities might have relatively lower L2 anxiety compared to the learners with introvert personalities. Figure 1 illustrates that lower WTC has negative influence on L2 communication frequency. Therefore, some learners who have introvert personalities, higher L2 anxiety, lower motivation and negative perception on their L2 competence might lose the opportunities for genuine language practice even in smaller classrooms.
The phenomenon of lower WTC in L2 classrooms seems to be even more severe in East Asian EFL contexts due to the effect of face-protected predisposition. Wen and Clement (2003) pointed out that Chinese learners are less likely to be active in communication in L2 and to be sensitive to the judgment of the public as a result of the influence of face-protection orientation of Confucianism. Because of the pervasive Confucianism throughout East Asian countries, some EFL learners with relatively enough motivation to learn English might also have an unwillingness to communicate in English in the classroom when they are aware of their lower L2 oral proficiency; they possibly want to protect their face in the classroom by not showing their less fluent L2 competence.

Thus, even for some classroom learners who have enough motivation, the increased numbers of speaking turns in a smaller EFL class does not appear to promise frequent L2 oral production, when they have high self-face protection. Such learners might need a different language learning environment that can reduce the negative effects of the affective factors and protect their face simultaneously.

Advent of technology in language learning

Pedagogical benefits and limitations of CMC. With the advent of the enabling technology, computer-mediated communication (CMC) was proposed to “compensate the deficiency of interaction both inside and outside of EFL classrooms” by using such tools as “e-mails, chat rooms, discussion rooms, video or audio conference,” which make possible synchronous or asynchronous communication with no boundaries of time and space (Chen, 2005). According to previous research on the benefits of CMC
in language learning, CMC engages learners in authentic social L2 interaction to practice TL (Blake, 2000; Campbell, 2004; Leh, 1999; Lightbown & Spada, 1999). Also, learners can learn more L2 pragmatic knowledge by communicating in target social and cultural contexts (Chen, 2005) and improve L2 proficiency in more student-centered learning environments (Braine, 2004). Thus, students can be involved in more “authentic and meaningful L2 interaction with other L2 users worldwide via the internet” (Chen, 2005), and it can also promote the learners’ motivation of L2 learning, which ultimately helps the learners to gain learner autonomy (Blin, 1999; Lightbown & Spada, 1999; Toyoda, 2001).

Yet the relationship between the L2 oral fluency development and synchronous or asynchronous CMC in the language learning classroom has been investigated very rarely (Payne & Whitney, 2002). Indeed, Payne & Whitney (2002, p. 25), which found some benefits of CMC for oral proficiency, argued that their study does not suggest that “speaking skills can be developed in the absence of face-to-face conversational interaction,” but that chatroom sessions can only “serve as a communicator stimulator for foreign language learners.” It seems that written communication in CMC can be necessary but not sufficient to facilitate the learner’s L2 speaking ability. In other words, CMC does not appear to substitute for the role of oral interaction, which enforces learners to produce genuine oral output to develop L2 fluency. Thus, whether or not the text-based language interaction directly promotes L2 oral fluency is still a debatable issue.

Another problem related with CMC in language learning is that there are some EFL classrooms that have no access to the technology. Also, the younger or older EFL students who do not have the ability to use computers (i.e., slow typing in L2, knowledge of online resources, ability to read and write, etc.) would have difficulties interacting in cyber space. Even though voice chat rooms are able to minimize some
disadvantages related to text-based communication, voice chatting requires more equipment (i.e., headphones, microphones, computers, voice chat interface application, web camera and fast, reliable cable modem connection to the Internet) to start computer-based interaction. Therefore, students in the classrooms without technological support might not experience the facilitative role of CMC activities in language learning. Consequently, CMC does not appear to be an optimal learner-friendly environment in general.

**Pedagogical use of telephone in language learning.** In 1966, Ordinate Corporation (www.ordinate.com), a company developing assessment tools to evaluate learners’ use of spoken English, first introduced telephones as a measuring tool for speakers’ L2 speaking and listening language skill. This L2 assessment system is called PhonePass, and it administers a test over the phone for 10 minutes. During those 10 minutes, PhonePass provides a number of interactive tasks, such as reading aloud, repeating sentences, producing antonyms of cue words, and answering questions. The learners’ interaction over the phone is graded by a computer system, and the results of the test reports the test takers’ overall scores in speaking ability and subscores in listening vocabulary, repeat accuracy, pronunciation, reading fluency, and repeat fluency (Lazaraton, 2001. p. 112).

Another use of the telephone in L2 pedagogy can be found in phone-based English interaction (PBI), which was introduced in Korea in 1990s. According to news reports (Hong, 2008), its market has expanded in accelerated speed, and more than 160 PBI companies have been established due in part to the learners’ preference for English speaking ability. Now, Korea estimates that the companies invest more than two hundred million dollars for this business annually (Hong, 2008). Hong (2008) reported that among the advantages of PBI are fewer space restrictions and fewer scheduling difficulties for English communication on the phone, which always enables
one-on-one oral conversation between an English speaker and an EFL learner. Table 1 below lists basic information about the PBI services offered by three different companies in Korea.

Table 1 Basic information of PBI in Korea

<table>
<thead>
<tr>
<th></th>
<th>Company A</th>
<th>Company B</th>
<th>Company C</th>
</tr>
</thead>
<tbody>
<tr>
<td>Class</td>
<td>Beginner, Intermediate, Advanced, Business</td>
<td>Level 0 (novice) – Level 7 (proficient)</td>
<td>Level 1 – Level 3</td>
</tr>
<tr>
<td>Contents</td>
<td>practical grammar checkup</td>
<td>pronunciation training</td>
<td>vocabulary training</td>
</tr>
<tr>
<td></td>
<td>one-on-one conversation</td>
<td>vocabulary learning</td>
<td>intonation training</td>
</tr>
<tr>
<td></td>
<td>diary/writing checkup</td>
<td>one-on-one conversation</td>
<td>read aloud</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>pronunciation correction</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>one-on-one conversation</td>
</tr>
<tr>
<td>Schedule</td>
<td>10 or 20 minutes a day</td>
<td>10 or 20 minutes a day</td>
<td>10 minutes a day</td>
</tr>
<tr>
<td>Price (month)</td>
<td>3 days a week: 130,000 won</td>
<td>3 days: 60,000 won</td>
<td>3 days: 120,000 won</td>
</tr>
<tr>
<td></td>
<td>(1,100 won = appx.) 5 days a week: 150,000 won</td>
<td>5 days: 100,000 won</td>
<td>5 days: 180,000 won</td>
</tr>
<tr>
<td>Textbook</td>
<td>developed by the company</td>
<td>developed by the company</td>
<td>developed by the company</td>
</tr>
<tr>
<td>English Speakers</td>
<td>Native English speakers in Korea</td>
<td>Fluent English speakers in the Philippines</td>
<td>Native English speakers in Canada</td>
</tr>
</tbody>
</table>

However, there has been no research investigating the role of a telephone as a technological tool to develop the learners’ L2 oral proficiency. Also, there is a very limited amount of research studying how telephone-based speech communication
contributes to a learner’s L2 speaking ability. Regarding PBI’s characteristic one-on-one conversation, it might provide more time of L2 production than L2 classrooms do. Also, in terms of providing each learner’s own private sphere to talk in English, such English learners with low WTC might feel less reluctance to start their English conversations.

Regarding those expecting benefits of a telephone in language learning, the present case study attempts to determine whether an adult EFL learner with low WTC can develop English communicative ability through 40 days of PBI.

For this, the current study investigated the following research questions:
1. Is PBI effective for developing the communicative competence of a low WTC EFL learner?
2. Does PBI affect the EFL learner’s WTC? If it has positive or negative influence on the learner’s WTC, what characteristics of PBI cause that result?

Method
The current case study investigated an EFL learner’s development of communicative competence through 10 minutes of PBI for two months. The session was given five days a week, Monday through Friday, for eight weeks for a total of 40 days. The following subsections provide more detailed information about participants, equipment and materials, procedures, data collection and data analysis.

The Participant
There was one participant, whom we will call Min, who has learned English in Korean classrooms for two years. He is a 26-year-old Korean adult majoring in Electronic Engineering at Hong-ik University in Seoul, Korea. What follows is detailed information about his prior English ability, prior experience in English education,
personality, aptitude, anxiety and motivation. Table 2 below indicates the various techniques used in this current research to collect his biodata before treatments. The interviews and questionnaires were completed in his L1, Korean, to gather more accurate and reliable data. Some English-translated questions in those questionnaires are attached in Appendix A and appendix B.

Table 2 Potential learner-internal variables and techniques

<table>
<thead>
<tr>
<th>Learner-internal variables</th>
<th>Techniques to collect the data</th>
</tr>
</thead>
<tbody>
<tr>
<td>Motivation for English learning</td>
<td>Questionnaire &amp; Interview</td>
</tr>
<tr>
<td>Degree of anxiety</td>
<td>Questionnaire &amp; Interview</td>
</tr>
<tr>
<td>Language aptitude</td>
<td>Modern Language Aptitude Test (MLAT)</td>
</tr>
<tr>
<td>Personality factors</td>
<td>Myers-Brigg Types Indicator (MBTI)</td>
</tr>
</tbody>
</table>

**Prior English Ability.** For his self-evaluation in English ability, Min graded his reading ability at the highest level (8 points out of 10), speaking ability the second highest level (5 points), and they were followed by listening skill (3.5 points) and writing skill (2.5). Also, he reported that he is more confident in grammar than other aspects of language, such as speaking, writing and listening. However, based on TEPS (Test of English Proficiency Developed by Seoul National University), which was taken on December 2nd, 2007, he had a lower score in grammar compared to other English abilities. Table 3 below shows the information about his self-evaluation and the standardized English test taken in Korea.

Table 3 Min's prior English ability

<table>
<thead>
<tr>
<th>Domain</th>
<th>Assessment</th>
<th>Listening</th>
<th>Grammar</th>
<th>Vocabulary</th>
<th>Writing</th>
<th>Reading</th>
<th>Speaking</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Self-evaluation</strong></td>
<td>3.5 points</td>
<td>have confidence</td>
<td>—</td>
<td>2.5</td>
<td>8</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>---------------------</td>
<td>-----------</td>
<td>----------------</td>
<td>---</td>
<td>-----</td>
<td>---</td>
<td>---</td>
<td></td>
</tr>
<tr>
<td>(out of 10)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Overall**: 7 points out of 10

<table>
<thead>
<tr>
<th><strong>TEPS</strong></th>
<th>Level 2*</th>
<th>Level 3+**</th>
<th>Level 2*</th>
<th>—</th>
<th>Level 2*</th>
<th>—</th>
</tr>
</thead>
</table>

**Overall**: Level 2*

* Level 2 indicates high intermediate level of communicative competence.

** Level 3+ indicates mid intermediate level of communicative competence

— indicates no evaluation was given in that criterion.

The TEPS score report shows that Min’s overall English ability is at level 2, which is high intermediate communicative competence. Because TEPS does not assess test takers’ speaking ability, it is not clear if his oral proficiency also satisfies the same average level of English.

**Prior Experience in English.** Min has no experience studying abroad in English speaking countries, but has had two years of experience in formal English learning; he attended two different institutes in 2005 and in 2007. The interview with him found out that the grammar-based classroom in 2007 did not include any L2 interaction, and that the output-driven classroom in 2005 adopted more activities based on the audiolingual method (i.e., listen and repeat and memorizing the dialogue). He reported that neither the grammar-based class nor the output-based class was very helpful in improving his oral proficiency, and that at the same time, he was not satisfied with the class environment, either. Obviously, he experienced only a little amount of meaningful L2 interaction due to less contextualized teaching approaches and the reduced number of speaking turns.

**Personality.** The reason that he showed little satisfaction in the classroom in 2005 might have caused by his introvert personality and high face-protected predisposition.
The interview reflecting his participation in the previous class showed that he did not like to reveal himself in public. Indeed, MBTI (Myers-Brigg Types Indicator), which reports an individual’s eight personality preferences over four dichotomies (i.e., extraversion – introversion, sensing – intuition, thinking – feeling and judging – perceiving), indicates that Min has the tendency of ISTP (introversion, sensing, thinking and perceiving); ISTP type of person is a quiet observer who analyzes what makes things work, organizes facts using logical principles, and has high value efficiency.

Table 4 Min's Preference Clarity Index (PCI)

<table>
<thead>
<tr>
<th>Extraversion E</th>
<th>Introversion I</th>
<th>Sensing S</th>
<th>Intuition N</th>
<th>Thinking T</th>
<th>Feeling F</th>
<th>Judging J</th>
<th>Perceiving P</th>
</tr>
</thead>
<tbody>
<tr>
<td>clear</td>
<td>clear</td>
<td>clear</td>
<td>clear</td>
<td>clear</td>
<td>clear</td>
<td>clear</td>
<td>clear</td>
</tr>
</tbody>
</table>

PCI results: Introversion (28) Sensing (26) Thinking (18) Perceiving (20)

The PCI results such as Min’s in Table 4 indicate how clearly an individual chooses one preference over its opposite. A longer bar indicates that a person may have a strong preference on that criterion. Min’s PCI shows that his level of introversion is 28; Min seems to have very low preference on drawing energy from activities and people outside.

In sum, Min is a highly introvert learner who also has relatively high orientation in face protection; therefore, these facts support that the previous L2
classroom instruction, which requires self-exposure for interaction, might have not been an optimal environment for him to practice English oral production.

**Language Aptitude.** Aptitude is considered a very important factor determining a learner’s success in second language acquisition; according to Skehan (1989), “aptitude is consistently the best predictor of language learning success.” Regarding Min’s aptitude, the present study research administered the Modern Language Aptitude Test (MLAT), and all the subtests involving number learning, phonetic script, spelling clues, words in sentences and paired associates were tested.

**Table 5 Results of Min's MLAT score**

<table>
<thead>
<tr>
<th>Subtests</th>
<th>Aims of Measures</th>
<th>Scores in subtest</th>
<th>Score in complete test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Part 1: Number Learning</td>
<td>associative memory</td>
<td>22</td>
<td>43</td>
</tr>
<tr>
<td>Part 2: Phonetic Script</td>
<td>phonemic coding ability</td>
<td>11</td>
<td>30</td>
</tr>
<tr>
<td>Part 3: Spelling Clues</td>
<td>native language vocabulary</td>
<td>8</td>
<td>50</td>
</tr>
<tr>
<td></td>
<td>phonemic coding ability</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Part 4: Words in Sentences</td>
<td>grammatical sensitivity</td>
<td>29</td>
<td>45</td>
</tr>
<tr>
<td>Part 5: Paired Associates</td>
<td>associative memory</td>
<td>8</td>
<td>24</td>
</tr>
</tbody>
</table>

Min scored 78 in the complete set of the test, and the percentile of the score is ranked at 10% based on the norms for college freshmen. Min’s MLAT score indicates that he might not have high aptitude in language learning. Therefore, his poor performance in L2 production might have been caused by his low language aptitude.

**L2 Anxiety.** Krashen’s (1982) affective filter hypothesis in his Monitor theory claims that comprehensible input can only be effectively acquired when a learner has high motivation, self-confidence and low anxiety level. In the pre-interview, however,
Min expressed high anxiety level in his L2 speaking, which ultimately precludes him from reaching the ultimate goals in language learning:

**Min**: *I just aware my own mistakes every time whenever I make them, and I hate making mistakes in front of them. Every time when I aware my own errors, I feel myself very stupid. Maybe I learned too much grammar? …. Anyway I don’t feel good when I make mistakes, and it makes me nervous.* (Interview, Feb 27th, 2008)

Also, the questionnaire which measured his anxiety level in L2 found that Min feels uneasiness and confusion, and he seems to be easily embarrassed when he has to speak in English in the classroom or in front of other people. Regarding this, the level of Min’s L2 anxiety in this questionnaire recorded 5.21 out of 6 (here, closer to 6, higher L2 anxiety).

Min’s high anxiety seems to stem from the high degree of self-consciousness on L2 errors. He suspected that his grammar knowledge might have produced more anxiety by making him consciously compare what he knows in L2 and what he actually produces; he appears to focus on L2 structure more rather than its meaning. Regarding this, Hoffman (1986) stated that the anxiety distracts learners from focusing on “semantic content,” and makes them pay more attention to “physical feature of words (i.e., acoustic properties, order of presentation, phonetic similarities)” (Hoffman, 1986). As a result, his high anxiety level and high self-consciousness on errors might have hindered him from improving oral fluency.

**Motivation.** Similar to anxiety, motivation is another very influential factor determining the learners’ WTC in the classroom, and it ultimately affects their frequency of communication in the L2 classroom (MacIntyre & Charos, 1996). Gardner (1985) differentiated two types of motivation: integrative and instrumental.
According to him, integrative motivation comes from a desire to integrate with the TL community (i.e., positive disposition toward the L2 group, desire to interact in L2) while instrumental motivation comes from the rewards that might come from learning (i.e., getting a better job or higher salary). The types of motivation that each learner has are recognized as a significant individual variables of learning, and integrative motivation is considered to be a better predictor of SLA than instrumental (Gass & Selinker, 2001, p. 352).

In the present research, a questionnaire with six-point Likert scales was administered to measure the type of motivation; Min seems to have higher integrative motivation (3.6 out of 6) (here, closer to 6, higher integrative motivation) such as a desire to talk to foreigners, read books written in English, and understand pop music whereas he appears to have a lower degree of instrumental motivation (1.9 out of 6) such as getting a job and acquiring a higher profession, and being a more knowledgeable person. Also, in the pre-interview he was asked the reason that he wants to learn fluent English oral proficiency, and he answered:

**Min:** Yeah, I want to speak English fluently. The reason is because... well, I don’t know. I actually do not need to be fluent in English to get a job. I also have no wish to go to studying abroad in the future. I just want to speak English better than now. Actually, it makes me very envious if I see the people who have very high English ability. It’s just my feeling. I want to learn English nowadays, too. (Interview translated, Feb 28th, 2008)

As the results of the questionnaire and the interview show, Min has more integrative motivation than instrumental motivation; thus, his integral motivation can positively influence Min’s L2 learning, and it might compensate the negative effect from his low language aptitude.
Equipment and Materials

**English Interlocutor.** The researcher in the present study participated in this research as a PBI interlocutor. She is a graduate student studying Linguistics at Georgetown University.

**Telephones.** In this study, the English interlocutor’s cell phone and the participant’s telephone were used for PBI. At the fixed time on each day, the researcher called the participant, and they had 10 minutes of English interaction on the phone.

**Textbook for English conversation.** One book published in Korea was selected to help the learner’s English pronunciation and conversation. Each session covered four pages of the book, and only subsections such as new vocabulary and situational dialogue were used for the learner’s pronunciation and intonation practice. Whole sessions were targeted as meaningful interaction between the participant and the English interlocutor, and no grammar-based instruction was provided.

Procedures
Figure 2 illustrates one session of PBI. The identical procedure was repeated for 40 days except for pre/post-evaluation and exit questionnaires. In terms of the pre-evaluation, it was conducted a day before the 40 days of PBI treatment began. The post-evaluation and exit questionnaires were administered a day after PBI treatment.
**Pre/Post Evaluation Session.** To evaluate Min’s development of pragmatic use of English, an identical discourse completion test (DCT) was administered before the treatment and after the treatment. By using the same questions, the current research attempted to see Min’s different reaction to the same situation, which might indicate the influence of PBI. Also, Min was asked to debate on the same topic (the pros and cons of the movement in Korea to establish English as an official language) both before and after the treatment. By doing this, the study tried to evaluate any improvement in his discourse production affected by 40 days of PBI.

**Treatment Session (40 times of repetition).** The phone-based interaction was mediated only in English. It starts with four minutes of vocabulary and dialogue practice for the acquisition of lexicon and natural speaking. It was followed by 5–6 minutes of English conversation. Any relevant topic from the textbook (e.g., hobby, personality, invitation, food, dating, etc) was instantly chosen by the English speaker, and the participant had no planning time for English conversation. All English interaction between the English speaker and the learner over the phone was recorded on a digital recorder. After 10 minutes of PBI, the learner was asked to write one or two paragraphs in a journal every day reflecting his daily PBI experience.

**Questionnaires.** Questionnaires were provided before and after the treatment to observe any changes in the participant’s affective variables, which can influence his level of WTC. As previous research (MacIntyre, Clément, Dörnyei, & Noels, 1998) pointed out motivation and L2 anxiety as significant factors affecting learners’ WTC, current research measured whether Min experienced any changes in the types of motivation and degrees of L2 anxiety. Also, for more valid and clear comparison, the questionnaires previously used for Min’s biodata collection were used again.
Data Collection

*Pre/Post evaluation.* Each evaluation involved DCT and English debate. In terms of DCT, the test sheet was emailed to the participant, and the participant wrote the answer in English by using his computer. When he finished, the participant emailed the researcher the test sheet with his answers. Regarding the pre/post English debate, one debatable topic (i.e., the movement of English as an official language in Korea) was selected, and the English speaker and the participant had 10 minutes of English debate on the phone. Their debates were recorded by the researcher’s recorder, and were transcribed by the researcher.

*Phone-based interaction.* 10-minute English conversations were also recorded by the researcher’s recorder, and the researcher transcribed daily interactions. The transcripts involved not only the utterances produced by Min and the English speaker but also their pauses, hesitation and back channels, pitch and speed. Thus, the transcripts included both linguistic and extra-linguistic cues.

*Journals.* Journals were chosen to obtain information about learners’ internal processes. To capture the learner-internal processes and language learning experience, some previous SLA studies (Bailey & Ochsner, 1983; Cohen, 1997; Jones, 1994; Schmidt & Frota, 1986) utilized writing journals to reflect the learner’s own L2 learning. One major benefit of writing journals is that it can “yield insights into the learning process that may be inaccessible from the researcher’s perspective alone (Gass & Mackey, 2007, P. 28).”

In the current research, the learner wrote a journal reflecting his English development and effects of PBI. Any relevant topics related to the participant’s perception of his L2 development, difficulties in using L2, evaluation on the effectiveness of PBI, or any changes in his WTC, were selected by the researcher. To
gather more detailed and accurate data, the journals were written in the participant’s L1, which is Korean. The written journals were emailed to the researcher.

**Questionnaires.** The questionnaires that measure the learner’s motivation and anxiety were administered in L2. Similar to DCT, the questionnaires were emailed to the participant, the participant answered for each questionnaire, and then they were emailed back to the researcher.

**Analysis**

There were two separate types of data analysis for each research questions; 1) effectiveness of PBI for communicative competence development; and 2) influence of PBI on the learner’s WTC. To answer the first research question, the pre/post-evaluation and the transcripts of PBI were analyzed. Also, for the second research question, the current research analyzed Min’s journals and the exit questionnaires, which were particularly about his motivation and L2 anxiety.

**Effectiveness of PBI in developing communicative competence.** To investigate the benefits of PBI in developing communicative competence, the current study analyzed Min’s interactions on the phone during 40 days and compared the results of pre/post evaluations. To assess the improvement of his communicative competence, the present study adopted Canale’s (1983) theoretical framework of four components measuring L2 learners’ communicative competence: grammatical competence, sociolinguistic competence, discourse competence, and strategic competence. Table 6 below provides more detailed information about those four criteria. Also, it describes how each criterion was operationalized in this study.
Table 6 Data analysis for Min's development of communicative competence

<table>
<thead>
<tr>
<th>Communicative Competence</th>
<th>Definition</th>
<th>Data Analyzed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grammatical Competence</td>
<td>Mastery of vocabulary, rules of word formation, sentential grammar, linguistic semantics and pronunciation</td>
<td>Pronunciation</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Intonation</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Vocabulary use</td>
</tr>
<tr>
<td></td>
<td></td>
<td>English morphemes</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Any relevant journal</td>
</tr>
<tr>
<td>Sociolinguistic Competence</td>
<td>Mastery of language use in different sociolinguistic context depending on contextual factors</td>
<td>DTC</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Min’s pragmatic use of language</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Any relevant journal</td>
</tr>
<tr>
<td>Discourse Competence</td>
<td>Mastery of combine forms and meanings to achieve unified spoken or written texts</td>
<td>Min’s patterns of discourse</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Debate on social issue</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Any relevant journal</td>
</tr>
<tr>
<td>Strategic Competence</td>
<td>Mastery of verbal and nonverbal communication strategies to compensate for communication breakdowns</td>
<td>Min’s ability to recover from communication breakdown: paraphrase, requests for feedback, repetition, slower speech…</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Any relevant journal</td>
</tr>
</tbody>
</table>

According to Canale (1983), grammatical competence refers to “the degree to which the language user has mastered the linguistic code, including knowledge of vocabulary, rules of pronunciation and spelling, word formation, and sentence structure” (Hadley, 2001, p. 6). With the consideration of the telephone-based oral interaction, Min’s ability to produce accurate pronunciation and natural intonation and the development of his knowledge of English vocabulary were analyzed to measure his grammatical
competence. Also, regarding the mastery of Min’s proficiency in English grammar, his use of nine commonly studied grammatical morphemes (i.e., copula BE, progressive ING, auxiliary BE, past irregular, plural, 3\textsuperscript{rd} singular, article, possessive, past regular) were analyzed based on the data collected on the first and the last five days.

Sociolinguistic competence is defined as competence to use and understand a second language appropriately in various contexts to convey specific communicative functions (Hadley, 2001, p. 6). Thus, this study evaluates whether utterances are produced and understood appropriately in different sociolinguistic contexts, depending on contextual factors such as status of participants (i.e., student, teacher, neighbor, stranger etc), purpose of the interaction (i.e., request, refusal, compliment, greeting etc), and norms or conventions of interaction (i.e., in English, one does not normally ask strangers their age, marital status, or salary on first meeting) (Schmidt, 1983). This research observed Min’s appropriate use language in these different contexts. To measure sociolinguistic competence, the identical Discourse Completion Test (DCT) was administered before and after the treatment. Also, his pragmatic use of English was analyzed throughout the period.

Discourse competence, by definition, is “the ability to combine ideas to achieve cohesion in form and coherence in thought,” and it involves the learner’s use of “cohesive devices (i.e., conjunctions, adverbs, and transitional phrases) to achieve unity of thought and continuity in a text” (Hadley, 2001, p. 6). In this research, Min’s development in discourse analysis was analyzed by his use of pronouns and grammatical connectors in the English debates, which were conducted before and after the treatment. Also, any developmental patterns of English discourse that appeared during the treatment were also analyzed.

Lastly, in terms of strategic competence, which is “the use of verbal and nonverbal communication strategies to compensate for gaps in the language user’s
knowledge of the code or for breakdown in communication” (Hadley, 2001, p. 6), Min’s strategies to recover from communication breakdown was identified. For this, Min’s ability to paraphrase and his requests for repetition, feedback, clarification and slower speech were analyzed throughout the experimental period.

**PBI’s influence on the learner’s WTC.** To observe how PBI affects Min’s WTC, the present study analyzed Min’s daily journals. Among the journals collected, particular journals describing any affective factors, which can positively or negatively influence his WTC, were analyzed to investigate the relationship between PBI and WTC. Also, the analysis of exit questionnaires on his motivation and L2 anxiety was compared to that of Min’s biodata to draw out any significant affective impact of PBI treatment.

**Results**

The results of this study were organized in two different sections to answer the two research questions of the current study. In both sections, each research question is stated, and the results follow the research question.

To answer the first research question, whether PBI is effective for developing the communicative competence of a low WTC EFL learner, the results of the study are stated according to the four different criteria (i.e., grammatical, sociolinguistic, discourse, and strategic) composing communicative competence.

**Grammatical competence.** Regarding Min’s development of grammatical competence, his pronunciation, use of vocabulary, intonation, and accuracy of English morphemes were observed. In terms of Min’s pronunciation, he was more accurate in the articulation of English consonants rather than the vowels. While he clearly differentiated the sounds of /f/-/p/, /r/-/l/ and /b/-/v/, that most Korean EFL students
have difficulties with, he had some phonological errors in vowels; /o/ sounds are the most prominent examples.

Table 7 Comparison of Standard English pronunciation and Min's pronunciation

<table>
<thead>
<tr>
<th>Words (Day)</th>
<th>Standard English pronunciation</th>
<th>Min’s pronunciation</th>
</tr>
</thead>
<tbody>
<tr>
<td>sl&lt;e&gt;i (Day 4)</td>
<td>/sl&lt;e&gt;i / or /slêi/</td>
<td>/slêi/</td>
</tr>
<tr>
<td>once (Day 22)</td>
<td>/wʌns/</td>
<td>/uns/</td>
</tr>
<tr>
<td>allowed (Day 24)</td>
<td>/ælɔud/</td>
<td>/ælɔud/</td>
</tr>
<tr>
<td>plot (Day 31)</td>
<td>/plɔ/ or /plɔ/</td>
<td>/plɔt/</td>
</tr>
<tr>
<td>toast (Day 39)</td>
<td>/tɔust/</td>
<td>/tɔst/</td>
</tr>
</tbody>
</table>

As Table 7 indicates Min could not correct his non-target-like pronunciation of /o/ due to his lack of ability to generalize the vowel sound /o/. Also, he did not quite successfully differentiate front al and back vowels such as low-law, leave-live and slip-sleep. These phonological inabilities in vowel pronunciation seem to be driven by Korean L1 transfer, which has 1:1 matching between its orthography and sound and no significant differentiation between back and frontal vowels.

Min’s new vocabulary acquisition is another part indicating his less significant development. Regarding his own ability with the English lexicon, Min reported that he had some difficulties in the use of English vocabulary. His journal on Day 39 (Example 1) below indicates that he was not still able to find adequate words and expressions easily to describe his thoughts in English.

**Example 1 from his journal on Day 39**

The most difficult thing is that I cannot remember the words. I make many pauses, and that’s because I cannot match the word and the situation that I want to express. And I really don’t know which word I should use in a certain
situation. I cannot easily distinguish some words which have the similar meanings.

Also, throughout the period, he frequently showed overgeneralization in the use of English vocabulary; for instance, verb *attend*: *I attend the class, I attend gym, I attend church*; verb *promise*: *I promised my girlfriend, I have promise at 2 o’clock tomorrow*; noun phrase *a day before yesterday*: *I received a present a day before yesterday* (six months ago); prepositional phrase *at this point*: *I want to rest at this night*. Even though, his vocabulary errors were corrected by ES several times, he did not seem to acquire the ability to use English words accurately.

Additionally, Min did not use the new vocabulary that he learned daily, and his use of vocabulary was limited to those words that he learned previously from the prior English instruction.

**Example 2 from his journal on Day 2**

I think my English conversation will have no problem if I remember all of these words. The words here are so helpful. (Journal, Day 2)

**Example 3 from his journal on Day 40**

I think I just still use the English words that I already know. I don’t use new words. So I don’t think my English vocabulary ability improved that much.

The journal on Day 2 (Example 2) indicates that he recognized the usefulness of the new vocabulary which was introduced daily at the beginning of PBI. However, his journal on Day 40 (Example 3) describes that he seldom used the new vocabulary for his English conversation, and he admitted that there was very minimal progress in his use of English vocabulary. Consequently, he did not seem to acquire the useful vocabulary provided every day.
Although his pronunciation and vocabulary did not improve much, his intonation seems to have developed significantly. At the very beginning of the treatment period, Min did not express any emotions while he was in dialogue role-play.

**Example 4 from the role-play on Day 1**

1. Min: It’s been a while. How about yourself? (monotonous intonation)
2. English Speaker (ES): I can’t complain too much.
3. Min: Wendy, is that you? (monotonous intonation)
4. ES: What a small world! I was looking forward to seeing you.

Therefore, all sentences, whether statements (*It’s been a while.*, line 1) or interrogatives (*Wendy, is that you?*, line 3), Min read with monotonous intonation.

Also, due to his limited L2 speaking ability, he paused very frequently while he was producing an English sentence; thus, it caused Min’s unnatural intonation.

**Example 5 from the conversation on Day 1**

Min: I… miss always miss one person one person… first name is Park.
   Park…. first name is Park….Her family name…. is Park… She is … so … pre-… beautiful… and a perfect…. perfect…body li- body line.

**Example 6 from the conversation on Day 10**

Min: and please write…. the… correct correct my … please correct my …. pronunciation … and … write please write … please send … mail.

As Example 5 and Example 6 show, Min hardly produced English sentences without pauses, and the pauses seem to interfere with his natural speech. Therefore, in the early period of PBI, Min showed very unnatural oral English due to monotonous intonation.
and frequent pausing. As a result, when he was part of interaction, he merely transmitted English information without any emotional expression.

However, as time went by, he became more expressive in delivering his feeling through intonation. It was represented by his use of rising intonation when he read the interrogative sentences. Also, he reduced the number of pauses while producing sentences, and also gained in his ability to emphasize and prolong certain words that he wanted to stress. Consequently, it led him to acquire more natural intonation than before.

**Example 7 from the role-play on Day 28**

1. ES: What do they feel on your feet?
2. Min: These shoes hurt my feet. (falling intonation)
   Can you show me a larger size? (rising intonation)

**Example 8 from the conversation on Day 34**

1. Min: Oh! It’s it’s good. … I ALWAYS get up VERY earl early and and I
gave them a…
2. I gave them a breakfirst breakfast. And after tha:t I came to the university.
3. ES: How was your exam yesterday?
4. Min: TERRIBLE! It was TERrible.

In Example 7, Min successfully applied falling intonation for the statement sentence and rising intonation for the interrogative sentence. Also, in Example 8, Min spoke some words louder (i.e., ALWAYS get up VERY early, line 1; TERRIBLE! It was TERRible, line 3) and longer (And after that I came to the university., line 2) than others to express his current feeling. Compared to the previous days, these examples show Min’s significant development in speaking with natural intonation.
To measure Min’s accuracy on English morphemes, percentages of correct instances in all obligatory contexts were calculated.

Table 8 Accuracy order for nine grammatical morphemes in obligatory contexts

<table>
<thead>
<tr>
<th>Morphemes</th>
<th>Day 1 - Day 5</th>
<th>Day 36 - Day 40</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Copula BE</td>
<td>Acquired, simple sentence only</td>
<td>No change</td>
</tr>
<tr>
<td>2. Progressive ING</td>
<td>Acquired</td>
<td>No change</td>
</tr>
<tr>
<td>3. Auxiliary BE</td>
<td>Acquired</td>
<td>No change</td>
</tr>
<tr>
<td>4. Plural</td>
<td>Acquired (?)</td>
<td>No change</td>
</tr>
<tr>
<td>5. Possessive</td>
<td>Acquired (?)</td>
<td>No change</td>
</tr>
<tr>
<td>6. Past irregular</td>
<td>78%</td>
<td>81%</td>
</tr>
<tr>
<td>7. Past regular</td>
<td>75%</td>
<td>69%</td>
</tr>
<tr>
<td>8. 3rd singular</td>
<td>70%</td>
<td>71%</td>
</tr>
<tr>
<td>9. Article</td>
<td>47%</td>
<td>63%</td>
</tr>
</tbody>
</table>

Source: First and last five days conversation

As shown in Table 8, five grammatical morphemes, such as copula BE, progressive ING, auxiliary BE and plural, seemed to be used quite accurately whereas four morphemes, which are past irregular, past regular, 3rd singular and article, showed lower accuracy. Among those four less accurate morphemes, Min’s use of articles (i.e., *a, the*) was ranked lower than any other morphemes.

In terms of copula BE, Min showed very high accuracy when it is used in a simple sentence: *I am outside, she’s very talkative, what is my plan tomorrow?, etc.* However, his use of copula BE in conjoined or relative sentence, Min showed some inaccuracy: *waiter and waitress is very kind, *I like the man and woman who is reading books*, etc. In terms of Plural and Possessive, they were the most difficult
criteria to evaluate accuracy; for example, during the first five days and the last five days, there were not enough contexts which required the use of plural and possessive obligatorily. For instance, only two instances of obligatory plural contexts were found in the first five days, and four instances of plural contexts and three times of possessive contexts were found for the last five days. However, in those limited chances to use plural and possessive morphemes, Min showed 100% of accuracy for these items.

For the accuracy development in English morphemes, Min’s significant development is only limited to the use of articles; as Table 6 indicates, his accuracy for articles increased from 47% to 63%. Yet the data for the last five days showed that Min might have acquired a greater amount of chunk involving articles (a, an, the) rather than the conceptual meaning of those articles. This finding is contrastively shown by his relevant use of articles in such article obligatory contexts as *I’m very busy on Thursday, I had a meal, I can play the piano, the bible tells us, I have a dream, etc, and his irrelevant use of articles in the different contexts which identically require articles: *it’s a Friday, *Flute is not good for me, *we get together in restaurant, *we have a lunch, *I teach a mathematics, etc. The examples show that it was not possible for Min to accurately apply the rule of articles in different contexts. As a result, his development in English morpheme accuracy does not seem to be the result of his comprehension of its function.

In sum, during the 40 days of PBI, his grammatical competence was increased regarding natural intonation and the accuracy of English morphemes while other areas such as pronunciation, especially, vowels, and use of vocabulary were not developed much.

Sociolinguistic competence. For sociolinguistic competence, the results of pre-DCT and post-DCT were compared in order to measure the development of pragmatic use of English. Also, the current study observed his appropriate or
inappropriate use of English during the 40 days of PBI session, and analyzed any changes related with them.

Firstly, in terms of the two identical DCTs which were administered before the treatment and after the treatment, they show Min’s development of appropriate English use in different social contexts. For instance,

**Example 9 from pre- and post- DCTs**

Q5 (talking to a stranger): In your class, there is a girl whom you like so much, so you want to be a friend of her. Now, you got the chance. She is alone in the classroom! You are approaching her and saying…

(pre-DCT, before treatment)

**Min:** Hello!!

(post-DCT, after treatment)

**Min:** Hello!! Nice to meet you!! How’s this class??

**Example 10 from pre- and post- DCTs**

Q8 (refusal): You are on the bus to go to school now, and there is someone sitting right next to you. He is a stranger to you, but he says you look very familiar to him. But still, you don’t think you have ever seen him before. So, you say…

(pre-DCT, before treatment)

**Min:** Please think again.

(post-DCT, before treatment)

**Min:** I’m sorry but I can’t remember you. Could you think again?

Those answers in Example 9 and Example 10 in the post-DCT indicate that Min gained sociolinguistic ability to avoid face-threatening situation by using routinized
greeting (Hello!! Nice to meet you!! How’s this class??, Example 9) and more indirect utterances (I’m sorry but I can’t remember you. Could you think again?, Example 10) when he encountered unfamiliar to person. More specifically, in Example 9, Min seems to use a general way of greeting when he approaches a stranger. Also, in Example 10, Min used the statement of apology (i.e., I’m sorry) and questioning (i.e., Could you think again?), which are conventionally indirect forms of language use, as a sign of refusal.

Moreover, the following examples describe that Min’s reduced verbosity and increased conciseness for requests.

**Example 11 from pre- and post-DCTs**
Q 3 (request): You are a part-time worker in Mc Donald. There are five part timers including you and all of you are cleaning the floor. However, there is one part timer who are talking on the phone with his girl friend. He does not seem to finish the phone call very soon. What would you say to him?

(pre-DCT, before treatment)

Min: Hey, I waited for your attending the work. But you spent much time to talk with your friends. Do you find anyone who doesn’t attend work except you? Please stop the phone calling.

(post-DCT, after treatment)

Min: We need your hand. Sorry but could you finish the phone call?

**Example 12 from pre and post DCTs**
Q 5 (request): Mary borrowed 20 dollars from you last week, and she promised she will give you the money back on Tuesday. It’s Wednesday today, but she did not give you the money yet. Now, you should buy the textbook, so need to get that money back from her.
Min: Mary, I don’t want to say this. Please understand me to say this. I spent all money except 20 dollars which I borrowed you last week. Then I have to buy the textbook. So could you give me 20 dollars until tomorrow?

Min: Mary!! If I had enough money, I would not tell you, but nowadays I don’t have much money. I have to buy the textbook. Could you give me the money back?

The two questions above show that Min was quite verbose (i.e., *I waited for your attending the work. But you spent much time to talk with your friends. Do you find anyone who doesn’t attend work except you?*, Example 11) in delivering the message to keep politeness. Also, before the treatment, some examples of Korean transfer (i.e., *I don’t want to say this. Please understand me to say this.*, Example 12) in pragmatic language use can be found. However, as the treatment goes by, Min appeared to acquire the sociolinguistic skill to convey his messages more succinctly while still keeping politeness through apology and questioning (i.e., *Sorry but could you finish the phone call?*, Example 11; *Could you give me the money back?*, Example 12).

In this respect, the comparison of pre-DCT and post-DCT seems to indicate that Min acquired more appropriate use of English in different sociolinguistic contexts, such as the situations of request, refusal, and talking to a stranger; in those contexts, Min’s abilities of speaking succinctly, indirect and politely appeared to be learned.

However, not all sociolinguistic use of language is easy to acquire for learners. In Min’s case, learning how to answer negative questions seems to be very difficult. In other words, Min does not seem to grasp the American way of answering
yes or no for negative questions, and it seemed to be caused by his Korean background. In English, yes/no means positive-negative realization of the events or states of affairs at issue regardless of whether the question is asked positively or negatively. However, in Korean, yes/no refers to whether the content of the question is true or false.

For instance,

<table>
<thead>
<tr>
<th>Negative Question in English</th>
<th>Negative Question in Korean</th>
</tr>
</thead>
<tbody>
<tr>
<td>A: Did you have breakfast?</td>
<td>A: Did you have breakfast?</td>
</tr>
<tr>
<td>B: No. (I didn’t)</td>
<td>B: No. (I didn’t)</td>
</tr>
<tr>
<td>A: → Oh, you didn’t have breakfast?</td>
<td>A: → Oh, you didn’t have breakfast?</td>
</tr>
<tr>
<td>B: → No. (I didn’t)</td>
<td>B: → Yes. (your content is correct)</td>
</tr>
</tbody>
</table>

Thus, Min’s following errors can be considered as a sociolinguistic error that Min could not gain until the end of the treatment.

**Example 13 from the conversation on Day 7**

1 ES: Do you like cooking?
2 Min: Cooking?
3 ES: Yeah.
4 Min: I don’t like cooking.
5 ES: Oh, you don’t like cooking?
6 → Min: Yeah. (=I don’t like cooking)
7 → ES: Oh, you mean you like cooking.
8 → Min: No.

**Example 14 from the conversation on Day 35**

1 ES: I know. Do you have any break in this semester?
2 Min: No.
3 → ES: No? No break?

4 → Min: Yeah. (=I don’t have a break)

5 ES: No, No, I said you should say “No.”

6 Min: …. Oh! Right, right. No. No. S**t!

On Day 7 (Example 13), Min answered positively (Yeah., line 6) to express that he does not enjoy cooking because he regarded the content of the ES’ question was correct (Oh, you don’t like cooking?, line 5). However, in English, he should have said “No” whenever the realization of states is negative. Similarly, on Day 35 (Example 14), Min also answered positively (Yeah., line 4) to agree with the content of the ES’ previous question (No? No break?, line 14). However, in English, “Yeah,” here, means “yes, there is a break.”

On Day 35 shown in Example 14, Min produced the answer in non-target-like way (Yeah., line 4 in Example 14), and ES corrected him explicitly (No, No, I said you should say “No.”, line 5 in Example 14) to provide the appropriate answer in target-like way; ES’ utterance “I said you should say” indicates that this type of correction also took place previously, and indeed, Min was also corrected explicitly on Days 13, 22, 24, 29 and 34 due to his inappropriate answers for English negative questions. However, it seems that his yes-no production for negative questions is quite difficult as seen from his periodic errors. Therefore, he showed some frustration on his limited sociolinguistic ability (…. Oh! Right, right. No. No. S**t!, line 6 in Example 14) which does not change flexibly.

It is interesting that similar errors were found on Day 7 and Day 35, and it means that Min could not have acquired the appropriate way to answer Yes-No for negative questions. Regarding this, it seems to be caused by his Korean cultural background. Therefore, sociolinguistic competence of L2 can be more difficult to
obtain when L1 and L2 involve totally opposite linguistic rules for a certain pattern of discourse.

**Discourse competence.** Regarding discourse competence, the present study looked at Min’s changes in discursive patterns and sentence structures shown during the 40 days of PBI session. Especially, his discursive patterns (i.e., active, passive, informative, persistent, etc.) were analyzed to see if his ability in fluent speaking and coherent speaking were improved. Also, by comparing his L2 speaking shown in pre/post English debate any developmental pattern in discourse ability was supported.

Compared to other competences, discourse competence seems to be the part that Min demonstrated the greatest development. This outcome is obviously shown by his more fluent conversation skill compared to that of his early period. Consider the following examples below.

**Example 15 from the conversation on Day 2**

1. ES: Oh, you are tutoring a student.
3. ES: Oh, I see. That’s really good. You can make money?
4. Min: Yeah, a little.
5. ES: Alright, let’s look at the page 21.
7. → ES: What? What is that? Sorry?
8. → Min: Ah… no, no, no.

**Example 16 from the conversation on Day 7**

1. ES: So, can you describe your personality in English?
3. ES: Oh, you are honest?
As the examples from the early period of interaction shows, Min easily lost his confidence in speaking English when ES did not understand his non-target-like English. For instance, in Example 15, when ES asked him to clarify his message (*What? What is that? Sorry?, line 7*), Min had a tendency to avoid situations which might expose his weakest English (*Ah... no, no, no..., line 8*). Additionally, he tended to make very short answer for ES’ questions, stop talking when his utterances were overlapped with the English speaker’s, and wait until the other speaker brought up the topics (*......, line 6 in Example 16*).

Min’s passive participation is also supported by his pattern of discourse in pre-English debate. Regarding the topic of “The Movement of English as an Official Language in Korea,” he produced more English sentences such as “*...I don’t remember the word.*” “*It’s not easy.,” “I think it’s not easy to speak in English,” or “*I think it’s not easy to speak in Korean, too,“ and he did not produce a strong argument indicating his stance toward that social issue. Therefore, during the English debate, there were more pauses than L2 utterances in his turn, and this might represent his very passive participation in the conversation.

Also, during the early period, his limited capability to understand English caused frequent miscommunication. Two examples are given below.

**Example 17 from the conversation on Day 1**

1 Min: So I said... I like her.
2 ES: Alright. we are going to finish our talk here. Was it good? Was it embarrassing? Did you enjoy it?

3 Min: I was very embarrassed that time at that time.

Example 18 from the conversation on Day 8

1 ES: Like. What does Daegu (a city in Korea) like?

2 Min: What does Daegu like?

3 ES: Yeah.

4 Min: Yeah, I like all things in Daegu.

In example 17, ES started to sum up the conversation and asked Min if today’s English conversation was embarrassing (Alright. we are going to finish our talk here. Was it good? Was it embarrassing? Did you enjoy it?, line 2) right after he finished his story how he proposed to his girlfriend (So I said... I like her., line 1). However, Min does not seem to understand the English speaker’s use of the pronoun it in was it good?; in other words, Min could not comprehend what the pronoun it indicates, exactly, in that sentence. Thus, he brought up the topic that he discussed previously, and answered he was very embarrassed when he proposed to his girlfriend. Consequently, Min’s misunderstanding of pronoun it introduced incoherent communication between the two speakers.

Similarly, in Example 18, he could not distinguish two different uses of like (i.e., Do you like (v.) it? or Is it like (prep.) that?) in the conversation. As a result, regarding the ES’ question asking his impression about the city, Daegu (What does Daegu like?, line 1), he answered that he liked the city (Yeah, I like all things in Daegu, line 4). Thus, it shows another miscommunication between those two interlocutors, and their conversation also ran incoherent.
As a result, in the early period of interaction, Min was passive in participating in English conversation. Also, due to his limited ability of listening and speaking, communication breakdown occurred from time to time. Therefore, in those days, less coherent conversation took place frequently.

However, as PBI treatment goes on, his discourse ability appeared to be more fluent and coherent than the previous time. At the middle of the period, he acquired the skill to use relative clauses with less effort, and it seemed to increase his ability to unify English sentences cohesively.

**Example 19 from the conversation on Day 12**

1. ES: How’s the weather in Korea now?
2. Min: It’s warm…. Nowadays it’s warm in Korea.
3. ES: Oh, that’s good.
4. Min: Yeah. So I can… wear… the… jacket, Jacket and… girlfriend… she bought… bought a jacket.
5. ES: She buys a jacket? Does she buy a jacket?
6. Min: She bought a jacket for me.
7. ES: Oh, you mean you are wearing the jacket that she bought for you!

Before he acquired the skill to use relative sentences effortlessly, as shown in Example 19, Min could not produce modifying clauses as a unified idea: for instance, he could not deliver the idea of “I can wear the jacket that my girlfriend bought for me” (*Yeah. So I can… wear… the… jacket, Jacket and… girlfriend… she bought… bought a jacket.*, line 4) cohesively. Due to less cohesion between such clauses as “wearing a jacket” and “a jacket from his girlfriend,” ES experienced communication breakdown (*She buys a jacket? Does she buy a jacket?*, line 5).
However, some days later, Min became able to produce English sentences more fluently and cohesively by being more proficient in the use of relative clauses. Therefore, modifying sentences were produced more coherently as shown in the following two examples.

**Example 20 from the conversation on Day 19**

1. ES: Where are you going now?
2. Min: I’m going to … I’m going to the station …. which I … take the bus.

**Example 21 from the conversation on Day 26**

1. ES: So did you decide which laptop you are going to buy?
2. Min: No, I I I found a laptop which I want to buy but it was a little expensive so I will call and I want discount the price.

In Example 20, he successfully modified “the station” with the clause starting with “which…” (I’m going to … I’m going to the station …. which I … take the bus., line 2). Also, in Example 21, Min easily explained the laptop that he wants to buy is not cheap (I I I found a laptop which I want to buy but it was a little expensive…, line 2). In this sense, his ability to produce English sentences more coherently was improved compared to the early days of PBI.

Moreover, Min’s passive participation changed into active participation while he was gaining more competence in performing English discourse. In contrast to his previous behavior of avoidance and short answering, his active participation was shown by his increased willingness to communicate; his speaking turn was not discouraged by the other speaker’s overlapping utterances; he voluntarily shared his interesting episodes that he experienced at school; and he made jokes and teased ES on the phone.
Example 22 from the conversation on Day 22

1 Min: I yeah I will go to library and I will I have I have have to finish my homework … I have two homeworks so I have to finish. And [after-
2 ES: [Is your-
3 Min: And after [that-
4 ES: [Is your stu-
5 Min: Wait a minute! And after that I will study general che chemistry and organic chemistry and biology.

In Example 22, Min’s speaking in line 1 (after-) and line 3 (that-) were overlapped by ES’ utterances in line 2 (Is your-) and line 4 (Is your stu-); thus, Min’s utterances were interrupted each time. However, his turns of speaking were not stopped by ES’ interruption, and he even asked her to wait for a while (Wait a minute!, line 5) until he finishes what he wanted to say.

These tendencies of Min’s L2 speaking, such as active talking and coherent speaking, were also shown in his post-English debate. For instance, he stated his own idea more often by uttering “In my opinion,” “I think,” “I don’t think,” and “Actually.” Also, by using relative sentences (i.e., people who need to study English, person who don’t need to speak English, Korea have our own language which is Korean, etc), he could deliver more coherent and unified idea in English.

Finally, in the latter period of PBI, Min became a better story teller as shown in the following example.

Example 23 from the conversation on Day 31

1 Min: I have funny story.
2 ES: Sorry?
3 Min: I attended the English class.
ES: Okay, how’s it going?

Min: Ah he the teacher told me that if I grow my hair like like….like… if I grow my hair I like Jama Jamaican in three months he ga he would give me he would give me A A plus. A plus. I have a I have a pencil case has has the doll The doll the hair has Jamaican hair. He saw he saw that and told me. I know I know it’s impossible I don’t like that kind of hair style. So I told him that if I …. Thought playboy I can I can I can grow up I could grow up grow up my hair very soon.

The Example 23 shows that Min voluntarily started to talk about a funny episode right after the role-play session. His narrative in line 5 indicates that his speech involves not only the funny situation, such as reading porn magazines and growing hair quickly (*I told him that if I .... Thought playboy I can I can I can grow up I could grow up grow up my hair very soon.*), line 10-11), but also its background regarding his pencil case with a Jamaican boy (*I have a pencil case has has the doll The doll the hair has Jamaican hair. He saw he saw that and told me.*, line 7-9).

**Example 24 from the conversation on Day 28**

ES: Oh, I didn’t know. I thought you were still talking about your muscle!

Min: I finished that subject.

ES: Oh, I totally understood now.

Min: Please improve your hearing skill (laughing)

ES: What? (laughing) Improve what?

Min: hearing skill.

ES: My hearing skill? My listening skill? (laughing)

Min: Yeah, my hearing is very good. (laughing)
In Example 24, Min teased ES when she did not catch up the topic that Min was talking about (Oh, I didn’t know. I thought you were still talking about your muscle!, line 1). At that time, Min teased her about her wrong understanding (Please improve your hearing skill. (laughing), line 4; Yeah, my hearing is very good. (laughing), line 8). Compared to the early conversation, his laughing and teasing in this excerpt seems to indicate that he now enjoys talking in English.

In sum, during the 40 days of interaction, Min’s discourse competence significantly improved in terms of speaking actively, persistently and cohesively; each fact can be supported by Min’s tendency to deliver more information at once and his use of some coherent markers (i.e., relative sentences, logical connectors) in his discourse. Also, he attempted to become a better story teller who can deliver not only the core of the story but also the background of the story. Compared to the early days of his English discursive behavior (i.e., quiet, passive, easily interrupted), his development in discourse competence (i.e., more talkative, active, keeping his own flow of conversation) can be regarded as considerable improvement.

**Strategic competence.** In terms of strategic competence, Min’s use of communication strategies to recover from a communication breakdown, such as a request for clarification or corrective feedback, were analyzed throughout the PBI period. Because the interaction on the phone provides no gestures or facial expressions to guess the other speaker’s utterances, Min’s strategic competence was only observed on the verbal level, and communication strategies at the nonverbal level were not analyzed. Similar to discourse competence, any significant changes of Min’s pattern of communication strategies were investigated in this section.

In response to clarification requests, Min seemed to develop his own strategies in later days. Due to Min’s insufficient English communication ability in the early period, communication breakdowns frequently took place. Therefore, Min often
requested ES to repeat what she said. In early days, he used *pardon? Could you say that again? I can’t hear your voice very well, I don’t understand, Could you say that again?*, and *Slowly, please*, and it made ES repeat or paraphrase the whole sentences that she uttered.

However, as time went by, he attained his own strategy to ask specific parts that he did not comprehend.

**Example 25 from the conversation on Day 28**

1 ES: How do you get the news?
2 Min: How do I…
3 ES: getting the news, get the news.
4 Min: How do I getting…
5 ES: the news.
6 Min: How did I get the news?
7 ES: Yeah.

**Example 26 from the conversation on Day 30**

1 Min: I’m fine. I’m a little tired.
2 ES: Oh, why are you a little tired?
3 Min: Why I am…
4 ES: Tired.
5 Min: Why I am a little tired?

**Example 27 from the conversation on Day 39**

1 ES: Okay, so what do you do when you meet your friends?
2 Min: What do I do what?
3 ES: What do you do if you meet your friends?
4 Min: Ah, I always have a meal with my friends.

In Examples 25, 26 and 27, Min repeated what he comprehended to indicate to ES the specific part that he did not comprehend (How do I..., line 2; How do I getting..., line 4 in Example 25; Why I am..., line 3 in Example 26; What do I do what?, line 2 in Example 27). Therefore, ES could know the particular part that Min was unable to understand, and repeated only that part (getting the news, get the news., line 3; the news., line 5 in Example 25; Tired., line 4 in Example 26). As a result, Min seemed to gain the ability to compensate for the communication breakdown more quickly and effectively.

Also, Min developed the skill of asking corrective feedback for his non-target-like English use. In the early period of PBI, Min tended to avoid speaking some English expressions that he was not sure of. For instance, his utterances were usually finished with “I don’t know (Day 3)” “I can’t speak in English (Day 12)” or “I will tell you later (Day 18).” However, as time went by, he was likely to look for ES’ corrective feedback when he encountered uncertain moments for the use of English words and expressions. Therefore, at that time, he indirectly requested for corrective feedback with his own strategy: I go to bed early in night? on night? (Day 24); I like my T-shirts, T-shirts, or shirts? Or T? (Day 32) etc.

In sum, regarding Min’s strategic competence, he seems to be able to create his own style of conversational strategies to recover from communication breakdown and his non-target-like use of English. Even though both strategies were not created in 100% accurate grammatical English sentences, it reflects that he acquired his own strategies to overcome those uncertain moments. As a result, his strategic competence seems to be improved to some extent.
So far, this section has described Min’s development of communicative competence in terms of four different components: grammatical, sociolinguistic, discourse, and strategic competence. The following Table illustrates the summary of Min’s communicative competence development.

<table>
<thead>
<tr>
<th>Communicative Competence</th>
<th>Significant development</th>
<th>Insignificant development</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Grammatical Competence</strong></td>
<td>Intonation:</td>
<td>Pronunciation:</td>
</tr>
<tr>
<td></td>
<td>- unnatural → natural</td>
<td></td>
</tr>
<tr>
<td>Use of English morphemes:</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>- accuracy increased</td>
<td></td>
</tr>
<tr>
<td></td>
<td>- especially, articles (a, the)</td>
<td></td>
</tr>
<tr>
<td><strong>Sociolinguistic Competence</strong></td>
<td>Pragmatic use of L2</td>
<td>Pragmatic use of L2</td>
</tr>
<tr>
<td></td>
<td>- increased politeness</td>
<td>- Answering appropriately</td>
</tr>
<tr>
<td></td>
<td>- decreased verbosity</td>
<td>(Yes/No) for negative</td>
</tr>
<tr>
<td></td>
<td>- increased conciseness</td>
<td></td>
</tr>
<tr>
<td><strong>Discourse Competence</strong></td>
<td>Discursive pattern</td>
<td></td>
</tr>
<tr>
<td></td>
<td>- involves more information</td>
<td></td>
</tr>
<tr>
<td></td>
<td>- increased cohesiveness</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(i.e., use of relative clauses)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>- increased persistence in L2 speaking</td>
<td></td>
</tr>
<tr>
<td><strong>Strategic Competence</strong></td>
<td>Recovery of communication</td>
<td></td>
</tr>
<tr>
<td></td>
<td>break down</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(i.e., requests for clear)</td>
<td></td>
</tr>
</tbody>
</table>
comprehension, requests for
corrective feedback)
- created the learner’s own strategies
to recover communication breakdown
more effectively

For the second research question, if PBI affects the EFL learner’s WTC, this section analyzes Min’s daily journals to find out any changes in his affective factors. For the two months of PBI session, he reported 10 times of changes in his L2 motivation and anxiety.

Regarding his motivation, Min reported that PBI stimulates him more to become a proficient English speaker compared to the previous L2 classroom instruction.

**Example 28 from his journal on Day 3**
In the classroom, when I had English interaction with my classmates, I just believed that I am kind of a better English speaker than those students. So I thought I will be fine if I don’t speak that much. But now, the English speaker that I am talking with is extremely good at English, and it makes me do my best to become a good English speaker like her.

This segment of Min’s journal indicates that he gains more motivation when he interacts with more proficient ES than himself. According to Min, he compares his English ability to the interlocutor’s, and he gets more motivation by noticing the obvious gap between his interlanguage and the other speaker’s TL. In this sense, the classroom setting which provides more time to interact with classmates did not give him a lot of motivation because those students’ ability was not significantly better.
However, Min found a contrast in proficiency in the PBI context, and it motivated him to make more effort in learning English communicative skill.

Also, on the same day, he wrote “I really want to have more confidence in English speaking,” “I really want to be a very fluent English speaker,” “I want good pronunciation!” and “at least, I want to be fluent in English free talking.” Moreover, on Day 16, he noted that

**Example 29 from his journal on Day 16**
I like speaking in English because learning English improves me myself. And I think this PBI is a very rare but good opportunity to practice English everyday.

Thus, it can be represented that his intrinsic and integrative motivation was reinforced more than his extrinsic and instrumental motivation.

In terms of L2 anxiety, Min seems to have tended to decrease his L2 anxiety as time goes by. Before PBI began, his self-report described that he has very high L2 anxiety shown by high self-consciousness on his errors, face-protected predisposition, and worries about his insufficient English competence. Therefore, Min talked in English in front of others very seldom. However, on Day 14 (Example 30), he wrote that:

**Example 30 from his journal on Day 14**
I have only 10 minutes of session everyday but for those 10 minutes I think and talk in English as much as I did in one-hour English classroom. In the classroom, I was passive, but now I just try to talk even though I make some mistakes, because I can realize what is wrong with my English and I can
correct it. There is no one else looking at me, so I can make mistakes without any fear. And I can also talk as much as I want.

His journal on Day 14 shows that Min feels more comfortable at making errors and producing English because PBI does not involve any other people looking at him. Therefore, Min became more active in speaking in English, and he was not discouraged by the errors that he made. Rather, Min might have considered those moments as a good chance to correct himself. Moreover, on Day 25 (Example 31):

**Example 31 from his journal on Day 25**

It is very obvious that I have reduced the fear of speaking English. So I might be less annoyed if I talk to a foreigner now. And, actually, I might want to take an English conversation course in the next semester, which has a native English instructor. I got excited!

This journal describes a significant change of his L2 anxiety. In the journal above, Min stated that talking with a native speaker of English is not very annoying for him, and he does not have much fear in talking to an English speaker. Regarding this, it illustrates that Min’s L2 anxiety is definitely reduced compared to the previous days.

Additionally, in the journal that he wrote on Day 34 (Example 32), he reported that:

**Example 32 from his journal on Day 34**

I just feel free to talk in English these days with you. There is only you and no one else around me. I mean, no body is watching me because it’s one-on-one conversation with you. I used to hate talking in front of a lot of people watching me. Face protection. Actually that was the reason.
In this journal, it is very obvious that one-on-one conversation in Min’s own private place was very helpful for him to free himself from face-protecting predisposition. Also, this fact made Min more comfortable at producing English, and reduced the worries about his lower L2 competence.

Because of his decreased L2 anxiety, Min seems to perceive his L2 competence positively. The following journals support the fact that Min gained self-confidence in L2 use and L2 competence.

**Example 33 from his journal on Day 19**
I just speak out anything that I want to say whatever it is right or wrong. I think I have some confidence now.

**Example 34 from his journal on Day 25**
Now I use relative clauses very often. And it is surprising that I can use relative clauses! I was, seriously, surprised when I realized that I am using relative clauses. I thought, “oh, look! I am modifying the phrase with the relative sentence!” It was amazing. I think my English speaking ability has increased a lot.

**Example 35 from his journal on Day 34**
I think I do not make a lot of mistakes nowadays. I feel confidence because I can see myself making most sentences that I want to say. Having confidence is really good because the feeling itself helps me improve myself, and by doing so, I can have more and more confidence.

In those journals in Examples 33, 34 and 35, he recognizes himself gaining confidence in L2 speaking, and his self-perception on his L2 competence is getting very positive. Also, his report about his more active participation with less consciousness on his L2 errors might represent Min’s enhanced WTC.
The outcome of exit questionnaires on Min’s motivation and L2 anxiety provides more convincing results to support his changes in affective variables. In terms of motivation, he seems to gain more integrative motivation from PBI; before treatment, his tendency to integrative motivation was quantified as 3.6 out of 6 (here, closer to 6, more clearly integrative) whereas instrumental motivation was 1.9 out of 6 (here, also, closer to 6, more obviously instrumental). However, after the 40 days of treatment, his integrative motivation was quantified as 4.8 while instrumental motivation was 2.1. Both types of motivation were increased, yet, the results of questionnaire illustrate that Min’s integrative motivation increased more significantly than instrumental motivation.

With respect to L2 anxiety, Min showed more considerable outcome. Before treatment, Min had very high L2 anxiety, and the level of L2 anxiety was calculated as 5.2 out of 6 (here, closer to 6, higher L2 anxiety). However, after 40 days of PBI, Min’s L2 anxiety was calculated as 3.3 out of 6; the results show that his L2 anxiety was significantly reduced in comparison with the beginning of the treatment. These results strongly support the fact that Min reported in his journals: more motivation and less L2 anxiety.

In sum, these results indicate that 40 days of PBI led him to increase motivation, decrease L2 anxiety, and enhance self-confidence in L2 speaking. Those changes in affective variables might raise Min’s level of WTC, and it could be represented by his more active participation in English conversation.

Discussion
In this section, the effectiveness of PBI is discussed based on Min’s improvement in communicative competence and WTC. More specifically, in response to Min’s development in communicative competence and his WTC, the current section
discusses PBI’s potential benefits that can contribute to their development. Also, by discussing the results of insignificant development, the present section brings some suggestions to increase its benefits. Table 10 below summarizes Min’s development of communicative competence, and PBI’s potential contributors.

Table 10 Potential contributors to Min's development

<table>
<thead>
<tr>
<th>Potential contributors</th>
<th>Significant development</th>
<th>Insignificant development</th>
</tr>
</thead>
<tbody>
<tr>
<td>Interaction in authentic L2 contexts</td>
<td>Grammatical Competence: - Intonation</td>
<td>Grammatical Competence: - Vocabulary</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Sociolinguistic Competence:</td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Politeness ↑,</td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Verbosity ↓,</td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Conciseness ↓</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Strategic Competence:</td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Conversation strategies ↑</td>
<td></td>
</tr>
<tr>
<td>Output in intensive one-on-one conversation</td>
<td>Grammatical Competence: - Use of English morphemes</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Discourse Competence:</td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Information ↑,</td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Cohesiveness ↑,</td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Persistence ↑</td>
<td></td>
</tr>
<tr>
<td>L1 transfer</td>
<td>Grammatical Competence:</td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Pronunciation</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Sociolinguistic Competence:</td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Yes/No for negative questions</td>
<td></td>
</tr>
</tbody>
</table>
As shown in Table 10, PBI seems to facilitate learners’ development of communicative competence by providing authentic contexts for L2 interaction and intensive one-on-one communication enforcing learners’ output.

**Interaction in an authentic L2 context.** It seems that PBI’s authentic L2 contexts contribute to the learner’s development of natural intonation and pragmatic knowledge of how to use L2 appropriately (i.e., less verbosely, more politely, and more concisely). Also, Min developed effective ways to recover from a communication breakdown by individually creating some conversation strategies (i.e., repeating only the part that he comprehended). Regarding some abilities in which Min did not show significant improvement, such as vocabulary, PBI’s provision of new vocabulary in decontextualized manner might have been less effective for the learner to acquire the new words.

In terms of natural intonation, PBI’s provision of genuine use of L2 might have shown him English-specific pitch, intonation and voice to deliver certain messages in particular contexts. According to Brown (2001, p. 269) L2 interaction involves multiplicity of sounds, words, phrases, and discourse forms that characterize any language. Regarding this, the learner’s development of natural intonation might have derived from genuine L2 interaction and authentic contexts, which provides English-specific use of paralinguistic factors.

For pragmatic L2 knowledge, which constitutes to the learners’ sociolinguistic competence, Bardovi-Harling et al. (1991) stated that L2 learners “must be exposed to language samples which observe social, cultural, and discourse conventions, or in other words, which are pragmatically appropriate” (1991, p. 4) Also, Williams (1988) observed that L2 activities “focusing on using language in ongoing discourse, in a particular context, for a particular purpose” can develop learners’ understanding of L2 speech act and their appropriate language use in a particular context (1988, p. 46). In
this sense, provision of genuine L2 context can develop learners’ pragmatic knowledge in L2 use by involving them in certain speech acts (e.g., compliments, apologies, refusals, regret, gratitude).

In this respect, PBI appears to involve various speech acts so that learners can acquire socially and culturally adequate rules of TL. Because PBI provides natural L2 interaction, instead of explicit L2 instruction, to reach the communicative goal of L2 learning, PBI necessarily includes various authentic L2 contexts which take place spontaneously. For instance, in the current research, such natural contexts as compliments (e.g., when a learner performed better English speaking ability than before), apologies (e.g., when a speaker could not understand the other speaker’s utterances, when a speaker could not talk on the phone longer due to personal reasons) and refusals (e.g., when a speaker turned down the other speaker’s suggestion) frequently occurred, and the learner could learn how to interact appropriately in L2 in a direct manner.

The effectiveness of interaction in genuine L2 context also seems to play a crucial role for Min’s development of conversation strategies. According to Selinker (1972), conversation strategies appear in spontaneous L2 speech due to L2 learners’ limited knowledge of the target language. Also, effective conversation strategies take place when students are in negotiation of meaning that L2 interaction involves (Hadley, 2001, p. 6). In this sense, the L2 learner’s development in strategic competence in this research indicates that PBI obviously involves authentic L2 interaction, and the interaction includes some conversation activities which spontaneously take place. Therefore, PBI can be also effective for learners to develop their conversational strategies through the interaction in authentic L2 contexts.

The effectiveness of authentic L2 interaction might be supported by Min’s less significant improvement in new vocabulary acquisition. In this current research, new
English vocabulary was provided in decontextualized manner rather than contextualized manner. For instance, 1–1.5 minutes of vocabulary session at the beginning of the conversation introduced meaning and form of the words in audiolingual manner, for instance, listen and repeat. Therefore, PBI in this research did not provide any specific information about how each word can be used in what contexts. Therefore, the learner might have learned how to pronounce them accurately rather than how to use them appropriately.

Related to this, Prince (1996) found that L2 students in decontextualized vocabulary learning were outperformed the students in contextualized vocabulary learning in terms of quantity, but the decontextualized group was weaker to transfer vocabulary knowledge in the L2 contexts. Also, Sternberg (1987), who discussed the development of vocabulary development based on three basic frameworks of vocabulary learning, concluded that even though learning vocabulary in context is not a rapid process, vocabulary leaning from context can stimulate L2 learners’ transfer mechanism, enabling learners to use L2 words in appropriate contexts (p. 104).

In this sense, it might be possible for the learner to have more development in vocabulary acquisition if contextualized vocabulary learning session were provided in the current PBI. Also, Min reported in his journal that finding appropriate English words or expressions to describe certain events is still difficult for him. Regarding this, it seems that Min did not learn how to use new L2 vocabulary in certain contexts. As a result, the importance of authentic L2 context and interaction in that context can be supported by Min’s less significant improvement in decontextualized vocabulary acquisition.

In summary, PBI can provide authentic L2 contexts, which can be a better environment for learners to acquire not only L2 pragmatic knowledge and conversation strategies but also new vocabulary. As Brown (2001, p. 116), stated “EFL students do
not have ready-made contexts for communication,” there can be no provision of genuine L2 contexts in EFL society, especially beyond the L2 classroom. Regarding this, PBI can provide the communicative situations in which EFL learners can have authentic L2 input and natural L2 interaction. Thus, the opportunities to interact in genuine L2 context can be provided by PBI, which can lead the students to enhance the ability to use L2 appropriately.

**Output in intensive one-on-one conversation.** In this research, Min showed some increase in grammatical competence in terms of the accurate use of English morphemes. Also, After 40 days of PBI session, Min obviously showed the development of discourse competence; his discourse after the treatment period became more informative, cohesive and persistent. Regarding the improvement in those two criteria, the intensive one-on-one conversation system provided by PBI seems to play a significant role.

Firstly, Min’s development of grammatical competence, represented by morpheme accuracy, might have been caused by his increased amount of L2 chunks. According to Ellis (2005, p. 333), practices lead learners to bind adjacent perceptual and behavioral sequences more strongly when they process them. Consequently, “the constructed message itself will get chunked together,” and the chunked utterances become “an entrenched formula that can be looked up” when similar situation encounters. Thus, chunking itself can be a significant learning mechanism which contributes to language automaticity, and it can be facilitated by continual practice.

Related with this, Min’s development of morpheme accuracy was particularly shown by the use of articles (i.e., *a(n), the*), and the accuracy of articles was only increased in limited numbers of sentence structures that he used frequently: for instance, “have a good day,” “I have a meal,” “I have a dream.” However, when he encountered a similar sentence structure, he overgeneralized the rule: *I had a lunch.*
Also, in some contexts, which require articles obligatorily, he did not produce the morphemes: *We went to ___ restaurant*. In this sense, he did not seem to understand the grammatical function of certain morphemes, but he might have produced them based on the similarities of sentence structures: Subj **HAVE A** Obj (something to eat)” Therefore, PBI’s increased opportunities of output seem to play a role of intensive practice, and it might have reinforced his automatized use of English morphemes, which became chunked together due to his frequent use of certain phrases or sentences.

Regarding discourse competence, Min also showed significant development in terms of his increased ability to deliver ideas with more information, cohesion and persistence; it might be facilitated by his skill acquisition, which enables him to produce L2 effortlessly. In other words, by producing English sentences with less attentional effort, he could produce more L2 sentences at once with better organization.

According to Schmidt (1992), learners’ repeated practice can facilitate the development of L2 oral ability, which can ultimately draw automatized processes of language production. Related with this, Swain (Swain, 1995; Swain & Lapkin, 1995) discussed Output Hypothesis, and it explains that learners’ comprehensible output can play a role of practices which can contribute to language fluency and automaticity. Similar to Swain, de Bot (1996) also noted that output plays a direct role in enhancing fluency by turning declarative knowledge into procedural knowledge. Therefore, previous studies suggest that output itself can be a significant L2 learning mechanism which facilitates automatic language process.

Therefore, Min’s active production of output also seems to be a significant reason he improved his discourse competence during 40 days of PBI. For instance, to raise cohesiveness of his discourse, he frequently used relative clauses such as “I am going to station… **which… I… take bus**. (Day 19)” “Jeju Island have many place… **which is… beautiful**. (Day 21)” “I found a laptop…. **which I want**…. to buy. (Day
26)” and “I met JS who is my friend. (Day 37)” As the examples show, his number of pauses that he made while he was producing coherent relative clauses was reduced as time went by. His ability to produce English automatically, effortlessly and with less additional attention might led him to have more L2 utterances at once, which might contribute to better informativeness, cohesiveness, and persistence.

In this respect, PBI seems to provide more opportunities for learners to produce output, which can contribute to the learners’ L2 automaticity. Unlike the L2 classroom setting, which provides limited time to talk to a native instructor, 10 minutes of PBI involves intensive one-on-one conversation environment with a highly proficient English speaker. In other words, in PBI, taking turns among students does not occur, and all English questions and answers are concentrated on the only one English learner. Therefore, as Min’s journal on Day 14 indicates, English learners are able to spend those 10 minutes speaking with the English interlocutor. As a result, PBI can provide very intensive 10-minute interactions providing time for output, and it would positively affect learners’ L2 processing and production.

**L1 transfer.** The results show that Min did not experience significant improvement regarding pronunciation and appropriate use of Yes/No when negative questions were given. The reason contributed to these results might be Min’s L1 transfer from his native language, Korean.

According to Ellis (1994), L1 transfer can take place due to different language systems and different sociolinguistic factors. Regarding different language systems, Ellis states that “L1 transfer is more pronounced at the level of the sound system than at the level of syntax” (1994, p. 316). Brown noted that acquisition of an accent-free skill in L2 speaking is almost impossible (2001, p. 268) especially for adult L2 learners. Regarding this, Min’s difficulties in L2 phonology acquisition also seems to be driven by his L1 phonology transfer. Particularly, when the Korean phonology system, which
has 1:1 matching between orthographies and sounds, is taken into consideration, acquiring the rule of English orthographies and their sounds can be more difficult for him because English does not have a one-on-one matching system, especially for vowels. Thus, this explanation of L1 transfer might account for the fact that Min has more difficulty in learning English vowels than English consonants.

Additionally, his difficulties in acquiring appropriate answers for negative questions might be also driven by L1 transfer. According to Ellis (1994, p. 317), L1 transfer arises when L1 and L2 interpret certain social contexts differently. Min’s inappropriate answering behavior for negative questions might have been the result of different interpretations of certain linguistic context (i.e., Americans considers the realization of events to answer for negative questions while Korean considers the contents of the negative questions). Answering questions requires learners’ quick and spontaneous reactions with either yes or no. However, sociolinguistically different interpretations of negative questions might interfere with L2 learners’ appropriate interpretation of certain linguistic contexts. Therefore, L2 transfer caused by cultural and social differences in language use might trigger L1 transfer, and L2 in that sense seems to be relatively harder to acquire.

In sum, Min’s insignificant development such as pronunciation and some sociolinguistic skill in using L2 seem to be taken place due to his L1 transfer, and this can be a result of a general limitation of adults’ SLA, not a limitation of PBI only.

Furthermore, the current study focused on the learner’s changes of affective variables, particularly motivation and L2 anxiety, which were identified as the influential factors contributing to the level of L2 WTC (MacIntyre & Charos, 1996).
Table 11 Min's changes in his affective variables

<table>
<thead>
<tr>
<th>Affective variables influencing WTC</th>
<th>Positively affects WTC when…</th>
<th>Min’s affective variables after PBI</th>
<th>Potential Contributors</th>
</tr>
</thead>
<tbody>
<tr>
<td>Motivation</td>
<td>High</td>
<td>Higher</td>
<td>Interaction with a highly proficiency ENG speaker</td>
</tr>
<tr>
<td>L2 Anxiety</td>
<td>Low</td>
<td>Lower</td>
<td>Talking in a private sphere</td>
</tr>
<tr>
<td>Perception on L2 competence</td>
<td>Positive</td>
<td>More positive</td>
<td>Increased self-confidence in his L2 ability</td>
</tr>
</tbody>
</table>

**Increased Motivation.** In this research, the comparison of exit questionnaires with the questionnaires for biodata collection shows that Min gained both in integrative motivation (3.6 → 4.8) and instrumental motivation (1.9 → 2.1) after PBI; his integrative motivation increased more significantly than instrumental motivation. As Garner & MacIntyre (1993, p. 4) stated, “motivation itself is dynamic” rather than static and restricted, Min’s motivation also showed some changes throughout PBI. Also, as Gass & Selinker (2001) mentioned that integrative motivation is a better indicator for SLA, Min can possibly expect more positive development in L2 learning after PBI.

The contributor of the increased motivation might be the one-on-one conversation with a highly proficient English interlocutor provided by PBI. Unlike traditional L2 classroom in EFL countries, which provides more speaking activities with classmates rather than the English native instructor, PBI offered 100% of one-on-one interaction only with an English speaker who has much higher L2 speaking competence than the learner. Regarding this, Min’s journal reported that having a very proficient L2 speaker stimulated him to become a more fluent L2 speaker like the English speaker that he encountered everyday; by comparing his L2 speaking
competence with the other speaker, he seems to stimulate himself more when he perceives more contrastive L2 ability between him and the English interlocutor.

In this sense, through 40 days of PBI, Min could have a positive disposition toward the L2 group represented by the English speaker, and he also seems to enhance the desire to interact with the speaker. Therefore, by supporting the learner’s chances to interact not with classmates at equivalent L2 level but with an English speaker at more proficient L2 level, PBI seems to motivate the learner to become a more fluent L2 speaker, as the English interlocutor is.

**Decreased L2 anxiety.** As the measurement of L2 anxiety shows, Min experienced a significant decrease in his L2 anxiety (5.2 → 3.3) after 40 days of PBI. Also, the journals he wrote during those days indicate that the embarrassment and annoyance that he used to feel while talking with an English speaker were also reduced. Both of his questionnaires and journals obviously show Min’s decreased level of L2 anxiety after PBI session, and it seems to be one of the significant results that PBI provided for him.

According to Young (1991, p. 427), anxiety can arise not only from a personal level (i.e., low self-esteem) but also from an interpersonal level (i.e., competitiveness, comparison with others). Additionally, Leary (1982) stated that L2 anxiety can also come from “audience anxiety, speech anxiety, and communication apprehension” (p. 102) when L2 learners perform or speak in L2 before others. Thus, he provided the constructs of L2 anxiety with social concerns (i.e., social evaluative anxiety). In this sense, L2 anxiety can be greatly influenced by how learners perceive others’ evaluation. It seems that this notion is very closely related with the social values of East Asian countries, which is face-protecting predisposition, and students’ language anxiety and fear can be inflated when a student with lower proficiency speaks in L2 in front of others.
Regarding this, PBI can be very effective for those learners who have high language anxiety when they speak in L2 in public. In PBI, the interaction between an English speaker and a learner is performed without any audience; in other words, learners in PBI are engaged in English conversations in a completely private sphere with no one else’s supervision or observation. Therefore, the learners who have language anxiety, such as social evaluative anxiety and stage fright, might have some benefits from PBI by interacting with an English speaker in their own private space.

Also, MacIntyre & Charos (1996) stated that the learners’ L2 anxiety can be reduced when they perceive their L2 competence positively. Regarding this, when learners experience some improvement through PBI, they can have more self-confidence in their L2 ability. By doing so, they might positively perceive their L2 ability, and eventually, it will contribute to a decrease of L2 anxiety. Therefore, ultimately, learners’ decreased L2 anxiety, increased self-confidence and positively perceived L2 ability can contribute more frequent L2 communication with an English speaker, which can be applied to non-PBI settings, such as future classroom L2 activities or L2 communication in their life.

**Conclusion & Implication**

The current research found that the learners who experienced no significant improvement in their communicative competence due to limited degree of WTC can have a positive effect from PBI in terms of the development of communicative competence. In this study, the learner showed overall development in each component of communicative competence, such as grammatical, sociolinguistic, discourse and strategic competence, and he seems to have experienced the most significant improvement in discourse competence. Regarding his overall development of communicative competence, PBI’s provision of authentic L2 contexts, which can
demonstrate the genuine use of L2, and increased opportunities to produce oral output, which might help the learner’s automatic processing of L2, appear to be the main contributors that facilitated the participant’s communicative competence.

The present study also found that PBI can have some positive influence on the learners’ WTC: the participant’s motivation was raised and L2 anxiety decreased, thus, both of them seem to contribute to the increase of WTC. This study shows that PBI’s provision of personal space to have English conversations might bring a better environment for those learners who have lower WTC driven by their introvert personality, low motivation and high L2 anxiety. Moreover, those learners who show a high tendency toward face-protecting predisposition or stage fright would experience more effective L2 learning in this setting because PBI does not involve any other audience around them. Therefore, they can feel free to talk in L2, and it brings more L2 oral development. As a result, PBI can be expected to have positive effects on L2 learners with lower WTC’s: both an increase of WTC and development of communicative competence.

However, even though the learner in this study showed significant development in both communicative competence and WTC, the results of this research cannot be generalized to all L2 learning populations who have lower WTC. For instance, different from Min’s case, some learners might have more anxiety in none face-to-face interaction because PBI does not provide any non-verbal cues which can help the comprehension of verbal communication. Also, some learners who do not have enough knowledge of L2 grammar might not feel language security because PBI does not provide explicit explanation of rules of TL.

Regarding this, future studies should discuss how the effectiveness of PBI can be applied differently to different learners in different contexts. For example, it is possible that each learner can experience significantly different degrees of
improvement according to their prior language proficiency, prior language education experience, and age. It is also possible that learners can experience different degrees of development according to who they talk to: native speakers of English, very proficient non-native speakers of English, less proficient non-native speakers of English, male English speakers, female English speakers. Moreover, learners might show different levels of improvement based on the curriculum of PBI: textbook-based instruction, task-based instruction, free conversation-only instruction. As a result, the effectiveness of PBI can be discussed from different perspectives to provide more valid expectations about learners’ L2 development.

Another limitation of this research is the limited length of the period to investigate the learner’s overall development of communicative competence. In this study, only 40 days were given to observe his communicative competence development. However, a previous case study, which also investigated one L2 learner’s improvement of communicative competence (Schmidt, 1983), observed the participant for more than one year, and that study provided more insightful examples to describe the learner’s developmental tendency, more detailed cultural and social effect, and more elaborate explanation about the learner’s L2 development.

Even though the current study also attempted to illustrate the learner’s L2 development with more insightful observation, PBI for two months did not provide a great deal of data to generalize his development of communicative competence. For instance, as Min’s morpheme accuracy rate indicates, the current study could not provide enough contexts which obligatorily require plural and possessive. Therefore, Min’s enhanced ability to use the plural and possessive morphemes accurately remains vague. Therefore, future study observing PBI’s impact on the improvement of L2 communicative ability should investigate learners’ developmental tendency for
prolonged period. By doing so, it can bring more insightful results to increase the validity of PBI’s effectiveness.

In a first attempt to provide evidence of the pedagogical usefulness of a telephone, the study suggests that telephones can be an effective educational device to develop learners’ L2 communicative skill and increase their desire to communicate in L2. Obviously, this result is more meaningful for learners in EFL contexts, which offer very few opportunities for authentic L2 use. Also, for those learners who have lower WTC, PBI seems to gain more potential by letting them speak more. Based on these preliminary results, PBI can be identified as a better pedagogical tool when some related issues, such as who should be the interlocutors? What materials should be used? What contents of language should be or can be taught? are taken into consideration.

Note1 Min’s level of anxiety and motivation were calculated in this research by averaging the points that he marked on the questionnaires. In the current research, 6-point Likert scales were used.

References


http://www.asian-efl-journal.com/march_05_yhc.php


Participation in and Opposition to the Ideology of English in South Korea: Insights from Personal Narratives

Elizabeth Root

Oregon State University, USA

Bio Data
Elizabeth Root is an assistant professor of intercultural communication at Oregon State University in Speech Communication. Before beginning her doctoral program, she taught EFL/ESL for many years in the USA, China, and South Korea.

Abstract
The purpose of this study is to investigate how the ideology of English influences English language classroom interactions in South Korea. The ideology of English is a set of beliefs about the necessity of English language skills, particularly the ability to converse in English. In order to explore how Korean students and their native-English-speaking teachers (NESTs) enact this ideology of English, personal narratives of experience were collected from 26 Korean university students, from two different universities, and 27 NESTs, teaching at several different institutions in Seoul. Through ideological analysis, six acts of participation in and seven acts of opposition to the ideology of English were identified with regard to Korean students. For NESTs, six acts of participation in and five acts of awareness of the ideology of English were identified. Critical reflection on how these acts influence classroom interactions is necessary to better understand the intercultural relationships between Korean students and NESTs.

Keywords: South Korea, EFL, language ideology, intercultural communication
Introduction

South Korea is a complex nation working to preserve its unique culture along with intense attempts to keep pace with the international market. One way in which Korea is attempting to stay competitive is to stress English language education:

the idea of what it means to be South Korean is transforming: increasingly, to be South Korean means to be South Korean ‘in the world’—a prospect that calls for the mastery of English as an index of cosmopolitan striving. (Park & Abelmann, 2004, p. 650)

This priority on English ability is demonstrated through the national curriculum, especially with the emphasis on communicative ability in English that began with the 6th national curriculum, implemented in 1995, and continuing with the 7th national curriculum, implemented in 2001 (Yook, 2010; Yoon, 2004). Under President Lee Myung-bak, the current South Korean government is pushing English education like no other government has before (Yook, 2010). For example, one goal is to eventually have a native-English-speaking teacher (NEST) in every primary school in the country, under the “Teach and Learn in Korea” (TaLK) program. This program targets foreign college students and has eliminated the requirement that students must have completed at least two years of undergraduate education (Card, 2008). This demonstrates that the desire to have a native-English speaker in the classroom is a high priority in the effort to have an English-speaking population. Native-English-speaking teachers (NESTs) from the United States, Canada, the United Kingdom, Ireland, New Zealand, and Australia are currently employed at private institutes, at public elementary, middle or high schools; and in university institutes or departments. Therefore, in South Korea, there is an emphasis on a native/non-native speaker dichotomy with regard to English language education.
This creates a situation in which South Korean students, besides learning English in the public school curriculum, are gaining more opportunities to interact with NESTs, thus creating an intercultural communicative situation within classrooms. While there can be many positive outcomes from exposure to other cultures, there are also complex aspects of the situation that influence how these two populations, students and teachers, interact with each other. NESTs and South Korean students do not meet each other on neutral ground; their interactions occur within a greater context. The purpose of this study is to highlight how intercultural interactions between NESTs and Korean students reinforce beliefs regarding the English language. First, aspects of English as a global language and the related ideology of English will be discussed. Then, through analysis of personal narratives, ways in which both teachers and students participate in and oppose this ideology of English will be described. In the classroom setting, NESTs benefit from their status as native speakers while students must work hard to improve their speaking skills.

**English as a Global Language**

The emphasis on English ability in South Korea is a result of English being the current dominant language of globalization, international trade, production and financial markets. The rise of English as a dominant language has not necessarily happened by chance: “English is not simply another language or an innocent term” (Holborow, 1999, p. 60). English is now the language of the global economy because of the lasting effect of colonial education and trade (Spring, 1998). However, as colonialism progressed, from the British Empire onward, the spread and dominance of English receded from overt attention. Emphasis was instead placed upon the economics of how the world was becoming interdependent (Gibson-Graham, 1996; Stromquist & Monkman, 2000). In fact, the now “fashionable” use of the term *globalization* is one way in which the
spread of English has been somewhat concealed; definitions and debates of globalization tend to focus on economics, culture, identity, politics, and technology (Block & Cameron, 2002). However, aspects of language are also an integral part of globalization; worldwide social relations cannot occur without some common language within which to communicate (Block & Cameron, 2002). In fact, Murata and Jenkins (2009) highlight how unprecedented and diffuse the spread of English has become; English is most often chosen as the lingua franca in international settings because of its global spread.

Nevertheless, the link between English and current globalization is a debated and contested topic (Block & Cameron, 2002; Canagarajah, 1999; Phillipson, 1992, 2009). For example, some would like to consider the dominance of English as somehow accidental, as Kaplan (2001) explains: “It is unlikely that there is some grand conspiracy among English speakers to disseminate English worldwide, on the contrary, the spread of English is largely accidental…” (p. 17). While this is a safe view to hold, this does not seem to adequately address the history of colonialism. Others are more willing to acknowledge the role of the past and how that relates to the current state of globalization. Macedo, Dendrinos, and Gounari (2003) stress how a neoliberal ideology, using the concept of globalization, continues the promotion of language policies that favor English. In fact, aspects of English language imperialism have been raised (Canagarajah, 1999; Phillipson 1992; 2009). As Phillipson (2009) describes, linguistic imperialism is a concept that considers how and why certain languages are dominant in the international scene, specifically focusing on whether or not there is a link between linguistic dependence and economic dependence. Bhatt (2001) specifies that another way to view the spread of English, as compared to linguistic imperialism, is the econocultural model, based on Quirk (1988) and Brutt-Griffler (1998). Instead of emphasizing imperialism, linguistic pragmatics is the focus of this model which traces
the economic influence of industrial capitalism based in England and the United States in the nineteenth century through to the establishment after World War II of such institutions as the United Nations, the World Bank, the World Health Organization, and, eventually, the Commonwealth and the European Union. Based on Bourdieu’s (1977) notion of cultural capital, school systems in commonwealth countries reproduced the English language as “symbolic capital” (Bhatt, 2001, p. 533). In this vein, Brutt-Griffler (1988) stresses how English learners and speakers have agency in their use of language: “…to use a language is to become an agent of language change.” (¶ 33). It is obvious from these different approaches, linguistic imperialism or pragmatics, how the link between the English language and aspects of globalization is debated.

Whether it is acknowledged that English is part of a lingering colonial legacy or not, the fact that one language has spread globally automatically creates inequalities for those who do not grow up in an English-speaking environment. At the very start, inequalities exist because some people live in an atmosphere where English is their first language and therefore have a linguistic advantage. For others, the amount of time, effort, and money it takes to acquire even a rudimentary ability in English automatically puts the rest of the world at a disadvantage (Van Parijs, 2000). Even the distinction between native/non-native speaker of English which exists in English teaching and learning contexts sets up disparity between those who can be labeled as native speaker and those who cannot. While there have been debates about these labels (Brutt-Griffler, 1998; Brutt-Griffler & Samimy, 2001; Jenkins, 2006), the end result is that despite the widespread distribution of the research into linguistic imperialism, it has not so far led to noticeable changes in English teaching and teacher education policy. The best that can be said to date is that it has raised many teachers’ and teacher educators’ awareness of the extent to which the spread of English works in native
speakers’ interests and sometimes marginalizes nonnative speakers. (Jenkins, 2006, p. 169)

Therefore, while some might prefer to see the spread of English in supposedly neutral terms, simply as an unstoppable and ongoing process, others are willing to voice concerns that these processes and changes are simply creating more inequalities in the world.

The Ideology of English

One way in which to view the inequalities that exist because of English-language dominance is through the lens of ideology. An ideology is a “pattern of ideas, belief systems, or interpretive schemes found in a society or among specific social groups” (Hall, 1989, p. 307). Certain beliefs about the English language and about how it should be taught create an ideology of English. For example, Phillipson (1992) describes certain tenets of English education that have been upheld strongly in the English-teaching profession even though research does not fully support these beliefs. These tenets include the idea that English should be taught monolingually, that use of other languages should be completely discouraged within the classroom, that the best teacher should be a native speaker of English, and that greater results will occur if more English is taught.

These tenets of English education are strongly adhered to in the South Korean context (Root, 2009). For example, one study conducted by Ko (2008) examined Korean teachers’ reactions to a potential English-only policy considered for their university English-language program, demonstrating how certain administrators support the belief that English should be taught monolingually, even if taught by Korean instructors. Another aspect of the ideology of English is the acceptance of
English skills as a necessity. In Korea, Yook (2010) asserts that learning English has become somewhat of a national obsession, describing how

   English has become the most important foreign language in Korea, and it permeates almost every aspect of Korean life. English has become a critical part of high-stakes tests, deciding major opportunities in the lives of Koreans (e.g., college entrance, employment, and job promotion). (p. 4)

   Within the ideology of English, conversational skills are seen as the most important determiner of language ability, a belief demonstrated by the current president’s efforts to reform foreign language instruction so that all high school graduates would have the ability to carry on a conversation in basic English with foreigners (Yook, 2010). The ideology of English also includes aspects of how to best study the language, which involves committed effort and actual practice using the language in real-life situations, with native-English speakers often being depicted as the most advantageous teachers to help students improve in their oral skills (Root, 2009). One example of this aspect of the ideology of English is enacted through the existence of “English villages.” These “multi-million-dollar” projects, developed by some provincial governments in Korea and staffed by NESTs, are meant to provide Korean students with an experience of English language immersion (Yook, 2010, p. 16). Another example comes from a study conducted by Yoo (2011), which examined attitudes Korean students held toward different English accents. The results indicate how the students favored North American English as the “standard” pronunciation and that the least preferred accents and speakers were identified as being from non-native-English contexts, once again indicating how students tend to favor native-speakers of English as language models. These examples demonstrate that strong ideological beliefs regarding the English language exist in South Korea.
In any context, however, multiple ideologies exist. Some ideologies are advantaged over others and some are subdued. This creates a more privileged perspective, which is the hegemonic ideology. Hegemony is “the condition by which dominant social groups construct a social consensus” and this is often “legitimated through positions in the media, government, churches, and schools” (Kubow & Fossum, 2007, p. 76). The ideology of English is a dominant ideology in language education in South Korea. Since it is a hegemonic ideology, it creates aspects of privilege and dominance. In such a situation, NESTs are the more privileged participants. NESTs can take their native-speaker status for granted; this status enables them to obtain jobs as English instructors whether or not they have had formal education training. Since the main goal is to help students improve in conversational and oral skills, NESTs in Korea can run classes simply by showing up and speaking with students, if they so desire. However, Korean students, within the ideology of English, are in the less privileged position. By accepting this ideology, students submit themselves to years of study that can cost them greatly, both in time and money.

Participation within an ideology is a complex act, since the dominance of any particular ideology is not necessarily a deliberate and conscious process (Foss, 2004). A dominant ideology is often presented in multiple ways as being the “natural” or “accepted” norm. At the same time, there also exists resistance to dominant ideologies. However, Canagarajah (1999) points out a need to clarify the distinction between resistance and opposition, based on Giroux. Giroux (1983) describes resistance as “displaying ideological clarity and commitment to collective action for social transformation,” but opposition is “unclear, ambivalent, and largely passive” (as cited in Canagarajah, 1999, p. 98). Canagarajah (1999) explains his own analysis of English language students’ opposition:
In most respects, theirs is a vague, instinctive, oppositional form of behavior that lacks ideological clarity. It fails to sustain consciousness-raising or collective action for change. Students’ behavior in the class is ambivalent, containing elements of accommodation as well as opposition in response to the conflicting pulls of socio-economic mobility, on the one hand, and those of cultural integrity on the other. (p. 98)

Therefore, with regard to the ideology of English, teachers and students engage in acts of participation and opposition to the dominance of English on a daily basis within the classroom setting. The purpose of this study is to describe ways in which participation and opposition happen within the specific context of language classrooms in South Korea.

**Data Collection and Analysis**

This paper is based on a larger study conducted in Seoul, South Korea, which initially focused on the intercultural relationship between NESTs and South Korean students in language classrooms. This specific context was chosen because of the importance of the role relationship between teacher and student within a foreign language classroom setting (Ellis, 2007). The initial study was designed as an instrumental collective case study (Creswell, 1998; Stake, 1995). While most case studies examine a bounded system, such as an event or a relationship that is bounded by a specific time and place, the data in this study are not bound by shared times or locations. In other words, the student participants are not telling narratives about their experiences with the specific teacher participants in the study, or vice versa. Instead, the boundaries for the narratives are based on the participants having an experience with either Korean students or NESTs. This allows for a wider range of experiences and narratives to be collected. The data prompt for the initial study was designed to collect stories about
what was memorable for participants, not to bind them to sharing stories specifically
about a set population of teachers or students. Also, since the data analysis method for
the study was based on rhetorical criticism, any narrative generated by a participant
that fit the boundaries of the study, a recalled experience about a teacher-student
interaction, was considered valid for data analysis. The results from this study, as
consistent with qualitative research, are not meant to be automatically generalized to
wider populations of teachers and students, but instead the results are meant to
highlight issues that exist within this data and to provide a basis from which to
continue research in these areas.

Personal written narratives were collected from 26 Korean university students
and 27 NESTs throughout a six-month sojourn in Seoul. Narratives were chosen as the
main source of data based on Fisher’s narrative paradigm (1984, 1985, 1987) and
Bruner’s (2002) description of the use of narratives. Stories are useful to consider with
regard to research, since “Stories are surely not innocent: they always have a message,
most often so well concealed that even the teller knows not what ax he may be
grinding” (Bruner, 2002, pp. 5-6). This approach is similar to Ellis’ (2002, 2008)
approach to studying learner beliefs based on metaphor. Metaphors are important
because they provide insight into underlying conceptual constructions that people hold
(Lakoff & Johnson, 1980). Since these conceptual constructions are not often in our
conscious awareness, research based on metaphors is useful to highlight beliefs that are
often overlooked or unnoticed. While this may seem to be an indirect approach to
studying learner beliefs, it actually allows researchers to access beliefs that would be
difficult for learners to directly identify, state, or explain. Narratives are similar in this
manner, since they, like metaphors, allow participants to share aspects of their beliefs
and underlying expectations of a situation.
The original narratives collected from participants were based on a broad story prompt: “Write a story about a memorable experience you had in a language classroom with a NEST/Korean student.” Since this research was not meant to mimic interviewing, the purpose of this vague and indirect story prompt was to generate a personal narrative. A more specific story prompt, such as asking for a positive or negative experience, would have led the data in a specific direction. Therefore, this story prompt was specifically designed to be vague to discover what personal memories the participants would focus on. Also, this story prompt was designed for the initial goal of the larger research study, which was to focus on the intercultural relationship between teachers and students. Participants were not informed that the purpose of the study was to examine beliefs about teaching and learning English, since one purpose for narrative research is to access participants’ underlying beliefs and expectations without their overt attempts to directly state or indicate these beliefs.

Once the data was collected, there was evidence within the data of certain beliefs held about the English language and about how to best learn English. Since one important feature of qualitative research is its emerging and evolving tendency (Creswell, 1998; Denzin & Lincoln, 2000; Lincoln & Guba, 1985), it became clear that the data provided insight into ways in which the participation in and opposition to the ideology of English occurred. Therefore, the research results examined in this article highlight the analysis that occurred once aspects of the data were identified as highlighting ideological components in the relationship between Korean students and NESTs. At this point in the study, ideological criticism was used to analyze the data (Foss, 2004, 2006; Gerzon, 1996; Sillars, 1991). Both teacher and student narratives were initially analyzed separately by identifying ideological aspects of each group. The stories were then examined for any mention of how certain actions or communication supported these ideological beliefs about English. At the same time, any distinct
representation of opposition was also documented. Once these were noted, the acts were combined together to present as complete a picture as possible of both participation and opposition to the ideology of English as presented in this particular data set.

For the student narratives, 19 females and seven males participated in this study, ranging from 20 – 40 years old. All participants were university students, the majority being fourth-year English language majors. Students indicated that they had studied English anywhere from 2 – 17 years. Most of the students did not meet a NEST until their first year of university, and so the majority of stories focused on experiences that happened in university. The narratives were produced through a writing assignment from two different university classes taught by Korean professors.iii Both courses were fourth-year classes in language translation for English majors. The two different universities which these students attended represented different rankings of prestige within Korea, therefore providing a range of socio-economic backgrounds with regard to students. In order to respect the students’ native language, the stories were submitted written in Korean. One hired translator worked with all the stories to translate them into English. After a back-translation check, the English versions appeared to represent the Korean versions as accurately as possible. These English versions were the stories used for data analysis.

For the teacher narratives, 20 males and seven females volunteered to participate in the study, ranging from 26 – 60 years of age. All teachers were recruited through word of mouth and flyers distributed at two universities and through the KOTESOL (Korea Teachers of English to Speakers of Other Languages) organization in Seoul. Since participation in the study was based on volunteers, any teacher who was willing to participate was accepted into the study, regardless of where they taught. These teachers were from five different countries of origin: ten from the USA, nine
from Canada, five from Australia, two from Britain, and one from New Zealand. The participants had a range of experience teaching English in elementary schools, private institutes for all ages of students, and also university classes; therefore, their stories cover a wide range of settings. 11 stories were submitted through e-mail. The other 16 stories were collected through a process that involved a recorded oral telling of the story. In an attempt to make these recorded stories as similar to a written story form as possible, the transcribed stories were e-mailed back to the participants to allow them to change, add, or edit any information that they might desire. Therefore, all participants who orally told their stories still had a chance to read and edit their stories in a written form. These were the narratives that were then used for data analysis.

**Research Findings**

Whether conscious or not, the participants provided evidence of how they both acquiesced to and also questioned their participation in the ideology of English. In order to explain and provide evidence of these different actions, they will be separated into different categories. However, these actions do not necessarily happen separately from the other. Instead, individuals within this ideology act both in participatory and oppositional ways, sometimes all at the same time. First, the students’ participation in and opposition to the ideology of English will be outlined. The teachers’ actions will be discussed second.

*Students’ Participation in the Ideology of English*

Throughout the student narratives, the Korean student participants provide examples of how they rationalize their participation in a belief system that privileges English. These examples indicate that students accept the dominant belief that oral skills are the
important marker of English-language ability. Six specific acts will be briefly explained.

Acknowledgment of the importance of English.
The students in this study candidly discuss how English is not an option for them to study. It is their accepted reality that they need to spend time and energy to improve their English skills in order for them to be successful. One student, for example, articulates that “the ability to speak English is a must” (SS 15). Not only do students acknowledge the necessity of English, but they also do not overtly question this assumption. English is also privileged over other languages, as this student explains: “I really would like to learn another language, but my current English skill is bothering me from doing that” (SS 2). Therefore, throughout their stories, these students reaffirm their belief that learning English is very important to them.

Personal blame for lack of progress.
When things do not go well in an English-speaking situation, participants quickly take the responsibility or blame for their lack of progress. Many students voice concerns that they could have worked harder or that they could have done better in class or on a test: “I feel shameful for not being able to improve my English skills very much” (SS 1). Another example of a reaction to a lack of language ability is expressed in this narrative: “I still blame myself after having a conversation with foreign instructors and say to myself, “Oh, you one big stupid animal, I should have said like this, not like that” (SS 4). From this data set, it appears that students feel a strong personal responsibility to make progress in English, both in and out of the classroom. Their comments demonstrate that the research participants have internalized the ideology of
English and are convinced that they must put in the effort themselves to make progress.

_Discussing the benefits of studying with a foreign teacher._

Many narratives indicate that students can slowly grow to favor NESTs over Korean instructors, once students adjust to studying with a foreign teacher. As these students explain, the benefits of studying with NESTs include the chance to learn exact English expressions and correct pronunciation, to have instant feedback, and to learn about another culture. Some NESTs are described as creating relaxed atmospheres so students are more comfortable to speak English. Some students even describe their preference for how foreign teachers grade based on effort and progress: “I like…their ability to see the improvement a certain student has made and evaluate the student based on that” (SS 5). Therefore, some students are not reticent to list reasons why they might prefer taking classes with NESTs.

_Criticizing the Korean way of learning._

While this is less prominent than praise for foreign teachers, there are comments in the narratives about dissatisfaction with the traditional way that students have studied English in the Korean school system. One student describes a moment when “I really felt embarrassed by the Korean way of teaching, which mainly focuses on memorizing the contents” (SS 14). This focus on memorization is also critiqued by another student who focuses on vocabulary and how serious misunderstandings can occur when Korean students memorize only one meaning for a vocabulary word (SS 7). These critiques focus on how little preparation the students had received to handle spontaneous conversations in English. These small criticisms are aimed at the disparity between their background educational culture and how the main determiner of English ability is based on conversational skills.
Setting goals to improve speaking skills.

Within the narratives are goals that research participants have set to improve their English abilities. These personal goals indicate an internalization of a priority to reach a fluent level of conversational English: “I have to work on my English conversation skills until I can proudly answer my future instructors that I am a student majoring in English language and culture whenever they ask me” (SS 12). Another student explains how determined he is to study harder: “The effort of trying to catch what the instructors were saying could be compared to a suspense game. After a semester of this ‘suspense,’ I felt that I should level up my speaking skills, even through private language school programs” (SS 9). This act of studying harder and putting more effort into the process is evidence of specific decisions made on the part of the students to find a way to excel in an atmosphere where oral English skills are highly regarded.

Lying in English.

While only one student mentions this aspect of participating in conversation classes, it provides important insight into students’ behavior in the classroom:

I once heard that Korean students tend to give answers in English conversation classes which are actually a lie since it is hard for them to communicate accurately in English. Basically they don’t mean to lie, but when they face difficulty in communicating with the instructor and the instructor is leading the conversation, sometimes they have no choice but to give a false answer whether they want to or not. (SS 10)

While some might consider “lying” to an instructor a form of opposition, these “false” answers stem from the students’ overwhelming belief that they must speak up in class, no matter what they say. This excerpt illustrates that students can be so convinced they must participate in class, it ultimately does not matter whether or not their answers are
accurate representations of themselves or their opinions. Therefore, this tactic is in support of the ideology of English, where conversational skills are prioritized highly.

These six acts demonstrate ways in which students have accepted aspects of the ideology of English and how they enact their participation to support English as a global language. These examples demonstrate ways in which language students rationalize their situation and come to terms with an undisputed fact that they must learn English. These actions are strategies to ultimately help them make progress in a subject that has been set as a requirement for them. While these actions are present in their stories, it cannot be assumed that students are overtly aware of how these actions continue the existence of English hegemony.

**Hints at Students’ Opposition to the Ideology of English**

Research participants in this study also show moments when they resist or question the assumptions of the ideology of English. Since the hegemonic emphasis on English has been adopted and integrated into the structure of the Korean educational system, it is difficult for students to engage in active resistance to this. Resistance in this context could block students’ routes to success in the education system. Therefore, while the mentions of opposition are subtle, it is still important to pull out these acts and focus on them. Seven specific acts compiled from the student data will be briefly explained.

*Silence.*

Silence is the strongest form of opposition that appears in these narratives. Not only is this mentioned the most frequently throughout the stories, but silence also seems to be the most effective way for students to capture the attention of foreign teachers. The presence of silence in the classroom puts pressure on teachers to figure out what is happening culturally with their Korean students. One student explains how there are
times when “long silence attacks the whole classroom” (SS 11) when a NEST asks for student responses: “If the foreign instructors do not have that much experience teaching Korean students, they feel very anxious when they experience this long silence in class” (SS 11). Another student describes it as teachers getting into “a panic when they see Korean students staying silent when asked questions” (SS 14). Besides silence being the easiest form of opposition for students to enact, there is no other situation present in these narratives in which they are so able to affect the teacher’s behavior and attitude in class.

Students do not explain the silence in the classroom as being a form of opposition, however. Instead, they explain their silence based on shyness or worrying about giving a wrong answer. Students also discuss in their narratives how difficult it is for them to feel comfortable speaking up spontaneously in class, since their educational background is not one that fosters this behavior in the classroom. Therefore, the situation of silence in the classroom is complex to analyze since silence can result for a number of reasons. Are students silent in class because they do not want to speak English or interact with the foreign teacher, or are they silent because while they want to speak English well and desire to engage with the teacher, they do not yet have that capacity? For whatever reasons that students are silent in class, however, silence is still depicted as one of the most powerful strategies students have to resist participating in the dominant ideology. Silence, from the teacher’s perspective, is seen as non-engagement. When the dominant ideology favors direct interaction, one way to resist is to not engage in interaction.

*Avoiding English classes with foreign teachers.*

These students hint at something they call “foreigner fright,” which happens before their initial contact with a foreign teacher and involves feelings of fear and dread. This
negative anticipation can result in students doing their best to put off enrolling in required classes. One student actually admits to having dodged ever studying with a foreign teacher: “I would have to make an embarrassing confession as a student majoring in English. That is I have not taken a single class instructed by a foreigner…the basic reason seems that I have a great fear about taking a foreign instructor’s class” (SS 12). While this may not be a typical situation, it does demonstrate that avoidance is a possibility for some students who have little confidence in their English-speaking ability. Other students also avoided classes that they had registered for at private institutes. Therefore, when it is in their power, students avoid classes with NESTs as a way to circumvent being forced to interact in English.

Doubting the situation.

In these narratives, a few students voice doubts about their study experience. As one student states: “I sometimes thought at that time, ‘I don’t have any reason to go overseas and my school does not require students to take conversation exams, so why go to a prep school and learn English’” (SS 2)? This student cannot see a specific need in his life to have an ability to converse in English, and so therefore wonders why he should work hard to gain this skill. While Korean students may not be overt in their doubting or disagreeing, these stories demonstrate how they still do doubt and question their circumstances.

Grumbling.

There is one story that mentions how students grumble about certain situations in the classroom. This particular example focuses on the fact that the instructor is British: “Many of my classmates grumbled that they could not understand what the instructor
was saying in class” (SS 2). Some students indicate favoring a North American accent, even though they are not guaranteed a NEST with an American accent. This group response of grumbling indicates that when students are not happy with the situation they do have a small outlet to complain about their situation, at least to other students.

Speaking Korean in class.
One result of the difficulty of interacting in English is explained by a student who specifically speaks Korean in defiance: “…whenever the instructor required us to talk in groups, my eyes would roll back and forth to figure out whether the instructor was close or far away from me. Whenever possible, I talked in Korean in that class in group debates” (SS 5). Instead of trying to participate in English, this student makes a conscious choice to speak Korean. While this might seem to be the most obvious form of opposition available since many NESTs are not fluent in Korean, most students do not mention this strategy in their narratives.

Describing cultural differences.
These narratives include comparisons between Korean teachers and foreign teachers to help rationalize why classes with foreign teachers are so difficult for many students. These cultural observations are not made merely because they are of interest to the students; these differences are pointed out as a way to help students cope with their discomfort and awkward feelings when introduced to classes with NESTs. These observations help students realize there are two systems of values and expectations functioning simultaneously. These observations also alert students to the fact that NESTs are operating under different rules and perspectives than they are. This awareness gives students a more complete picture of their situation.

Giving advice.
One final way that these students indicate opposition is through the messages in their stories. Student participants express their opinion about how to make their educational experience better and more successful. For example, one student wants to make a point that NESTs are perhaps not a good choice for beginning language students: “For beginners, Korean instructors seem to be better than foreign instructors” (SS 20). These hints of advice, or of students’ voicing their opinions, are also ways that students try to subtly resist their experience within the ideology of English.

While it might not be an easy task for students to oppose participation in the ideology of English, these narratives demonstrate that students have moments when they doubt this system that privileges English to such a degree. As mentioned previously, these acts might occur in separate moments, but there also ways in which students engage in participation and opposition at the same time. For example, the one student who directly questions why he needs to study English at all followed that question with this immediate thought: “Of course, all that proved to be nothing but an immature justification to stop learning English” (SS 2). Therefore, while these actions are separated in this section for the sake of distinction, students’ enactment of them are not always so distinctly either in support of or engaging in opposition with the ideology of English.

*Teachers’ Participation in the Ideology of English*

Just as students give evidence of their participation and opposition in the ideology of English through their stories, the teachers also are participants in this hegemony. Throughout their stories, these NESTs provide examples of how they rationalize their participation in a belief system that privileges English and their native-speaker status. These examples show how teachers reinforce their privilege and their status in their role as NESTs in South Korea. Six acts of participation will be briefly described.
Critiquing Korean culture.

Many teachers include some criticisms about Korea in their narratives. The majority of critiques focus on the Korean educational system or student behavior. Some of these teachers are not impressed with the general educational system in Korea. Comments focus on a lack of academic accountability, low standards, and curriculum and courses that do not challenge students enough. Another topic that is critiqued is Korean culture in general. While this type of commentary demonstrates a certain level of frustration that foreign teachers have, these comments also demonstrate a level of awareness regarding Korean culture and Korean systems. However, the cultural awareness in these cases is communicated in a negative manner, since teachers are describing their awareness in the form of a critique. While NESTs are aware of their presence in a foreign culture, they still feel free to admit that they either do not understand, respect, or accept aspects of Korean culture. Therefore, while the ideology of English focuses on language use, alongside the privileging of English there is a privileging of certain cultural values. These narratives demonstrate how some NESTs critique Korea based on their worldview of how the educational system should operate, how students should behave, and also how certain general aspects of life should occur.

Comparisons to personal culture.

Not only do NESTs sometimes criticize Korean culture in their stories, they also include commentary about how they might have handled a situation differently. This comparison back to their standard of behavior or to their home culture indicates how they feel their norms or values should apply to the Korean context. One teacher explains how, as an undergraduate student, he would usually go up and introduce himself to his professors, spend time talking with them, and would specifically visit
them in their offices so that they would know who he was. This is in contrast to how this teacher views his Korean students’ behavior: “…they are very good at obscuring themselves. When you ask a question in class to the whole class, you will never get an answer…every student in class buries their head, muttering to themselves, ‘Not me, not me, don’t ask me!’” (TS 5). This teacher values his approach to his education as advantageous compared to the students he teaches now. When teachers set up comparisons between Korea and their home culture like this, they present their own culture as the standard by which to gauge behavior and decisions. These stories indicate how some teachers attempt to apply their cultural standards to the situation in Korea and how they then might expect these standards to hold firm through their experience there. Since these NESTs hold the position of teacher, they have privilege over their students to grade them based on their own cultural standards instead of the students’ cultural standards.

*Maintaining a separation from Koreans.*

Within these narratives, there were moments when NESTs would foster a feeling of separation from Koreans. For example, one teacher describes his experience at a private elementary school: “The Korean teachers followed the curriculum set by the government, but the foreign teachers could pretty much do what we wanted to, focus on what we wanted to focus on” (TS 4). Separation from Korean society means a certain freedom from systems of accountability for NESTs. The work environment at times fosters this feeling of separation, since in some situations NESTs are used as a way to advertise and bring in customers for certain schools and private institutes. While NESTs may feel used by the system, such as objects of hire and advertisements for certain schools, they also can benefit from this situation. There were moments in the narratives when some NESTs indicate that they do feel different, separate, and
perhaps even “above” the Korean society in which they live. Not being required to answer to all the dictates and demands of Korean society is a privilege that certain teachers voice.

*Acting as expert.*

A few teachers demonstrate how they internalize an attitude of language expert. This is communicated in how teachers advise students in language study. One teacher tries to steer students away from too much memorization, since that is an “arduous grind” (TS 5). Instead, he wants students to spend “20 minutes a day reading something that is interesting…, talking to someone, watching a TV show…, listening to music and singing that music” (TS 5). Another instructor is very honest about how he feels Koreans are studying English for the wrong reasons:

> They’re learning it for jobs, not for everything else that is beautiful about English, for travel, for meeting foreigners—I mean everyone speaks English, really, around the world…so if they (Koreans) could focus more on traveling, on meeting people, on enjoying movies, books, cultural aspects like music, and to enjoy all that, then everyone would want to learn English. (TS 4)

Whether or not this is helpful advice for language learning is not the specific point of this analysis. Instead, these comments about how to learn English and why students should learn English indicate certain beliefs about the inherent goodness of English, and how it should be something that is enjoyed and also used in a practical manner. The other complex issue with regard to how advice is given is that many NESTs in Korea are monolingual. Whether or not NESTs are fluent in any language other than English, they still communicate opinions about how students should study English. This advice comes from the status that native speakers hold in such a situation.
Descriptions of favored students.

The students and classes that are favored by these teachers, or are presented in the most complimentary manner, are those who have a higher level of English. Some teachers are quite willing to admit this partiality: “I just love the advanced classes, it is my admitted bias. A lot of people do” (TS 20). With more advanced students, these teachers can experience more depth of interaction and often are more challenged in the classroom in a way that they enjoy. Descriptions of high-level English speakers indicate how NESTs prefer students who can interact more successfully in an English-speaking world. When students are able to adjust and interact more on the teachers’ terms, the situation is more comfortable for the teacher.

Positive feedback from students.

When teachers are treated well by students, they interpret this as reinforcement that they are doing the right thing. This treatment spills over into life outside the classroom, as one NEST describes how many students came to visit him in the hospital when he had broken three ribs: “Probably more than a hundred students came to visit me in the hospital. They were always bringing little cardboard boxes of juice…It was really touching.” (TS 18). Another teacher describes how students gave him Christmas presents his first year in Korea, as students “wanted to help me adjust to Korea” (TS 8). These acts of kindness tend to be interpreted as reinforcement that the teachers are in the right place, at the right time, doing the right thing by teaching conversational English. When students show signs of welcome and friendship, these teachers feel they must be engaging in a useful, helpful act of teaching English. Students’ positive responses can be interpreted as somehow contributing to the greater good.

These six acts demonstrate ways in which NESTs show their acceptance of the ideology of English and how they enact privilege within this ideology. By engaging in
these actions, teachers support beliefs that they are language experts based on their native-English-speaker status. They also continue to privilege their cultural expectations and foster situations in which the more advanced speakers of English can rise to the top of the class. While these actions are present in their stories, it cannot be assumed that NESTs are overtly aware of how these actions continue the existence of English hegemony.

Hints at Teachers’ Awareness of the Ideology of English

Also present in the teachers’ stories are moments when teachers seem to doubt their actions. These moments are called “hints of awareness,” meaning that the slight mentions of these in their stories indicate that NESTs do not completely accept the ideology of English without any questions or doubts in their mind. These moments are also indications of consciousness of privilege and how teachers are involved in a larger global action; these moments provide insight into how teachers consider their role in the enactment of the ideology of English. Five acts will be described briefly.

Admission of job as easy.

There are moments in the data when NESTs articulate that their jobs are easy. In these moments, these teachers demonstrate a sense of privilege in living as a NEST in Korea. As one teacher puts it: “…I ended up at the university, and it was a nice cushy job” (TS 2). Other NESTs not only indicate that their job is easy, but also that it is highly enjoyable. One instructor describes his time with a private student that he taught on the side: “For me, personally, it was like being a middle school kid again, having these fun conversations, having snacks, and just enjoying life” (TS 18). While not all jobs are this relaxed, there are still definite instances in these stories when NESTs explain how
their jobs are not always so stressful or difficult; these statements hint at an awareness of the privilege with which they are living.

*Questioning the situation.*

In this data, NESTs are not always completely confident in aspects of their jobs. There is sometimes an awareness of the influence that teachers have, or a concern about trying to better understand Korean culture or their place within that foreign culture. One teacher voices these questions after a situation where a student found out his birthday and surprised him with a cake and some gifts in class: “This is what I don’t know how to figure out. Do students act out of obligation or because they want to? How much is it that students have to curry favor with me” (TS 3)? This instructor’s awareness of his status as a teacher and his influence over students’ lives is communicated through these questions. When asking these types of questions, NESTs demonstrate their difficulty in understanding experiences with students. These questions are examples showing that NESTs do try to analyze and figure out aspects of their influence and how they might have affected the students, the situations, and the circumstances around them.

*Desire to act appropriately.*

Also in the teachers’ stories are indications of attempts to adapt to Korean culture and to honor the students’ culture in the classroom. These moments demonstrate a desire on the part of NESTs to be conscious about the fact that they are living and teaching in a different culture. One example is a story about how a teacher waited too long for a student to provide an answer in class; the student ended up bursting into tears. While the teacher was trying to be patient and provide time for the student to think of an answer, the end result was uncomfortable for everyone (TS 21). This NEST has now
become more aware about changing some of his classroom techniques to better meet Korean students’ needs. Some NESTs, then, are aware of differences and the fact that they are in a foreign culture, and whether successful or not, they still try their best to be appropriate in different situations.

Admission of change or lessons learned.

Related to the previous category of trying to be culturally appropriate, NESTs also indicate they learn lessons and change somehow because of these experiences. The one teacher, mentioned previously, who is horrified to have a student cry in his class explains how he has learned to handle “wait time” differently in class (TS 21). Another instructor, who received many strong complaints about her approach to a teacher-training course, admits:

I’ve made many changes to my teaching approach since that time. It seems to be working for the trainees. There have been no complaints brought to my supervisor. Although the problem now is that I have no idea if I’m satisfied with my course. (TS 13)

Whether or not she is pleased with the course, this teacher demonstrates her willingness to change and adapt to the situation, no matter how difficult this seems. Therefore, in the data, there are those admissions that teachers have learned lessons about how to approach Korean students in the classroom and that they are working to change and adapt their methods to deal with these new realizations.

Realizing students have limited resources.

While very few teachers made any comments about students’ opportunities to speak English outside of the classroom, one teacher does describe his acknowledgement that
advanced skills in English require certain opportunities: “Another thought that came...was a realization of how much English is tied up with international experience” (TS 18). While this teacher does not go into much detail, he at least is aware that it is difficult to reach certain levels of English while only being immersed in the Korean context. English, then, is available more to some than to others, such as those students who have funds to travel or opportunities to have foreign acquaintances in Korea.

Therefore, these moments of awareness of the privilege of NESTs are evident in the teacher narratives, even if these moments are not mentioned frequently. These five acts all indicate that teachers have occasional insight into the fact that they have certain privileges with regard to their job and also how they are aware of their present role in living in another country. To more completely understand the teachers’ participation in the ideology of English, however, these signs of awareness must be considered with their acts of participation. Similar to the students, these teachers exist within this ideology of English without any guarantee of their overt identification of such an ideology. Therefore, teachers engage in participation and show signs of awareness all at the same time, existing somewhere between the tension of the two.

Conclusion

The results from the ideological analysis of these two sets of narratives provide insight into how people in an intercultural relationship enact aspects of an overarching ideology. The ideology of English is a strong mental framework that favors English language skills over any other language skills, specifically the ability to converse in English (Phillipson, 1992; Root, 2009). As depicted through this data, both Korean students and foreign teachers participate in this hegemonic ideology by reinforcing the necessity for students to work hard, to practice speaking, and to learn from native speakers of English. At the same time, the participants also show signs of questioning
and doubting this ideology. As one student articulated: “I don’t have any reason to go overseas and my school does not require students to take conversation exams, so why go to a prep school and learn English?” (SS 2). NESTs, on the other hand, might question their privilege within their job position, as this one teacher demonstrates: “Do students act out of obligation or because they want to? How much is it that students have to curry favor with me?” (TS 3). While the opposition to the ideology of English appears to be less of a factor than the total picture of upholding of the ideology, the presence of opposition allows for a more complete picture of the situation.

Both parties in this student-teacher relationship exist within this ideology. Their narratives show that on a daily basis, they engage in actions that reinforce aspects of English language dominance. These results support Jenkins (2006) assertions that, while there are theoretical arguments to refute ideological beliefs around English, “the belief in native speaker ownership persists among both native and nonnative speakers” (p. 171). NESTs’ enactment of the ideology of English is based on foundational assumptions that seem natural to them, such as the fact that they can assume that their cultural assumptions are the basis for standards in the classroom. They have the capacity to critique this culture within which they are “foreigners” and to take for granted their native-speaker status of English. Much of these assumptions are outward-focused; teachers can focus on their reactions to Korea or their reactions to students, since they can take for granted their authority in the classroom. Students, on the other hand, tend to internalize the ideology of English more; their enactment of the ideology of English can take a greater toll. Students internalize the need to learn English and place a high amount of personal responsibility on themselves for progress. They even critique their own educational culture and make goals to study harder and longer.
The spread of English as the language of globalization has far-reaching consequences. It is not a neutral language that was somehow destined to become the world language (Canagarajah, 1999; Holborow, 1999; Macedo, Dendrinos, & Gounari, 2003; Phillipson, 1992, 2009; Spring, 1998). As countries such as South Korea require citizens to learn English in order to stay globally competitive, more and more opportunities for intercultural interactions arise. To meet the demands for English language education, more foreign instructors are hired (Jenkins, 2006). However, when culturally different teachers and students interact, they do not approach each other on equal grounds on which to form a neutral “middle space” of culture. Instead, they both enter the interaction as members of complex situations, products of history, economics, politics, and culture, in which certain overarching ideologies dictate privileging of different aspects of their situations. The results of this study show how one ideology, the ideology of English, dictates inequalities within the interactions between NESTs and students in South Korea. This ideology provides the framework to view who can dictate norms and who feels the pressure to adapt to whom. To continue to investigate aspects of privilege and power with regard to how global and local settings are reacting to globalization is important. Decisions regarding language policies and language methodology should reflect and consciously address how the existence of language ideology influences current situations.
References


**Notes**

i The words “foreign” and “foreigner” are used quite commonly in South Korea. The term, wae-gook-een, originally came from Chinese roots and is defined as “someone from another country; a person with a foreign nationality” (Harris, 2004, p. 19). NESTs are usually referred to as “foreign teachers.”

ii The larger study included research results from both narrative analysis and ideological analysis conducted on the data. This article only reports results from the ideological critique conducted on the data.

iii An important aspect of this research design was to collect assignments from Korean students in a class taught by a Korean professor. Students were not asked to write their stories and then turn them in initially to a NEST. After the assignment was completed for the Korean professor students were then asked if they wanted to submit their narratives for data analysis.

iv All stories are referenced by number. SS represents “student story.” Later in the article, TS represents “teacher story.” The stories were numbered in order to protect the identity of the participants, not to try to dehumanize the narrators.
A Quantitative Reanalysis of Data on the Structure of L1 and L2 Language Ability in Multitrait-Multimethod Studies

Yo In'nami & Rie Koizumi

Shibaura Institute of Technology & Juntendo University, Japan

Bio Data

Yo In'nami Ph.D. is an Associate Professor of English in the College of Engineering at Shibaura Institute of Technology, Japan. He is interested in meta-analytic inquiry into the variability of effects and the longitudinal measurement of change in language proficiency.

Rie Koizumi Ph.D. is an Associate Professor of English in the Faculty of Medicine at Juntendo University, Japan. She is interested in validating speaking and vocabulary assessment and modeling factor structures of language ability and performance.

Abstract

To gain a better understanding of the structure of L1 and L2 language ability, we conducted an extensive literature search to collect multitrait-multimethod (MTMM) studies. We specifically examined the unitary, correlated, uncorrelated, and higher-order structures of language ability to determine which of these best fit the data across studies. We used confirmatory factor analysis to reanalyze 58 correlation or covariance matrices from 39 studies, among which we retained 17 correlation or covariance matrices from 14 studies for the main analysis. Empirical evidence showed that the unitary and higher-order models best defined both L1 and L2 ability. The support for the unitary model was surprising but consistent with the findings of Davidson (1988). Moderator variable analyses failed to identify clear relationships between the examined models and moderator variables.
Keywords: multitrait-multimethod, language ability structure, confirmatory factor analysis, first language, second language

Introduction

For the past five decades, the nature of language ability has been one of the most widely discussed issues in the field of language learning, particularly in language testing. Lado’s (1961, 1964) seminal work on defining language ability advocated the structural linguistic approach to language, which described language as consisting of several elements or components such as phonemes and morphemes, and regarded language learning as the process of mastering these elements or components. More specifically, one example of the structural approach was to divide language into two types of elements represented in three-by-four matrices: the first type consisted of phonology or orthography, vocabulary, and structure; and the second type of listening, speaking, reading, and writing skills (Fouly, Bachman, & Cziko, 1990).

Obviously, the structural approach focused heavily on grammatical or linguistic aspects of language. This approach was supplanted with an attempt to define language from the wider perspective of communicative competence. Canale and Swain’s (1980) model of communicative language ability included sociolinguistic and strategic competences as well as grammatical competence. Sociolinguistic competence referred to the ability to use language appropriate to any given social context, and strategic competence referred to the ability to use appropriate communication strategies. Canale (1983) further expanded this model to include the dimension of discourse competence, which was the ability to construct cohesive and coherent text.

Finally, a more extended model of communicative language ability was proposed by Bachman (1990) and Bachman and Palmer (1996). In this model, communicative language ability comprised organizational and pragmatic competences.
Organizational competences included grammatical and textual competences, and pragmatic competences included illocutionary and sociolinguistic competences.

Along with the aforementioned models of language ability, which were mainly theory-based rather than data-based, empirical studies have also investigated the structure of language ability. According to Carroll (1993), almost all L1 studies (e.g., Horn, 1988; Jensen, 1970) have focused on investigating the structure of intelligence, a component of which is language comprehension. An exception is an inquiry conducted by Carroll (1993), which directly investigated the structure of L1 ability by exhaustively collecting, reviewing, and reanalyzing correlation matrices, largely through exploratory factor analysis (EFA). The results of the study showed that in almost every case, L1 ability was hierarchical and affected lower-order abilities.

Carroll’s (1993) model is called a higher-order model. Figure 1 (See Appendix) is a schematic representation of the model, in which the rectangles represent the observed variables represented by test scores, the ovals represent language abilities, traits, or latent factors, and the circles represent measurement errors; a single-headed arrow indicates an influence from one variable to another, whereas a double-headed curved arrow—as in Figure 3a, for example—shows a correlation between two variables. Figure 1 has three first-order trait factors (Traits 1 to 3) with a general, higher-order factor (Higher-order trait). It should be noted that although the higher-order structure of language ability remains the same, the number of observed variables and latent factors in Figure 1 may differ from that in Carroll’s (1993) study because this number differed across the models that were reanalyzed. This difference in the number of observed and latent variables also applies to the other models in Figures 2–6 (See Appendix).

In L2 research, a larger number of studies have been conducted on the structure of language ability. Oller (1983) conducted a well-known initial study that performed
principal components analysis (PCA) on a university placement test consisting of composition, vocabulary, grammar, phonology, and dictation or cloze tasks. The results of the analysis showed that L2 ability was unitary in nature, and the same conclusion was obtained across three other different datasets in the same study. Oller argued that this unitary view of language ability, known as the unitary trait hypothesis, could refer to learners’ ability to predict and process sequences of language elements that were constrained by immediate context.

The unitary trait model is presented in Figure 2. It was later rejected because Oller (1983) analyzed his data using PCA, which is known for overestimating the significance of the unitary factor (Carroll, 1983; Farhady, 1983). Scholz, Hendricks, Spurling, Johnson, and Vandenburg (1980) also claimed that their data supported the unitary trait hypothesis of language ability, but when Carroll (1983) used an appropriate method (principal factor analysis) to reanalyze their data, he found evidence of not only a general, higher-order factor but also several additional factors.

The unitary trait hypothesis was also refuted by studies employing more advanced analytical methods such as confirmatory factor analysis (CFA). Bachman and Palmer (1982) adopted the multitrait-multimethod (MTMM) design in the CFA framework in order to test the following three models for their capacity to explain the structure of language ability: a unitary model (see Figure 3a), a higher-order model (Figure 6a), and a correlated model (Figure 4a), which assumes the correlation of multiple abilities. They found the higher-order model (Figure 6a) the most appropriate for defining language ability.

This higher-order model of L2 ability has also been widely accepted by other empirical studies (e.g., Bachman & Palmer, 1982, 1989; Llosa, 2005, 2007; Sawaki, 2007; Sawaki, Stricker, & Oranje, 2009; Shin, S.-K., 2005). For example, Sawaki (2007) investigated the structure of L2 speaking ability using a unitary ability model
(see Figure 3b), a correlated model (Figure 4b), an uncorrelated ability model (Figure 5b), and a higher-order model (Figure 6b). The results favored the higher-order model (Figure 6b) and suggested that L2 speaking ability comprised multiple abilities and an underlying higher-order ability. Her models used the MTMM design and included uncorrelated test method factors, both of which are explained in the Confirmatory Factor Analysis (CFA) and Multitrait-Multimethod (MTMM) Design section below.

In contrast, Bachman and Palmer (1981) and Sang, Schmitz, Vollmer, Baumert, and Roeder (1986) found that a correlated trait model with correlated method factors (see Figure 4a) best described the structure of language ability. Although the debate between the higher-order and correlated models has not yet been resolved, researchers (e.g., Chalhoub-Deville, 2003; Kunnan, 1998; Purpura, 2004) have argued that the structure of L2 ability in both these models is composed of multiple components; that is, the L2 ability structure is multicomponential. Bachman (1998) observed “the emergence of a general consensus that language proficiency consists of a number of distinct but related component abilities” (p. 177), and Douglas (2000) noted, “As has become clear in recent years through empirical studies conducted by language testers and others, language knowledge is multicomponential” (p. 25).

However, this does not explain how these multiple components relate to each other; that is, it does not clarify whether the structure of L2 ability is hierarchical (i.e., has a higher-order factor) or merely consists of multiple correlated abilities without any hierarchical structures (Bachman, 1998; Chalhoub-Deville, 2003; Douglas, 2000; Kunnan 1998). These two ability structures (hierarchical or non-hierarchical) may look similar, but are different and worth investigating. The hierarchical structure suggests that the multiple abilities are separate but closely related, and can be combined into one higher-order ability, whereas the non-hierarchical structure indicates that they are separate but not strongly related enough to assume one higher-order ability.
Ability structure has practical implications for test score reporting. If it is hierarchical and a single higher-order factor underlies the multiple abilities assessed in a test, a single total score should be reported along with separate scores for each ability. However, if it is non-hierarchical, and the abilities are correlated but distinct, only separate scores should be reported for each ability.

In addition, one study casts doubt on the multicomponentiality of L2 ability. Davidson (1988) conducted a statistical analysis of 21 item-level datasets from nine L2 test batteries in order to determine whether they contained the attributes of multicomponential language models. EFA showed that 12 of the 21 datasets displayed a unitary ability factor and that the remaining nine datasets revealed multiple factors (Davidson, 1988). Therefore, Davidson (1988) argued that while L2 ability could be unitary, it is also possible that factor analysis methods are too insensitive to detect multicomponential ability structure, especially when item-level data are employed.

To conclude, L1 studies have shown that a hierarchical model of language ability is the most appropriate model of L1 ability, whereas L2 studies have been unclear as to whether a correlated trait model or a higher-order model, or even a unitary model, most appropriately describes the structure of L2 ability.

Confirmatory Factor Analysis (CFA) and Multitrait-Multimethod (MTMM) Design
As discussed above, the methods commonly used for examining the language ability structure have been exploratory factor analysis (EFA) and confirmatory factor analysis (CFA). CFA is an analytical method for testing the structures of observed variables (test scores in the present study) and unobserved (latent) factors predicted by the theory (see e.g., Brown, 2006). CFA extends EFA in that it helps researchers construct complex models based on theory, compared to the data-driven nature of EFA. CFA
also offers flexible approaches to modeling. One example is the multitrait-multimethod (MTMM) design.

The MTMM is an experimental design aimed to investigate the extent to which test results are due to either learner abilities (more often called traits in the MTMM design) or the test method effects (Campbell & Fiske, 1959). It is devised to measure more than one trait with more than one test method. One example is when the reading and the listening abilities of students are tested, with each ability tested using short-answer as well as self-assessment methods. In this case, four test scores would be derived: two reading test scores assessed by short-answer and self-assessment methods, and two listening test scores assessed by these two methods. The four test scores can be considered to be affected by three factors: trait, method, and error. The MTMM design allows us to empirically separate the trait, the method, and the error factors, whereas non-MTMM designs only allow us to separate the trait and the error factors (Bachman, 1990). Although MTMM models are often unstable (Kenny & Kashy, 1992), we focused on previous studies employing MTMM models in order to examine the ability structure more clearly than would have been possible if we were using a non-MTMM design.

In the MTMM design, a variety of models can be created. Figures 3a to 6c shows examples. For example, in Figure 4a, nine test scores are explained by three trait factors (Traits 1 to 3), three method factors (Methods 1 to 3), and nine errors; the three traits are interrelated to each other, like the three methods are. Figure 4b is relatively similar to Figure 4a, but is different in that the three methods are not correlated. An examination of the fit between the model and the data using statistical fit indices is helpful in deciding which model most aptly explains the data. Researchers select and examine models based on the theory and previous studies. For example, Sawaki (2007) compared four models with uncorrelated test method factors: a unitary ability model
(see Figure 3b), a correlated model (Figure 4b), an uncorrelated ability model (Figure 5b), and a higher-order model (Figure 6b). She found the higher-order model to be the best-fitting one. The four models that she tested and other possible models based on the theory are called competing, alternative, or rival models.

In CFA, raw data are not required; a correlation or covariance matrix will suffice in most cases. A correlation matrix includes the correlations among test scores, whereas a covariance matrix includes the variances of and covariances among test scores; covariances are correlations multiplied by the standard deviations of the two scores. If previous studies report a correlation or covariance matrix, we can use it as data, and construct the models the researchers tested as well as the models they did not test, and examine the model fit. This is one of the most useful features of CFA.

Current Study

Previous studies have provided invaluable information on the appropriate models for defining the structure of language ability; however, there still exist four limitations in the ability to select the most appropriate models. First, all the competing theoretical models, including those originally tested, need to be examined and compared so that the best one is selected. Some studies did not systematically attempt to falsify alternative models of language ability, probably because this was not the focus of their investigations or because the analytic approaches employed did not permit a close investigation of this issue.

Second, it is difficult to judge the adequacy of the models adopted in previous studies because we do not know the extent to which these models are supported by multiple statistical fit indices. Some early models were tested with only one statistic—the chi-square statistic, which fails current standards for model testing (e.g.,
Kline, 2005). The fit between data and these previous models needs to be reanalyzed with the various fit indices that are now available.

Third, despite its invaluable finding related to the unitary trait structure, the large number of variables (i.e., item-level data) analyzed in Davidson’s (1988) study makes it difficult to find more than one factor (as Davidson himself mentioned, based on Bengt Muthen’s personal comment). Thus, we need to conduct a reanalysis with a smaller number of variables, such as those at the subscore level (e.g., a reading section score, not a score of each item).

Fourth, the findings of previous studies are yet to be supported across studies. If the findings are replicated by a large number of studies, we would be able to understand the structure of L1 and L2 language abilities better. Such a cross-study replication approach is rare in language-related studies, but is prevalent in psychological studies (see e.g., Lance, Lambert, Gewin, Lievens, & Conway, 2004).

The current study analyzes previous MTMM studies by comparing L2 competing models from four model families—something that previous language studies have not attempted to do. We use multiple statistical fit indices and subscore level data (as represented in correlation or covariance matrices) to examine the families of models most often deemed appropriate across studies.

We aim to contribute not only to theory building but also to test validation, of which investigation into language ability structure is an essential aspect (Messick, 1996). Test developers usually start designing a test by writing a specification postulating a certain ability structure. They then develop the test based on the specification, administer it, and analyze the test data. If the language ability structure derived from the data is consistent with the one postulated in the test specification, the test is considered to have evidence for the validity of the score-based interpretation. If, however, the structure derived from the data is inconsistent with the one postulated, the
test interpretation based on the test scores is considered invalid, and the test or test specification must be put under scrutiny for revision. The findings of this study will provide implications for test validation.

We reanalyze correlation or covariance matrices from each study to find the best model for each study. Then, we collate these models in order to determine the most commonly adopted model. We also aim to identify characteristics shared across the best models. Our research question is as follows:

Which structure of language ability—unitary, correlated, uncorrelated, or higher-order—is most often supported in L1 and L2 studies?

We understand that a variety of language ability models exist because language tests operationalize diverse ranges of abilities. Carroll (1983) acknowledged that the structure of language ability depends on the characteristics of the tests used, the nature of the sample, and other conditions. However, he also stressed that if all studies conducted proper data analyses, they would obtain similar results. It follows that if we properly reanalyze models of language ability, we will find that some models more aptly capture the structure of language ability than others do.

**Method**

*Data Collection*

*Data identification.* An exhaustive search for MTMM studies concerning language performance was conducted in three ways from July 2007 to January 2008. First, literature was retrieved by searching through databases: Educational Resources Information Center, FirstSearch, Linguistics and Language Behavior Abstracts, PsycINFO, ScienceDirect, and Web of Science. In order to avoid retrieving a large number of irrelevant studies by using a single keyword, we used a combination of
keywords: multitrait, multimethod, convergent, discriminant, divergent, and MTMM. This list of keywords was compiled by the authors based on the keywords and synonyms retrieved from the thesauruses provided in databases, books and articles reviewed, authors’ experiences, and feedback from colleagues. Abstract, title, and article keyword searches were performed and no date range restriction was imposed.

Second, we found MTMM studies concerning language performance by referring to books and journals in the fields of language testing, L1 and L2 language acquisition, and educational measurement (See Appendix Table 1). Different editions of the same book were also referred to. Third, we asked other researchers to suggest studies that were missing from the list we had compiled. To obtain additional relevant material, we inspected the reference lists of all retrieved empirical, theoretical, and review papers and chapters, including theses and dissertations as well as publications in English and other languages, regardless of whether these works were published or unpublished.

Criteria for the inclusion of a study. We identified and retrieved approximately 400 studies. Their titles, abstracts, and study descriptors were scrutinized to determine whether (a) the studies used an MTMM design; (b) they reported the information required to reanalyze an MTMM model (a correlation or covariance matrix); (c) all their observed variables measured language ability and/or knowledge; and (d) their MTMM model had at least three first-order trait factors—a minimum requirement for testing the higher-order model (e.g., Brown, T. A., 2006).

We reiterate that raw data is not necessary to reanalyze MTMM models, as long as a study provides its correlation or covariance matrix. We tried our best to contact the authors of the studies to request the matrix and/or study details when necessary. In addition, we examined a sample of 9.46% (n = 40) of the 400 retrieved
studies independently to ensure that they met all the aforementioned four criteria. Interrater agreement was high at 90%, and the kappa coefficient was high at .80. Disagreement was resolved through discussions. The first author reviewed the remaining studies.

As a result, 79 correlation or covariance matrices from 54 studies were selected, among which 21 correlation or covariance matrices from 15 studies failed to satisfy criteria (c) and (d). These 15 studies are marked with an asterisk in the references. In the end, 58 correlation or covariance matrices from 39 studies satisfied all criteria and provided the data for subsequent analysis.

Analyses

Models tested. Confirmatory factor analysis (CFA) was used to estimate the suitability of 12 models for each MTMM correlation or covariance matrix.¹ We used the same number of traits and the same number of methods specified by the original authors of the studies. The 12 models were based on Widaman (1985) and classified into four family types: unitary trait family, correlated trait family, uncorrelated trait family, and higher-order trait family. For each family, three method structure models were examined: a correlated method model, an uncorrelated method model, and a correlated uniqueness model.

The unitary trait family consists of a unitary trait correlated method model, a unitary trait uncorrelated method model, and a unitary trait correlated uniqueness model. A unitary trait model is based on Oller (1983) and assumes that test performance can be explained by only one trait. In the unitary trait correlated method model, both trait and method are considered, with the methods correlated, as shown in Figure 3a. On the other hand, a model without the correlations between the method factors is a unitary trait uncorrelated method model, as shown in Figure 3b. In a unitary
trait correlated uniqueness model as seen in Figure 3c, method factors are not specified but implied by the correlated errors (i.e., uniquenesses), where the observed variables sharing the same method are associated (e.g., Kenny & Kashy, 1992). A correlated uniqueness model has been used to avoid problems with parameter estimation and convergence in the MTMM design (e.g., Lance, Noble, & Scullen, 2002).

The correlated trait family is based on Bachman and Palmer (1981) and Sang et al. (1986) and includes a correlated trait correlated method model, a correlated trait uncorrelated method model, and a correlated trait correlated uniqueness model. A correlated trait model has multiple correlated trait factors. A correlated trait correlated method model, shown in Figure 4a, consists of trait and method factors correlated with the same type of factors, whereas a correlated trait uncorrelated method model, shown in Figure 4b, assumes correlated traits but uncorrelated methods. A correlated trait correlated uniqueness model is shown in Figure 4c.

The uncorrelated trait family is tested in Sawaki (2007), and is composed of an uncorrelated trait correlated method model, an uncorrelated trait uncorrelated method model, and an uncorrelated trait correlated uniqueness model. The models in this family are different from those in the correlated trait family only in that the trait factors are uncorrelated with each other in the uncorrelated trait model family (see Figures 5a, 5b, and 5c).

The higher-order trait family is based on Bachman and Palmer (1982), for example, and subsumes a higher-order trait correlated method model, a higher-order trait uncorrelated method model, and a higher-order trait correlated uniqueness model. A higher order trait model includes one general higher-order factor and several distinct first-order factors (e.g., Bachman & Palmer, 1982; Carroll, 1993; Llosa, 2005, 2007; Sawaki, 2007; Shin, S.-K., 2005). Figures 6a, 6b, and 6c show three types of
higher-order trait models. Table 2 summarizes the 12 models classified into the four model families.

*Model selection.* To assess the overall fit of the model, each model was evaluated using three criteria: two fit indices (the Comparative Fit Index [CFI] and the Standardized Root Mean square Residual [SRMR]) and either a chi-square difference test or the Consistent Akaike Information Criterion (CAIC). Since the chi-square statistic is sensitive to sample size, a CFI of .90 or above was used as an indicator of an adequate model fit. Since the Root Mean Square Error of Approximation (RMSEA) value is often large in small samples even when the model is correct, an SRMR of .08 or below was used to evaluate model fit (Hu & Bentler, 1999). When the models were nested, that is, when one model can be created from another by imposing constraints on the paths, the chi-square difference test was used to compare the model fit. When the models were not nested or when the chi-square difference test was not significant, the model that had the smaller CAIC value and that met the criteria of CFI and SRMR was selected. Additionally, each model pair was statistically not discriminable when three first-order factors were used in comparing (a) the correlated trait uncorrelated method model (Figure 4b) with the higher-order trait uncorrelated method model (Figure 6b; Rindskopf & Rose, 1988), and (b) the correlated trait correlated uniqueness model (Figure 4c) with the higher-order trait correlated uniqueness model (Figure 6c). An instance of (a) was observed in one matrix (Janssen & De Boeck, 1996), and an instance of (b) in another matrix (Jones & Day, 1997). These two matrices were excluded from further analysis.² They were all L1 data and did not affect L2 data analysis.

After one best-fitting model was selected by analyzing each correlation or covariance matrix, that model was classified into one of the four categories (unitary
trait, correlated trait, uncorrelated trait, or higher-order trait model families) based on
the hypothetical structure of trait. The category with the distinctively large number of
best-fitting models was selected to determine the most plausible model to be used
when defining the L1 and L2 language ability structures. We tallied the number of
model families (e.g., unitary trait model families and correlated trait model families)
and not the number of models themselves because we were interested in trait structures,
not method structures.

Preliminary analyses. To analyze 12 models for each correlation or covariance matrix,
we employed maximum likelihood estimation method with Amos 4.01 (Arbuckle,
1999). We could not examine univariate and multivariate normality, or use robust
estimation method (even when the original studies had used it) such as the
Satorra-Bentler scaled chi-square statistic or bootstrapping estimates of parameters in
(plausibly) nonnormal data (see Kline, 2005, for details), since we had no access to the
raw datasets. One of the factor loadings from each factor was fixed to 1 in order to
derive scale identification, a necessary condition for CFA analysis. While error
variances should exceed zero, we occasionally obtained below-zero error variances.
We decided that the models with a statistically significant negative error variance(s)
would not be used, and that the models with nonsignificant negative error variance(s)
would be considered acceptable and would be reestimated by confining the error
variance(s) to zero, based on Chen, Bollen, Paxton, Curran, and Kirby (2001). When
models did not converge, that is, if models failed to obtain proper solutions within 500
iterations (which was the default value for Amos 4.01), we changed the start value and
reanalyzed the models using a start value of either 0.1, 0.3, or 0.5 for the parameters
(Kline, 2005). Once a model converged and had no statistically significant negative
error variances, all the parameter estimates were checked to ensure they were within
the expected range. To ensure that the analyses were conducted properly, even those models that converged quickly or produced no statistically significant negative error variances were cross-checked with the three start values of 0.1, 0.3, and 0.5 because MTMM models are often unstable (Kenny & Kashy, 1992). Models that did not remain stable across the three different values were excluded.

From the 58 correlation or covariance matrices of 39 studies that remained after the initial stage of data inclusion, 41 matrices from 22 studies were excluded from further analysis. Five matrices from five studies were excluded because they were not positive definite (i.e., their variables correlated too strongly to be analyzed by CFA). One matrix from one study was excluded because it could not obtain positive definite results or identify the model from any of the 12 models tested. Thirty-one matrices from 20 studies were also excluded because their data could not be fit by any of the 12 models. Two matrices from two studies were excluded because their results varied when start values were changed. Another two matrices from two studies were excluded because each model pair could not be statistically discriminated. For studies excluded at this stage, see works marked with a double asterisk in the references. Finally, 17 matrices from 14 studies were retained and included in the current analysis (see Table 2 for these studies; they are marked with a triple asterisk in the References).

Moderator variables coded for each study. Although we aimed to identify the language ability structure that is most frequently observed in L1 and L2 studies, the structure may vary depending on the characteristics of the study, such as learners’ proficiency levels and mother tongue. Such study features are called moderator variables, and are used to examine the relationship between a selected model and its study characteristics. We coded 14 moderator variables of interest related to MTMM design, learner
characteristics, and research methods (see Table 3), based on Davidson (1988), Cooper and Hedges (1994), Sasaki (1996), and Norris and Ortega (2000).

In particular, the specific traits that were measured were coded because Davidson (1988) suggested that testing for similar traits (e.g., reading and listening rather than reading and speaking) tends to create a stronger relationship among traits, thus producing a unitary trait factor. The L2 study context was coded because Sasaki (1996) argued that L2 learners studying in countries where the L2 is commonly spoken are more likely to develop all-round skills, thereby producing a unitary factor. Finally, the reliability coefficients (e.g., Cronbach’s alpha) of observed variables (i.e., test scores) were coded because models tested with highly reliable variables tend to measure a single ability, resulting in the production of one ability factor, as pointed out by Davidson (1988).

Unfortunately, another interesting variable, L2 study length—could not be coded, since either the varied length of the included L2 studies or the lack of sufficient information hampered reliable coding. Likewise, L2 proficiency level could not be coded due to a lack of information on the relative difficulty of proficiency measures.

To code the 14 moderator variables, both authors independently examined a sample of 53% (n = 9) of the 17 correlation or covariance matrices included in the final data collection stage of the study. We recorded an agreement percentage of 90% and a high kappa coefficient of .80. Disagreement was resolved through discussion. The remaining studies were examined by the first author.

Results

Characteristics of MTMM Correlation or Covariance Matrices

We examined the study characteristics of 17 matrices from 14 studies that were retained for the final analyses. Although we do not show this in the tables, we found
that L1 studies have been conducted across a longer period than L2 studies have been. The publication date of L1 studies ranged from 1968 to 1990, while that for L2 studies were between 1982 and 2007. In addition, the sample sizes of L1 studies are much larger and more varied than those of L2 studies are. For L1 studies, mean sample size was 1110.22, sample size ranged from 57 to 4,045, and sample standard deviation (SD) measures 1656.518; for L2, these statistics were 199.64, 50 to 593, and 163.23 respectively.

**L1 and L2 Language Ability Structures**

Using 17 matrices from 14 studies retained for the final analyses, we assessed the overall fit of the model, selected one best-fitting model for each matrix, and classified it into one of the four categories (unitary trait, correlated trait, uncorrelated trait, or higher-order trait model families). Table 4 shows that the most frequently observed best-fitting model family in L1 matrices was the unitary trait model family, which fit approximately half of the correlation or covariance matrices \( k = 3; 50.00\% \), followed by the higher-order trait model family \( k = 2; 33.33\% \), and the correlated trait model family \( k = 1; 16.67\% \). The uncorrelated trait model family did not fit any correlation or covariance matrices, suggesting that there are some relationships among traits in the L1 ability structure. All matrices had a sample size of 100 or more, which is satisfactory for CFA. In sum, the unitary trait model family, the higher-order trait model family, and the correlated trait model family all received empirical support.

Although not shown in Table 4, the three matrices whose data were best explained by the unitary trait model family created statistical problems for the higher-order model family. That is to say, when the models from the higher-order model family were tested with these three matrices, one or more of the following three problems occurred: They did not converge, were not identified, or did not produce...
positive definite results for the matrices (see the discussion in the Analyses section). This is probably because the higher-order model family is statistically more complex to estimate. Since this family included three models (i.e., a higher-order trait correlated method model, higher-order trait uncorrelated method model, and higher-order trait correlated uniqueness model), we tested nine higher-order models for the three correlation or covariance matrices (3 models [Figures 6a, 6b, and 6c] x 3 matrices). Five of the nine models experienced statistical problems and could not be compared with the unitary models that best explained the data for the three matrices.

For L2 studies, the models from the unitary trait model family fit the data (k = 5; 45.45%) most frequently. They were followed by the models from the higher-order model family (k = 4; 36.36%), the correlated trait model family (k = 1; 9.09%), and the uncorrelated trait model family (k = 1; 9.09%). Although there were two matrices with a sample size that was rather too small (99 or below) to be analyzed using CFA, the results remained essentially the same when only matrices with a sample size of 100 or more (k = 9) were considered; among the nine matrices, four models were from the unitary trait model family (44.44%), three from the higher-order trait model family (33.33%), and one each from the correlated trait model family (11.11%) and the uncorrelated trait model family (11.11%). Therefore, the unitary trait model family and the higher-order trait model family were supported more often than the correlated and uncorrelated trait model families. The almost equal frequency of empirical support for the unitary trait model family and the higher-order trait model family suggests that these two model families are plausible structures of L2 ability.

Although this is not shown in Table 4, the five matrices whose data were best explained by the unitary trait model family had statistical problems in the higher-order model family. Since there are three models in this family, we tested 15 higher-order models (3 models x 5 matrices). Eleven of the 15 models experienced statistical
problems, which prevented us from comparing them with the unitary models that best explained the data for the five matrices.

Moderator Variables Related to Language Ability Structure

Model families with the most empirical support for L1 ability structure, as mentioned above, were the unitary trait, the higher-order trait, and the correlated trait model families. Those with the most support for L2 ability structure were the unitary trait and the higher-order trait model families. Since each study possessed various characteristics, we expected detailed analyses to help identify some key moderator variables related to the foregoing model families. Nevertheless, as seen in Tables 5 and 6, the results of these analyses failed to show clear patterns for any moderator variables, including the three aforementioned variables of specific traits measured, L2 study context, and reliability of observed variables. For example, based on Davidson (1988), we expected a unitary trait factor to emerge in studies where learners had been tested for similar traits. However, Table 6 shows that among the three studies in which only speaking was tested, the higher-order trait structure was observed in two L2 studies (Sawaki, 2007; Yamashiro, 2002), whereas the unitary trait structure was seen in only one study (Henning, 1983). The details of the three studies are given below.

Sawaki (2007) used five first-order trait factors (i.e., pronunciation, vocabulary, cohesion, organization, and grammar) and four method factors (i.e., two raters and two role-play tasks) to test 205 L2 adult learners of Spanish in the U.S. Yamashiro (2002) assessed the speech activities of 170 Japanese secondary school students who were learning English in Japan, and modeled three trait factors (i.e., non-verbal delivery, verbal delivery, and organization or purpose) and three method factors (i.e., teacher, peer, and self-ratings), which showed the higher-order trait structure. In contrast, Henning’s (1983) data showed the unitary trait structure. His participants consisted of 50 L2 Arabic adult learners of English in Egypt, and his four trait factors were (a)
accuracy or FSI-like scores, (b) fluency, (c) pronunciation, and (d) grammar, while his three method factors were imitation, completion, and interview tasks.

Although some may argue that the small sample size ($N = 50$) in Henning (1983) affected the resultant structure, this seems unlikely because the higher-order trait structure was also observed in the work of Dandonoli and Henning (1990, French), who used a sample of a similar size ($N = 60$). They tested 60 L2 adult learners of French in the U.S. with four traits (speaking, writing, listening, and reading) and four methods (i.e., two raters and two test formats—multiple-choice and open-ended).

Similarly, as shown in Table 5, we obtained both the unitary and the higher-order trait structures for L1 studies in which similar traits (i.e., only reading) were tested (Marsh & Butler, 1984, Studies 1 and 2). This result is notable in that the trait structures were different even when almost the same research design was used across Studies 1 and 2. Both studies tested L1 English speakers in Australia for five reading trait factors (i.e., comprehension, vocabulary, syllabication, blending, and sound discrimination). One difference between the studies was the participants’ level of schooling; in Study 1, they were kindergarteners, while in Study 2, they were primary school students. The second difference was that Study 1 used Levels One and Two of the Stanford Diagnostic Reading Test (SDRT) and teacher rating as three method factors, whereas Study 2 used Level One of the SDRT and two types of teacher rating.

In contrast, the unitary trait structure was observed even in studies that measured different traits, as seen in Table 6. The three L2 studies that each measured four different abilities (i.e., listening, reading, speaking, and writing) produced three different types of trait structure: unitary (Llosa, 2005, Grade 2), uncorrelated (Llosa, 2005, Grade 3), and higher-order (Dandonoli & Henning, 1990). Interestingly, two of Llosa’s (2005) studies, which tested Grade 2 and Grade 3 learners respectively,
produced different trait structures. For these studies, she used three trait factors (i.e., listening & speaking, reading, and writing) and two method factors (i.e., scores of a standardized test [the California English Language Development Test] and ratings of classroom teachers [English Language Development Classroom Assessment]).

In sum, the diverse trait structures obtained from data with a variety of study characteristics contradict Davidson’s (1988) suggestion that the unitary trait structure is derived when learners are tested for similar traits. On the contrary, the similarity of measured traits may not be related to language ability structure.

Comparison of the Best-fitting Models With Those in the Original Studies
The support for the unitary trait structure in some L1 and L2 studies may be surprising, considering that the literature advocates the multicomponential nature of language ability (e.g., Chalhoub-Deville, 2003). In order to explore the reasons for this divergence between the current study and the literature, we compared our best-fitting models with the ones reported in the original papers.

As indicated in Table 7, the models that were found to be the best fit differed across this study and their original studies in all the cases except those of Stevens and Clauser (n.d., 1996 data) and Yamashiro (2002). One possible reason for this divergence could be that the estimation methods used in the original studies and the current study differed, because different results may be obtained when employing (a) maximum likelihood estimation method, which we used, and (b) robust maximum likelihood method, which some previous studies used for nonnormal data. However, as shown in Tables 5 and 6, the number of studies using the robust method was low (zero out of six L1 studies and four out of 11 L2 studies); some studies (i.e., three L1 studies and one L2 studies) did not report their estimation method. Further, two previous studies (i.e., Stevens & Clauser, n.d., 1994 data; Shin, D., 1999) that had used the
maximum likelihood method produced the unitary trait structure even when we used the same method.

Another likely reason why our results diverged from those obtained in the original studies is that not all the models examined in the current reanalysis were tested in the original studies. For instance, Sawaki (2003) compared a correlated-trait and uncorrelated-method model (see Figure 4b), an uncorrelated-trait and uncorrelated-method model (Figure 5b), and a higher-order trait and uncorrelated-method model (Figure 6b); she reported that the model with the higher-order trait and uncorrelated method structure (Figure 6b) fit the data best. However, our reanalysis showed that the best-fitting model was the one with the unitary trait and correlated uniqueness structure (Figure 3c), which Sawaki had not analyzed.

We understand that some of the models we analyzed or the fit statistics we employed were not recommended or widely used at the time of the original studies. We also understand that some of the studies did not test alternative models because they were not the focus of their research. However, we would encourage researchers to analyze all theoretically possible rival models using better analytical methods whenever this is possible. The confirmatory nature of CFA allows us to test only the most plausible models, but the falsification of alternative models always strengthens one’s findings.

**Discussion and Conclusion**

To gather comprehensive empirical evidence on the structures of L1 and L2 language abilities, we used CFA to reanalyze 17 MTMM correlation or covariance matrices taken from 14 studies. We investigated the four structures of language ability—unitary, correlated, uncorrelated, and higher-order—to discover the one that is most often
supported in L1 and L2 studies. The results showed that the likeliest contenders for L1 language ability structure were the unitary, higher-order, and correlated structures, and the most probable candidates for L2 ability structure were the unitary and higher-order structures. The unitary and higher-order ability structures were consistently plausible across both L1 and L2 studies. The moderator variable analyses failed to identify clear relationships between best-fitting structures and moderator variables.

Comparisons With Previous Studies
The results for the most frequently occurring models that best fit the data are partially consistent with the literature. The support for the higher-order trait model family indicates that both L1 and L2 language abilities can be best characterized as consisting of a hierarchical structure where a general, higher-order ability influences lower-order abilities. This result accords with the most of the literature, such as Carroll’s (1993) study of L1 ability, and Sawaki et al.’s (2009) and S.-K. Shin’s (2005) works on L2 ability. Since the higher-order trait model family was more likely to have statistical problems than the other family models were (see the Results section), but frequently emerged as the best-fitting models in the current study, it is highly possible that the higher-order structure accurately described L1 and L2 language ability.

However, the almost equal degree of support for the unitary trait model family, which assumes that language ability consists of a single factor, supports Davidson’s (1988) conclusions but does not concur with most of the literature. This finding is surprising but may be tenable because we used CFA instead of PCA, which Oller (1983) was criticized for using when he argued that L2 ability was unitary in nature. Further, as we discussed in the Current Study section, we expanded on the work of previous studies by comparing four alternative model families, judging models in terms of various fit indices, analyzing data at the subscore level (as represented in
correlation or covariance matrices), and collecting a wide range of available correlation or covariance matrices. Additionally, it should be noted that the current results were derived from previous MTMM data, which had usually been obtained from tests designed to operationalize multiple traits. Thus, it is reasonable to expect to find underlying multiple traits in MTMM data, but our findings often showed the opposite.

Therefore, our findings of support for the unitary trait model in some matrices are not due to statistical artifacts, as Oller’s (1983) were. The finding that the unitary trait model family provided the best fit in studies with various study characteristics, even when these studies assessed different traits, could also strengthen the evidence for the empirical existence of the unitary language structure.

Unitary trait models are empirically as likely as higher-order models are, and both can provide a more accurate explanation of L1 and L2 ability structure. However, the unitary language ability structure does not necessarily mean that it is a single, simplistic construct. Carroll (1983), Davidson (1988), and Henning (1992) stressed that most test items require more than one ability for successful completion, even when the empirical structure of the data suggests that only one ability is involved. Further, Henning (1992) distinguished between the psychometric (empirically based) dimensionality and psychological (theory-based) dimensionality of data structure, and argued that psychometric unidimensionality can be present even in a test measuring a variety of skills. Reckase, Ackerman, and Carlson (1988) demonstrated this using both simulated and real data, and reported that the detection of a unitary trait factor means that test questions “measure the same composite of abilities, rather than only a single ability” (p. 202). Thus, while we have shown that the unitary ability model is very often the best model for both the L1 and the L2, we argue that this ability consists of various subcomponents. In Henning’s (1992) terms, we interpret the unitary trait
structure as reflecting psychometric unidimensionality, behind which psychological multidimensionality can be assumed.

It should be noted that this view of language ability as unitary but consisting of various subcomponents reveals a gap between theory and methodology. Although such a hypothesis about language ability can be constructed, empirical investigation may not be well equipped enough to test it. In other words, CFA or factor analysis in general may not be sensitive enough to detect multiple subcomponents of language. According to Davidson (1988), “factor analysis … cannot adequately model that complexity” (p. 65) of language ability, and it is important to search for other data analysis techniques.

Implications
This study has four implications. The first one is for test developers, regarding test score reporting. The current results for obtaining the unitary and the higher-order ability structure in L1 and L2 studies support the use of a single total score. The higher-order ability structure also suggests the appropriateness of reporting separate scores for each ability.

The second implication is for teachers. The present study provides possible predictions about students’ test scores and performances. For both the unitary and the higher-order abilities, we expect to have a single underlying ability that explains multiple abilities and test scores. It follows that in a test of reading and speaking, for example, those who have obtained a high reading score tend to receive a high/moderate speaking score. Score fluctuations among the abilities are likely to occur because even in the unitary structure, some tests or test sections are more highly related to the single underlying ability than other tests. Thus, teachers need not worry much about minor gaps in learners’ test scores or attempt to rectify all mistakes that students make, unless the gaps are noticeably wide. Instead, instruction should be designed to enhance
students’ overall ability. When there are large gaps in the scores or performances between tests or between test sections, teachers may need to provide remedial classes to strengthen the low-performing skills.

The third implication is for test developers and users. Since we have found a gap between the theoretical and the empirical structures of language ability, we may need to consider how this gap relates to test validation. Normally, researchers are able to hypothesize and empirically defend a test structure when they provide positive evidence supporting the validity of their claim. Conversely, they are unable to defend a test structure when they obtain negative evidence questioning the validity of their claim. However, when investigating the factor structure of a language test, postulating a model with multiple components but obtaining a model with a unitary structure may not necessarily be negative evidence or a threat to the validity of the hypothesis. Rather, we would argue that such evidence can neither be considered positive nor negative because the methodology may not have been able to capture the complexity of the multicomponentiality of the language ability being tested. Therefore, we submit that this “neutral” evidence does not contribute to a test validation process. In this case, one would need to find other types of positive evidence to back up one’s claim.

A fourth implication is methodological. Although we did not expect the unitary trait model to be one of the most frequently supported L2 models, this finding may be tenable, given our attempts to expand on the work of previous studies. These attempts include: (a) comparison of all theoretically plausible models; (b) the use of appropriate data analysis methods (CFA and fit indices); (c) analysis at the subscore level; and (d) the use of a relatively large collection of MTMM studies for reanalysis. Our reanalysis implies that a further investigation of the unitary trait model is warranted, even though support for it was refuted in previous studies (e.g., Bachman & Palmer, 1982). Likewise, it suggests that comprehensive compilation, reexamination, and reanalysis of
previous studies, in the fashion of the current study, are crucial for accumulating strong evidence in the field.

Limitations and Future Research

This study has two main limitations. First, we retained a small number of correlation or covariance matrices for reanalysis. We reiterate that among the 58 correlation or covariance matrices that were retained at the preliminary stage of data inclusion, only 17 produced models that explained the data well. Second, testing of the higher-order trait structure faced statistical problems, which may have resulted in adoption of a smaller number of higher-order trait family models. These problems may have been due to the complexity that the higher-order trait family models create for model estimation. These two limitations are not surprising because MTMM is a widely used yet often difficult procedure for obtaining statistically admissible (i.e., acceptable) solutions, and because similar problems are often seen elsewhere (see, e.g., Lance et al., 2004).

In order to overcome these two limitations, more MTMM studies need to be conducted and included in future reanalyses. This will increase the number of studies that can be retained for reanalysis and can be potentially used to overcome the statistical problems encountered in this study. The availability of a larger number of studies would enable researchers to conduct more rigorous moderator variable analyses that would be able to detect contextual factors related to ability structure. In addition, non-MTMM studies (see e.g., Sasaki, 1996; Shin, S.-K., 2005) should be added to the current analysis in order to further the examination of language ability structure. The inclusion of more MTMM and non-MTMM studies would enable researchers to examine the generalizability of the current results, and would provide stronger
evidence toward the nature and structure of L1 and L2 language ability, which has long been a topic of interest in the area of language learning.

Notes
1. Although it is possible to conduct EFA diagnostics using EQS, we did not do so because we hypothesized four types of language ability structure based on the literature, making our approach to data analysis theory-driven or confirmatory rather than data-driven or exploratory. If the number of factors extracted by CFA had been suspect, our fit indices should have rendered poor values. Since we removed poorly fitting models from our analyses, we believe our findings are stable even without the use of EFA diagnostics. We also did not use the scree plot to determine the number of factors because it is too subjective (see Fabrigar, Wegener, MacCallum, & Strahan, 1999).

2. If the two models represented by these two matrices are included into the correlated trait model family, the family has three (1 + 2) models in total (see Table 4). If they are included into the higher-order trait model family, the family has four (2 + 2) models in total. In the latter case, the higher-order model would occur more frequently than the unitary trait model family, which has three models, and would be a more plausible model of L1 ability structure. This treatment would change the order of selected models in terms of frequency, but would not change the fact that the selection of trait models varies across studies.

Acknowledgements
We would like to thank Fred Davidson, Takayuki Nakanishi, the journal editors, and the anonymous reviewers for their valuable comments on earlier versions of this paper.
This research was supported by Grants-in-Aid for Scientific Research (No. 20820020) provided by the Japan Society for the Promotion of Science.

References

When the same correlation or covariance matrix appeared in more than one reference, the reference with the broader availability (e.g., the reference that was published, rather than the one that was unpublished) was listed. References marked with an asterisk (*) failed to meet two of the four criteria (see the Method section). Those marked with a double asterisk (**) underwent reanalysis but did not provide a stable, best-fitting model, whereas those marked with a triple asterisk (***) provided a stable, best-fitting model and contributed to the main analysis.


Tables

Table 1 List of Books and Journal Used for the Literature Search

Examples of the language testing books

Examples of books on L1 and L2 language acquisition
Examples of books on educational measurement

Journals
Table 2 Classification of 12 Models That the Current Study Used

<table>
<thead>
<tr>
<th>12 models</th>
<th>Figure</th>
<th>Model family</th>
<th>Breakdown of reanalyzed L1 study</th>
<th>Breakdown of reanalyzed L2 study</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unitary trait correlated method</td>
<td>3a</td>
<td>Unitary trait</td>
<td>Stevens &amp; Clauser (n.d., 1994 data)</td>
<td>--</td>
</tr>
<tr>
<td>Unitary trait uncorrelated method</td>
<td>3b</td>
<td>Unitary trait</td>
<td>Marsh &amp; Butler (1984, Study 1)</td>
<td>Henning (1983), Llosa (2005, Grade 2)</td>
</tr>
<tr>
<td>Correlated trait correlated method</td>
<td>4a</td>
<td>Correlated trait</td>
<td>--</td>
<td>Chiang (1999)</td>
</tr>
<tr>
<td>Correlated trait uncorrelated method</td>
<td>4b</td>
<td>Correlated trait</td>
<td>Stevens &amp; Clauser (n.d., 1996 data)</td>
<td>--</td>
</tr>
<tr>
<td>Correlated trait correlated uniqueness</td>
<td>4c</td>
<td>Correlated trait</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Uncorrelated trait correlated method</td>
<td>5a</td>
<td>Uncorrelated trait</td>
<td>--</td>
<td>Llosa (2005, Grade 3)</td>
</tr>
<tr>
<td>Uncorrelated trait uncorrelated method</td>
<td>5b</td>
<td>Uncorrelated trait</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Uncorrelated trait correlated uniqueness</td>
<td>5c</td>
<td>Uncorrelated trait</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Higher-order trait correlated method</td>
<td>6a</td>
<td>Higher-order trait</td>
<td>--</td>
<td>Sawaki (2007)</td>
</tr>
<tr>
<td>Higher-order trait uncorrelated method</td>
<td>6b</td>
<td>Higher-order trait</td>
<td>Quellmalz, Capell, &amp; Chou (1982)</td>
<td></td>
</tr>
</tbody>
</table>
## Table 3 Moderator Variables

<table>
<thead>
<tr>
<th>Variable</th>
<th>Definition and example</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>MTMM-related characteristics</strong></td>
<td></td>
</tr>
<tr>
<td>Trait by method matrix size</td>
<td>e.g., 3 by 2 (3 traits tested using 2 methods)</td>
</tr>
<tr>
<td>Trait measured</td>
<td></td>
</tr>
<tr>
<td>Specific trait(s)</td>
<td>e.g., Grammar and vocabulary</td>
</tr>
<tr>
<td>Method</td>
<td></td>
</tr>
<tr>
<td>Methods type</td>
<td>e.g., Multiple-choice, self-rating, and teacher rating</td>
</tr>
<tr>
<td>Method difference</td>
<td>Whether the test methods used were</td>
</tr>
<tr>
<td></td>
<td>(a) similar (e.g., three multiple-choice methods, as in Bachman &amp; Palmer, 1989; three raters, as in Chiang, 1999),</td>
</tr>
<tr>
<td></td>
<td>(b) different (e.g., one multiple-choice and one open-ended methods, as in Tindal &amp; Nolet, 1990), or</td>
</tr>
<tr>
<td></td>
<td>(c) mixed (partially overlapping methods; e.g., two tasks and two raters, as in Sawaki, 2007)</td>
</tr>
<tr>
<td><strong>Learner characteristics</strong></td>
<td></td>
</tr>
<tr>
<td>Learners’ age</td>
<td>e.g., Primary, secondary, or adult</td>
</tr>
<tr>
<td>L1</td>
<td>e.g., English and Welsh</td>
</tr>
<tr>
<td>Indo-European or not</td>
<td>Indo-European or non-Indo-European</td>
</tr>
<tr>
<td>Learners’ L2</td>
<td>e.g., French</td>
</tr>
<tr>
<td>L2 study context</td>
<td>L2 is used in that study context as a common language or foreign language</td>
</tr>
<tr>
<td><strong>Research methods</strong></td>
<td></td>
</tr>
<tr>
<td>Reliability</td>
<td>Cronbach’s alpha and Kuder-Richardson 20</td>
</tr>
<tr>
<td>Publication status</td>
<td>Published or unpublished (e.g., dissertations, presentation handouts)</td>
</tr>
<tr>
<td>Sample size</td>
<td>99 or below, or 100 or above</td>
</tr>
<tr>
<td>Analytical method*</td>
<td>Correlation or confirmatory factor analysis (CFA)</td>
</tr>
<tr>
<td>Estimation method for CFA</td>
<td>Maximum likelihood (ML) or robust ML</td>
</tr>
</tbody>
</table>

*Note.* There were no matrices derived from studies using exploratory factor analysis.
Table 4 Best-Fitting Models for L1 and L2

<table>
<thead>
<tr>
<th></th>
<th>Unitary trait family</th>
<th>Correlated trait family</th>
<th>Uncorrelated trait family</th>
<th>Higher-order trait family</th>
<th>Grand total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>total</td>
<td>total</td>
<td>total</td>
<td>total</td>
<td></td>
</tr>
<tr>
<td>L1 (all)</td>
<td>N</td>
<td>4646</td>
<td>3933</td>
<td>--</td>
<td>298</td>
</tr>
<tr>
<td></td>
<td>k</td>
<td>3</td>
<td>1</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>%</td>
<td>50.00</td>
<td>16.67</td>
<td>0.00</td>
<td>33.33</td>
</tr>
<tr>
<td>L2 (all)</td>
<td>N</td>
<td>880</td>
<td>172</td>
<td>593</td>
<td>551</td>
</tr>
<tr>
<td></td>
<td>k</td>
<td>5</td>
<td>1</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>%</td>
<td>45.45</td>
<td>9.09</td>
<td>9.09</td>
<td>36.36</td>
</tr>
<tr>
<td>L2 (N of 100 or above)</td>
<td>N</td>
<td>830</td>
<td>172</td>
<td>593</td>
<td>491</td>
</tr>
<tr>
<td></td>
<td>k</td>
<td>4</td>
<td>1</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>%</td>
<td>44.44</td>
<td>11.11</td>
<td>11.11</td>
<td>33.33</td>
</tr>
</tbody>
</table>

Note. N = sample size (i.e., the total number of test-takers). k = the number of correlation or covariance matrices. aAll models in L1 studies had sample sizes of 100 or more. bResults are classified by sample size because small sample sizes often produce unstable results.

Table 5 Moderator Variable Analysis for L1

<table>
<thead>
<tr>
<th>Variable</th>
<th>Unitary (k = 3)</th>
<th>Correlated (k = 1)</th>
<th>Higher-order (k = 2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trait by method</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3 by 4</td>
<td>1 (Ward, 1982)</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>4 by 4</td>
<td>--</td>
<td>--</td>
<td>1 (Quellmalz et al., 1982)</td>
</tr>
<tr>
<td>5 by 3</td>
<td>1 (Marsh &amp; Butler, 1984, Study 1)</td>
<td>--</td>
<td>1 (Marsh &amp; Butler, 1984, Study 2)</td>
</tr>
<tr>
<td>Specific trait</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R</td>
<td>1 (Marsh &amp; Butler, 1984, Study 1)</td>
<td>--</td>
<td>1 (Marsh &amp; Butler, 1984, Study 2)</td>
</tr>
<tr>
<td>V</td>
<td>1 (Ward, 1982)</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>---------</td>
<td>----------------</td>
<td>----</td>
<td>----</td>
</tr>
<tr>
<td>W</td>
<td>--</td>
<td>--</td>
<td>1 (Quellmalz et al., 1982)</td>
</tr>
</tbody>
</table>

**Method type**

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>MC, short</td>
<td>1 (Ward, 1982)</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>MC, teacher</td>
<td>1 (Marsh &amp; Butler, 1984, Study 1)</td>
<td>--</td>
<td>1 (Marsh &amp; Butler, 1984, Study 2)</td>
</tr>
</tbody>
</table>

**Method difference**

<table>
<thead>
<tr>
<th>Mixed</th>
<th>2 (Marsh &amp; Butler, 1984, Study 1, Ward, 1982)</th>
<th>--</th>
<th>2b</th>
</tr>
</thead>
</table>

**Age**

<table>
<thead>
<tr>
<th>Kindergarten</th>
<th>1 (Marsh &amp; Butler, 1984, Study 1)</th>
<th>--</th>
<th>--</th>
</tr>
</thead>
<tbody>
<tr>
<td>Secondary</td>
<td>--</td>
<td>--</td>
<td>1 (Quellmalz et al., 1982)</td>
</tr>
<tr>
<td>Adult</td>
<td>1 (Ward, 1982)</td>
<td>--</td>
<td>--</td>
</tr>
</tbody>
</table>

**Reliability**

| .672c (Ward, 1982) | -- | 1 (.600)c (Quellmalz et al., 1982) |

**Publication status**

<table>
<thead>
<tr>
<th>Published</th>
<th>2 (Marsh &amp; Butler, 1984, Study 1, Ward, 1982)</th>
<th>--</th>
<th>2b</th>
</tr>
</thead>
</table>

**Analytical method**

<table>
<thead>
<tr>
<th>Correlation</th>
<th>1 (Ward, 1982)</th>
<th>--</th>
<th>--</th>
</tr>
</thead>
</table>

**Estimation method**


**Note.** In some cases, the number count reported here does not correspond to the one in Table 4 because information on moderator variables was not always reported in the original studies. G = grammar; M = mechanics; R = reading; V = vocabulary; W = writing; short = short-answer; teacher = teacher rating.  

a All matrices had three moderator variables in common (i.e., the L1 of the test-takers was English, Indo-European language, the sample size was 100 or more). (Marsh & Butler, 1984, Study 2; Quellmalz et al., 1982). b The value in parentheses indicates the grand mean (i.e., the mean of the means) of the reliability coefficients of the observed variables.
<table>
<thead>
<tr>
<th>Trait by method</th>
<th>Unitary ($k = 5$)</th>
<th>Correlated ($k = 1$)</th>
<th>Uncorrelated ($k = 1$)</th>
<th>Higher-order ($k = 4$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>3 by 2</td>
<td>1 (Llosa, 2005, Grade 2)</td>
<td>--</td>
<td>1 (Llosa, 2005, Grade 3)</td>
<td>--</td>
</tr>
<tr>
<td>3 by 3</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>2 (Bachman &amp; Palmer, 1989; Yamashiro, 2002)</td>
</tr>
<tr>
<td>3 by 4</td>
<td>2 (Bachman &amp; Palmer, 1982; Sawaki, 2003)</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>4 by 3</td>
<td>2 (Henning, 1983; Shin, 1999)</td>
<td>1 (Chiang, 1999)</td>
<td>--</td>
<td>1 (Dandonoli &amp; Henning, 1990)</td>
</tr>
<tr>
<td>5 by 4</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>1 (Sawaki, 2007)</td>
</tr>
<tr>
<td>Specific trait</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L</td>
<td>1 (Shin, 1999)</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>L, R, S, W</td>
<td>1 (Llosa, 2005, Grade 2)</td>
<td>--</td>
<td>1 (Llosa, 2005, Grade 3)</td>
<td>1 (Dandonoli &amp; Henning, 1990)</td>
</tr>
<tr>
<td>R</td>
<td>1 (Sawaki, 2003)</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>S</td>
<td>1 (Henning, 1983)</td>
<td>--</td>
<td>--</td>
<td>2 (Sawaki, 2007; Yamashiro, 2002)</td>
</tr>
<tr>
<td>W</td>
<td>--</td>
<td>--</td>
<td>1 (Chiang, 1999)</td>
<td>--</td>
</tr>
<tr>
<td>Method type</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MC, gap, short</td>
<td>1 (Shin, 1999)</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>MC, inter, self, W</td>
<td>1 (Bachman &amp; Palmer, 1982)</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>MC, rater, short</td>
<td>1 (Henning, 1983)</td>
<td>--</td>
<td>--</td>
<td>1 (Dandonoli &amp; Henning, 1990)</td>
</tr>
<tr>
<td>Imi, comple, inter</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Rater</td>
<td>--</td>
<td>1 (Chiang, 1999)</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Rater, role-play</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>1 (Sawaki, 2007)</td>
</tr>
<tr>
<td>Recall, summary</td>
<td>1 (Sawaki, 2003)</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Self-rating</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>1 (Bachman &amp; Palmer, 1989)</td>
</tr>
<tr>
<td>Self, peer, teacher</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>1 (Yamashiro, 2002)</td>
</tr>
<tr>
<td>Standard, teacher</td>
<td>1 (Llosa, 2005, Grade 2)</td>
<td>--</td>
<td>1 (Llosa, 2005, Grade 3)</td>
<td>--</td>
</tr>
<tr>
<td>Method difference</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Similar</td>
<td>--</td>
<td>1 (Chiang, 1999)</td>
<td>--</td>
<td>1 (Bachman &amp; Palmer, 1989)</td>
</tr>
<tr>
<td>Mixed</td>
<td>1 (Sawaki, 2003)</td>
<td>--</td>
<td>--</td>
<td>2 (Dandonoli &amp; Henning, 1990; Sawaki, 2007)</td>
</tr>
<tr>
<td>Different</td>
<td>4 (Bachman &amp; Palmer, 1982; Henning, 1983; Llosa, 2005, Grade 2; Shin, 1999)</td>
<td>--</td>
<td>1 (Llosa, 2005, Grade 3)</td>
<td>1 (Yamashiro, 2002)</td>
</tr>
<tr>
<td>Age</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Primary</td>
<td>1 (Llosa, 2005, Grade 2)</td>
<td>--</td>
<td>1 (Llosa, 2005, Grade 3)</td>
<td>--</td>
</tr>
<tr>
<td>Secondary</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>1 (Yamashiro, 2002)</td>
</tr>
</tbody>
</table>

(Continues)
<table>
<thead>
<tr>
<th>Variable</th>
<th>Unitary ($k = 5$)</th>
<th>Correlated ($k = 1$)</th>
<th>Uncorrelated ($k = 1$)</th>
<th>Higher-order ($k = 4$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>L1 [Indo-European or not]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Arabic [non-Indo]</td>
<td>1 (Henning, 1983)</td>
<td>--</td>
<td>--</td>
<td></td>
</tr>
<tr>
<td>English [Indo]</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td></td>
</tr>
<tr>
<td>Japanese [non-Indo]</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td></td>
</tr>
<tr>
<td>Learners’ L2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>English</td>
<td>4 (Bachman &amp; Palmer, 1982; Henning, 1983; Llosa, 2005, Grade 2; Shin, 1999)</td>
<td>--</td>
<td>1 (Llosa, 2005, Grade 3)</td>
<td>2 (Bachman &amp; Palmer, 1989; Yamashiro, 2002)</td>
</tr>
<tr>
<td>French</td>
<td>--</td>
<td>1 (Chiang, 1999)</td>
<td>--</td>
<td>1 (Dandonoli &amp; Henning, 1990)</td>
</tr>
<tr>
<td>Japanese</td>
<td>1 (Sawaki, 2003)</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Spanish</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>1 (Sawaki, 2007)</td>
</tr>
<tr>
<td>L2 study context</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Common</td>
<td>3 (Bachman &amp; Palmer, 1982; Llosa, 2005, Grade 2; Shin, 1999)</td>
<td>--</td>
<td>1 (Llosa, 2005, Grade 3)</td>
<td>1 (Bachman &amp; Palmer, 1989)</td>
</tr>
<tr>
<td>Foreign</td>
<td>2 (Henning, 1983; Sawaki, 2003)</td>
<td>1 (Chiang, 1999) 1 1 (Llosa, 2005, Grade 3)</td>
<td>3 (Dandonoli &amp; Henning, 1990; Sawaki, 2007)</td>
<td></td>
</tr>
<tr>
<td>Reliability</td>
<td>4 (.786)* (Bachman &amp; Palmer, 1982; Henning, 1983; Llosa, 2005, Grade 2; Shin, 1999)</td>
<td>--</td>
<td>1 (.925)* (Llosa, 2005, Grade 3)</td>
<td>3 (.841)* (Bachman &amp; Palmer, 1989; Dandonoli &amp; Henning, 1990; Yamashiro, 2002)</td>
</tr>
<tr>
<td>Publication status</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Published</td>
<td>2 (Bachman &amp; Palmer, 1982; Henning, 1983)</td>
<td>1 (Chiang, 1999) 1</td>
<td>--</td>
<td>3 (Bachman &amp; Palmer, 1989; Dandonoli &amp; Henning, 1990; Sawaki, 2007)</td>
</tr>
<tr>
<td>Unpublished</td>
<td>3 (Llosa, 2005, Grade 2; Sawaki, 2003; Shin, 1999)</td>
<td>--</td>
<td>1 (Llosa, 2005, Grade 3)</td>
<td>1 (Yamashiro, 2002)</td>
</tr>
<tr>
<td>Sample size</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>99 or below</td>
<td>1 (Henning, 1983)</td>
<td>--</td>
<td>--</td>
<td>1 (Dandonoli &amp; Henning, 1990)</td>
</tr>
<tr>
<td>100 or above</td>
<td>4 (Bachman &amp; Palmer, 1982; Llosa, 2005, Grade 2; Sawaki, 2003; Shin, 1999)</td>
<td>1 (Chiang, 1999) 1 1 (Llosa, 2005, Grade 3)</td>
<td>3 (Bachman &amp; Palmer, 1989; Sawaki, 2007; Yamashiro, 2002)</td>
<td></td>
</tr>
<tr>
<td>Analytical method</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Correlation</td>
<td>1 (Henning, 1983)</td>
<td>--</td>
<td>--</td>
<td>1 (Dandonoli &amp; Henning, 1990)</td>
</tr>
<tr>
<td>CFA</td>
<td>4 (Bachman &amp; Palmer, 1982; Llosa, 2005, Grade 2; Sawaki, 2003; Shin, 1999)</td>
<td>1 (Chiang, 1999) 1 1 (Llosa, 2005, Grade 3)</td>
<td>3 (Bachman &amp; Palmer, 1989; Sawaki, 2007; Yamashiro, 2002)</td>
<td></td>
</tr>
<tr>
<td>Estimation method</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ML</td>
<td>1 (Shin, 1999)</td>
<td>--</td>
<td>--</td>
<td>2 (Bachman &amp; Palmer, 1989; Yamashiro, 2002)</td>
</tr>
<tr>
<td>Robust ML</td>
<td>2 (Llosa, 2005, Grade 2; Sawaki, 2003)</td>
<td>--</td>
<td>1 (Llosa, 2005, Grade 3)</td>
<td>1 (Sawaki, 2007)</td>
</tr>
</tbody>
</table>

*Note. See Table 5. Note. *The value in parentheses indicates the grand mean (i.e., the mean of the means) of the reliability coefficients of the observed variables. P = pragmatic competence; Socio = sociolinguistic competence; L = listening; S = speaking; gap = gap-filling; inter = interview; self = self-rating; TF = True-false; trans = translation; peer = peer rating.*
Table 7 Comparison of the Best-fitting Models With Those in the Original Studies

<table>
<thead>
<tr>
<th>Best-fitting models in our reanalysis</th>
<th>L1 study</th>
<th>Best-fitting model in the original study</th>
<th>L2 study</th>
<th>Best-fitting model in the original study</th>
</tr>
</thead>
<tbody>
<tr>
<td>Correlated trait correlated method</td>
<td>--</td>
<td>--</td>
<td>Chiang (1999)</td>
<td>*</td>
</tr>
<tr>
<td>Correlated trait uncorrelated method</td>
<td>Stevens &amp; Clauser (n.d., 1996 data)</td>
<td>Correlated trait uncorrelated method</td>
<td>--</td>
<td></td>
</tr>
<tr>
<td>Uncorrelated trait correlated method</td>
<td>--</td>
<td>Llosa (2005, Grade 3)</td>
<td>Higher-order trait uncorrelated method</td>
<td></td>
</tr>
<tr>
<td>Higher-order trait uncorrelated method</td>
<td>Quellmalz et al. (1982)</td>
<td>Correlated trait uncorrelated method</td>
<td>--</td>
<td></td>
</tr>
</tbody>
</table>

Note. *One unitary (first-order) trait factor, two correlated (first-order) trait factors, and two uncorrelated method factors all appeared in one model. *This study originally used a correlational analysis technique and produced no best-fitting model.
Figure 1 Higher-order ability model

Figure 2 Unitary trait model

Figures Captions
Figure 3a. Unitary trait correlated method model.

Figure 3b. Unitary trait uncorrelated model.

Figure 3c. Unitary trait correlated uniqueness model.

Figure 3 Unitary trait model family

Figure 4a. Correlated trait correlated method model.

Figure 4b. Correlated trait uncorrelated method model.

Figure 4c. Correlated trait correlated uniqueness model.

Figure 4 Correlated trait model family
Figure 5a. Uncorrelated trait correlated method model.

Figure 5b. Uncorrelated trait uncorrelated method model.

Figure 5c. Uncorrelated trait correlated uniqueness model.

Figure 5 Uncorrelated trait model family

Figure 6a. Higher-order trait correlated method model.

Figure 6b. Higher-order trait uncorrelated method model.

Figure 6c. Higher-order trait correlated uniqueness model.

Figure 6 Higher-order trait model family

265
Word Knowledge: Aspects, Viewpoints and Performances

Zahra Fotovatnia and Alireza Barouni Ebrahimi

English Department, Najafabad Branch, Islamic Azad University, Najafabad, Iran

Bio Data
Zahra Fotovatnia is an assistant professor at the English department of Islamic Azad University, Najafabad Branch, Iran. Her areas of interest are psycholinguistics and pedagogical phonetics and phonology. She has published and presented papers in national and international journals and conferences.

Alireza Barouni Ebrahimi is an MA holder in TEFL from Islamic Azad University, Najafabad Branch, Iran. His areas of interest are applied linguistics, methodology and vocabulary learning and teaching. He is an English teacher in Sadr Institute of Higher Education. He has published and presented papers in international and national journals and conferences.

Abstract
This study collected learner and teacher beliefs about depth of vocabulary knowledge in L1 and L2 to see if and to what extent these aspects were taken as important, and whether there was any relationship between the ratings of these aspects in L2 and learners’ actual performance. 42 TEFL graduate students and teachers, 55 Persian Literature graduate students and teachers, and 20 native TESL teachers received the word-knowledge questionnaire based on Nation’s table of word knowledge (2001) in L1 and L2. Results show that the participants performed differently on some of the 18 items of the questionnaires and the same on others, that the participants did not perform on the reception and production tests based on what they rated as important regarding aspects of word knowledge, and that Nation's table was an important criterion for word knowledge in L1 and L2. The findings might help learners and teachers more easily recognize possible gaps in their depth of vocabulary knowledge and their beliefs about it in accordance with a sound theoretical framework.

Keywords: vocabulary knowledge, beliefs, performances on word

Introduction
What is involved in knowing a word has always been a matter of concern, for vocabulary acquisition is essential in language classes and is important in the development of the four skills of a foreign or a second language. Scholars entitle this involvement ‘depth of vocabulary knowledge’ (Mukarto, 2005; Nassaji, 2004; Shen, 2008). Depth of vocabulary knowledge refers to the quality of knowing words i.e., the quality of lexical knowledge (Meara, 1996; Read, 2000).
Traditionally the knowledge of a word refers to the meaning and form of that word. Schmitt (2008, p.333) indicates that "many teachers and learners consider a word ‘learned’ if the spoken/written form and meaning are known". This matter has obviously been observed in a foreign language setting, when an L2 word is matched with an L1 translation. For example, an Iranian English language learner claims he knows the word "book" as soon as he understands that the word "book" equals his mother tongue word /keta:b/. However, does it really mean he knows the word when he only knows the meaning of that word?

Miller (1999) indicates certainly when we know a word, we know its meaning, but there is more to word knowledge than meaning. A lot of scholars (Bogaards, 2000; Carroll, 2008; Henrikson, 1995; Nation, 1990, 1994, 2001; Richards, 1976) consider some variable factors involved in knowing a word. For example, Bogaards (2000) mentions form, meaning, morphology, syntax, collocates, and discourse as the aspects of vocabulary knowledge. In the same vein, Carroll (2008) divides word knowledge into phonological, syntactic, morphological, and semantic features. Richards (1976, as cited in Mukarto, 2005b) indicated seven assumptions of vocabulary knowledge:

1) knowing its relative frequency and its collocation, 2) knowing the limitation imposed on its use, 3) knowing its syntactic behavior, 4) knowing its basic forms and derivations, 5) knowing its association with other words, 6) knowing its semantic value, and 7) knowing many of the different meanings associated with the word (p.8).

Nation (1998) mentions a wide range of aspects of what is involved in knowing a word: first and foremost, familiarity with the written and spoken form of the word and ability to connect a concept with this form; second, ability to use appropriate and grammatically correct collocation with the word; third, ability to make other members of the word by using appropriate affixes; fourth, awareness of cultural, geographical and stylistic limitations of using the word; and finally, awareness of different meanings and associations of the word.

Nation (1998) insists on the awareness of these aspects of word knowledge. He also believes that this awareness should be based on some regular systems so that learners can easily recognize what to look for and can easily observe gaps in their knowledge. One of his suggested systems is to use the items of the table of what is involved in knowing a word.
(2001, Appendix A), which includes word knowledge in receptive and productive modes in both spoken as well as written channels.

Considering such a clarification of word knowledge, one can conclude that the aforementioned Iranian English student’s belief about word knowledge may be inadequate and problematic. However, the way those who are involved with languages (students or teachers) think about words has a great influence on how successful they learn or teach words. As Fishbein and Ajzen and Ajzen (1975, 1988, as cited in Bernat & Gvozdenko, 2005, para.2) indicate, “beliefs are a central construct in every discipline that deals with human behavior and learning”.

Teacher beliefs show their importance as they help teachers to have a better understanding of their class, their teaching objectives and their lesson plans. In this regard, Richards (1994, p.385) indicates “Teacher beliefs form a structured set of principles that are derived from experience, school practice, personality, educational theory, reading and other sources”. As studies show (Kelly, 1995, as cited in Shehni, 2008; Richards, 1994), there is a mutual connection between teacher’s beliefs and their classroom practice. However, there always are learners whose beliefs and ideas show their effectiveness not only in how successfully they acquire a language but also in teachers’ methods and lesson plans.

Learner beliefs, according to Richardson (1996, as cited in Huang, 2006, p.62), are “psychologically held understandings, premises, or propositions about the world that are felt to be true”. Huang (2006) also indicates that there is no clear definition of learner beliefs of language learning although there are varied understandings of learner beliefs. She continues that learners have their own beliefs about what they should study. As Dakin (1973, as cited in Huang, 2006, p.63) stated “though the teacher may control the experiences the learner is exposed to, it is the learner, who selects what is learnt from them”.

A great number of research studies on teacher and learner beliefs about different subjects have already been done (Farrell & Patricia, 2005; Flippo & Caverly, 2009; Harmon, 1998; Lacorte & Canabal, 2005; Thornton, 2009; Tanaka & Ellis, 2003; Woods, 1991, etc.). For example, Woods (1991, as cited in Zacharias, 2005) carried a longitudinal study of two TESL teachers with different theoretical beliefs. Both teachers taught the same course in a university in Canada. While one of the teachers had a ‘curriculum-based’ orientation, the other one had a ‘student-based’ orientation. Findings showed that the teacher with ‘curriculum-based’ orientation evaluated her teaching based on how successful she had
accomplished what she had preplanned according to the curriculum. However, the other teacher evaluated her teaching based on students’ responses. Regarding learner beliefs, Harmon (1998, as cited in Flippo & Caverly, 2009) was the first who investigated the connection between student beliefs and their vocabulary knowledge. Harmon asked middle-school students to participate in the think-aloud procedure to determine their perceptions of word meanings. She found that some students were unaware of appropriate processes that could reveal the best contextual meaning of a word. Flippo and Caverly (2009) indicate that while Harmon did not ask directly about the students’ beliefs about vocabulary knowledge, the descriptions used by the students to narrate the vocabulary acquisition process showed their fundamental belief systems about the uncertainty of vocabulary knowledge.

However, regarding beliefs about vocabulary knowledge, research studies have been so rare that Flippo and Caverly (2009) indicate “at this point in time, the area of beliefs about vocabulary knowledge is still in the infancy stage of research” (p.111). As a result, if we consider learner beliefs as an important aspect of learning process, which can condition their attitudes and behaviors, we first need to know what these beliefs are and to what extent they are similar to teacher beliefs in both L1 and L2. The present study is a reply to such a requirement. The following questions will be answered in this study:

1. To what extent are various aspects of word knowledge, as defined by Nation (2001), taken as important in L2 by TEFL graduate students and teachers as well as English native TESL teachers and professors?
2. To what extent are various aspects of word knowledge, as defined by Nation (2001), taken as important in L1 by TEFL graduate students and teachers, and Persian Literature graduate students and teachers?
3. Does each group agree with others on the importance they have given to the aspects of word knowledge in L1 and L2?
4. Is there a relationship between what TEFL graduate students and teachers rated as important in the reception and production questionnaires, concerning aspects of word knowledge, and their actual performance on the reception and production tests?
5. To what extent is Nation's table a significant criterion for what is involved in knowing a word in L1 and L2 for TEFL graduate students and teachers, in L1 for Persian Literature graduate students and teachers, and in L2 for native TESL teachers or professors?
Methodology

Participants

Five groups of participants took part in this research: (1) 21 TEFL graduate students (TEFLGS hereafter), (2) 21 Persian Literature graduate students (PLGS hereafter), (3) 20 native TESL Canadian and American English teachers and professors with M.A. and Ph.D. degrees in TESL from the English language institute of the University of British Columbia, Alberta, Calgary, Canada, Saint Michael’s College, and the University of Iowa (NTTP hereafter), (4) 20 TEFL graduate teachers who were among TEFL graduate students with at least four years of teaching experience (TEFLGT hereafter), and (5) 29 PL graduate teachers, who were among PL graduate students with at least four years of teaching experience (PLGT hereafter). The participants were chosen from students who were studying in Najafabad University (IAU), Iran, (except group 3) regardless of their gender, age or any other special criteria. However, the selected teachers had to have necessarily more than four years of teaching experience.

Instruments

The main instruments used in this research study were the reception and production proficiency tests (RPT & PPT hereafter) originally designed and developed by Nation et al. (2001, pp. 416-425) and simplified by the researchers, Nation’s table of word knowledge (2001, p. 27), the reception and production tests (RT & PT hereafter) made by the researchers, the English and Persian reception and production questionnaires (ERQ, EPQ, PRQ & PPQ hereafter) designed for TEFL graduate students and teachers, the Persian questionnaire (PQ hereafter) designed for PL graduate students and teachers, and the English questionnaire (EQ hereafter) designed for Canadian and American TESL teachers and professors. All of the questionnaires were developed according to Nation’s table of word knowledge. The following explanations fully clarify how these materials were prepared.

Reception and production proficiency tests

RPT and PPT which originally belong to Nation et al. were adapted by the researchers to show whether the participants surely had a knowledge of 3300 absolutely essential words, which equals the frequency bands of 5, 4 and 3 of Collins Cobuild dictionary (2001). This number of words equals the 4000 word level in “BNC-20 VocabProfile” (VP hereafter) adapted by Tom Cobb (2003) for the Web, which is based on Laufer and Nation's Lexical
Frequency Profiler (1995). The Lexical Frequency Profile measures the amount of vocabulary from different frequency levels used in learners’ composition writings.

For this purpose, the reduced form of the Vocabulary Levels Test (VLT hereafter), Test B and Productive Levels Test (PLT hereafter), Version C, presented in Nation (2001, pp. 416 & 425) were chosen. The original VLT made by Nation is described in Nation (1983 & 1990). Test B is a major improvement of the original test, which was designed and made reliable and valid by Schmitt, Schmitt and, Clapham (2001). PLT is based on VLT, but it tests vocabulary productively rather than receptively. The description of making and validation of this test is presented in Laufer and Nation (1999). Each test consists of 2000, 3000, 5000 academic and 10000 word levels in an increasing order of difficulty. Since for this research the participants’ performance on at most the 4000 word level of the VP was supposed to be evaluated, the “academic word list” and the “10000 word level” parts of these two tests were omitted. The tests were piloted with a different sample from the TEFL graduate population in IAU, became reliable by Cronbach’s alpha (r = .87), and were validated by three English professors to be given to graduate students in the same university.

There were some reasons behind choosing these two tests. First, this research aimed to choose those who had the knowledge of the VP 4000 word level (equals 3300 word level of Cobuild dictionary), and these two tests were appropriate to measure this precisely. Second, these two tests were equivalent forms, which became reliable and valid by their own designers. Last but not least, they were approved of being appropriate for Iranian English language students by three Iranian professors.

Since these two tests were diagnostic tests to diagnose where learners need help, Nation (communicated, consulted and informed via e-mail) preferred his own Vocabulary Size Test for this purpose. Nonetheless, the VST did not seem appropriate for the participants, especially its third 1000-word level seemed difficult for them. Moreover, the VST did not have an equivalent productive form.

**Nation’s table of word knowledge**

This research has been based on Nation’s table (2001, p. 27) of what is involved in knowing a word. Three main factors are considered as important in knowing the word: form, meaning and use. Each of these main factors has their own sub-factors further divided into receptive and productive categories as well. Totally, 18 items are considered as important in knowing a word.
English and Persian reception and production questionnaires designed for TEFL graduate students and teachers (Appendix B)

These questionnaires were designed to seek the participants' opinion on what is involved in knowing a word and the degree of importance they give to each item. Therefore, Nation’s table of word knowledge changed to two tables of the English reception questionnaire (ERQ), items 1, 3, 5, 7, 9, 11, 13 and 17 and the English production questionnaire (EPQ), items 2, 4, 6, 8, 10, 12, 14, 16 and 18. This division was made to ease the administration of the tests in two different sessions, one for reception and the other for production. Each item changed from a question (original form) to a statement (questionnaire form) and care was taken to have the related receptive and productive items as parallel statements. The participants were asked to rate the level of importance of each item in knowing a second language word by checking the box in the table. The importance of each item was considered to be rated on a scale of 1 to 5, very important, important, moderately important, less important, and not important. The participants’ full comprehension of the statements and what they were supposed to do were matters of concern. As a result, Persian directions, Persian translation of a few words in the questionnaire items, and some pieces of explanation for some questionnaire items were added to the questionnaires. Moreover, to enable understanding and to make the participants’ better perform on each item, it was mentioned in the directions that RT and PT had been designed based on the aspects (items) of these questionnaires, and each section of RT and PT (A to I) was an example of the following questionnaires’ aspects. Therefore, to give the participants this opportunity to compare each item by its own designed test section, item numbers were removed and letters of A to I, just like RT and PT, were used instead. Doing so, the researcher let the participants perceive questionnaire items practically.

Persian reception and production questionnaires (PRQ & PPQ) underwent the same procedures to be created. Much care was taken to translate the items as precisely and accurately as possible. Just like the English versions to make items more comprehensible, some Persian explanations were added to some items. However, this time word knowledge in the mother tongue was asked to be rated by the participants. Between 5 to 10 minutes was considered for each of the receptive and productive questionnaires to be rated by the participants.
Reception and production tests (Appendix C)

RT and PT were designed to measure the productive and receptive skills of the participants, and the results were used to see if there was any relationship between the participants’ performance and opinions on different items of Nation’s table.

In these two tests, all the target words chosen for the tests, the answers, and even the distractors were selected from high frequency words. The reason was that these tests were supposed to measure the different aspects of the participants’ word knowledge of the target words as presented by Nation both receptively and productively. Otherwise, the lack of success on the part of participants might have resulted from their unfamiliarity with the target words or their breadth of vocabulary knowledge (vocabulary size), which was not the focus. The target words were chosen from different sources like Vocabulary in Use (1999 & 1997), Study Skills (1998), 504 Absolutely Essential Words (2001).

The frequency of the words is based on the frequency bands of Collins Cobuild dictionary (2001). The considered word frequency for these tests is 3 or more among the 5 frequency bands equal to the 4000 word level of the VP.

Based on the dictionary explanation, the words in bands 5 and 4 (approximately 680 and 1040 words respectively) account for about 75% of all English usage, and it was the reason for choosing these two bands for this research. Band 3 includes approximately 1580 words. Knowing the words in this band extends the range of topics which one can talk about, and it was the reason behind choosing this band for this research as well. Considering the mentioned number of words in these three frequency bands (680+1040+1580), it should be noticed that the participants had to be at least at the level of 3300 words based on the frequency bands of this dictionary. Moreover, all the target words, their answers and distractors of these two tests, after being controlled by Cobuild’s frequency bands to be among bands 5, 4 and 3, were programmed into the VP software. Almost all were among the first, second, third, and fourth 1000-word level (totally 4000 word level).

Moreover, Nation’s instruction in the section entitled “How can depth of knowledge about a word be tested?” in his book Learning Vocabulary in Another Language (2001, p.354) was the basis of designing these two tests; however, there were a few changes due to the designers’ opinion and his negotiation with Nation through some e-mail communication. The one to one item validly of test items with the questionnaires was confirmed by Nation and three English professors.

273
Finally in this regard, some issues were considered as important in test design, which are included in Appendix D.

**Persian questionnaire designed for PL graduate students and teachers**
PQ was designed to seek the opinions of a group of participants on aspects of word knowledge in mother tongue rather than the second language to make the comparison with other opinions possible. In other words, since this group had low proficiency in English, they were expected to express their pure opinion on word aspects in L1 without contamination with the knowledge in L2. Therefore, the aforementioned PRQ and PPQ were put together to make an 18-item questionnaire just like the original table and translated into Farsi. A bio-data section was added to the questionnaire asking for information such as experience in teaching PL (if yes how long and in which grades) and familiarity with any foreign languages.

**English questionnaire designed for Canadian and American TESL teachers and professors**
The aforementioned ERQ and EPQ were put together to make an 18-item questionnaire as the original table and Persian translations of the items were removed.

The tests and the questionnaires were piloted with similar samples and their internal consistency was measured (Cronbach’s alphas between .75 to.87).

**Procedures**
Fifty six graduate students from three classes were chosen and received RPT, RT, ERQ and PRQ in one session and PPT, PT, EPQ and PPQ two weeks later in another session. Based on the Guide to Levels Test by Nation, only those who were able to score 27 or more out of 30 words in RPT, and 16 or more out of 18 words in PPT were supposed to be chosen for the RT and PT. However, the problem encountered was that the productive performance of students was lower than expected. As a result, for RPT, 27 or more out of 30 scores in parts 1 and 2, and 14 or more in part 3 were selected as the basic requirement of being in the data. These numbers in PPT were 14 or more out of 18 (3/4 of each section) in sections 1 and 2, and almost around 8 or more in section 3. The performance of 42 out of 56 participants, almost half of them were TEFLGS and the other half were TEFLGT, was in this range and probable to be considered as the research data. Therefore, their RT and PT were scored and their
answers to the items of the questionnaires were considered. The Persian questionnaire was
given to PL graduate students and teachers, and EQ was sent to native TESL teachers and
professors by email.

Data analysis
The following analyses, first descriptive then inferential, were applied to the data to answer
the first (a), second (b), third (a, b, c), fourth (d), and fifth (a, b, c) research questions. The
scales of the questionnaires were given values from 1 (not important) to 5 (very important).

a) The participants’ performance on all individual items (A to I) of ERQ and EPQ
b) The participants’ performance on all individual items (A to I) of PRQ and PPQ
c) The participants’ performance on all variables of the study including: group (TEFLGS,
   PLGS, NTTP, TEFLGT, PLGT), questionnaire part (form, meaning, use), mode
   (receptive, productive), and questionnaire type (English, Persian)
d) The participants’ performance on all the individual parts of RT and PT (A to I), and
   the relationship of each part with the related parts in ERQ and EPQ

Performance on ERQ items
The results of descriptive analysis of TEFLGS, NTTP, and TEFLGT are tabulated in Table 1
(left part) and illustrated in Figure 1.

Table 1 Descriptive Statistics of Participants’ Performance on English (right) and Persian
(left)
Questionnaires
| Q1 | TEFL GS | 4.19 | 2 | 5 | 3.95 | 2 | 5 | NTTP | 4.55 | 3 | 5 | 4.45 | 3 | 5 | TEFL GT | 4.42 | 3 | 5 | 4.80 | 4 | 5 | Total | 4.38 | 2 | 5 | 4.39 | 2 | 5 |
| Q2 | TEFL GS | 3.52 | 2 | 5 | 3.67 | 2 | 5 | NTTP | 3.25 | 2 | 5 | 3.30 | 2 | 5 | TEFL GT | 3.58 | 2 | 5 | 3.75 | 3 | 5 | Total | 3.45 | 2 | 5 | 3.57 | 2 | 5 |
| Q3 | TEFL GS | 3.57 | 2 | 5 | 3.52 | 2 | 5 | NTTP | 3.25 | 2 | 5 | 3.45 | 2 | 5 | TEFL GT | 4.05 | 2 | 5 | 3.75 | 3 | 5 | Total | 3.62 | 2 | 5 | 3.62 | 2 | 5 |
| Q4 | TEFL GS | 4.29 | 3 | 5 | 4.05 | 3 | 5 | NTTP | 4.65 | 3 | 5 | 4.05 | 2 | 5 | TEFL GT | 4.21 | 2 | 5 | 4.05 | 2 | 5 | Total | 4.38 | 2 | 5 | 4.05 | 2 | 5 |
| Q5 | TEFL GS | 3.48 | 2 | 5 | 3.29 | 2 | 4 | NTTP | 3.10 | 2 | 5 | 3.60 | 2 | 5 | TEFL GT | 3.63 | 2 | 5 | 3.45 | 2 | 5 | Total | 3.40 | 2 | 5 | 3.64 | 1 | 5 |
| Q6 | TEFL GS | 3.43 | 2 | 5 | 3.52 | 2 | 5 | NTTP | 3.90 | 2 | 5 | 3.60 | 2 | 5 | TEFL GT | 3.99 | 2 | 5 | 3.75 | 2 | 5 | Total | 3.40 | 2 | 5 | 3.50 | 2 | 5 |
| Q7 | TEFL GS | 3.67 | 2 | 5 | 3.52 | 2 | 5 | NTTP | 4.00 | 2 | 5 | 3.95 | 2 | 5 | TEFL GT | 3.89 | 2 | 5 | 3.70 | 1 | 5 | Total | 3.85 | 2 | 5 | 3.97 | 1 | 5 |
| Q8 | TEFL GS | 3.57 | 2 | 5 | 3.67 | 2 | 5 | NTTP | 4.00 | 2 | 5 | 3.90 | 2 | 5 | TEFL GT | 4.42 | 2 | 5 | 4.05 | 2 | 5 | Total | 4.00 | 2 | 5 | 3.95 | 2 | 5 |
| Q9 | TEFL GS | 3.67 | 2 | 5 | 3.86 | 2 | 5 | NTTP | 3.50 | 2 | 5 | 3.55 | 1 | 5 | TEFL GT | 3.63 | 2 | 5 | 3.90 | 2 | 5 | Total | 3.62 | 2 | 5 | 3.79 | 1 | 5 |

Generally, all items are given importance between 3.10 and 4.55. Also, no big difference is observed between the groups in all items except items 3, 7, and 8. Items 1 and 4 are considered as more important than others by all groups.
More specifically, while ERQ items 1 and 4 are rated as more than *important* but not *very important* by all groups, items 2, 5, 6, and 9 are rated between *moderately important* and *important* by the same groups. Item 3 is rated as *important* by TEFLGT and between *moderately important* and *important* by the other two groups. Item 7 is rated as important by NTTP and between *moderately important* and *important* by the other two groups. Item 8 is rated between *moderately important* and *important* by TEFLGS, as *important* by NTTP, and as more than *important* but not *very important* by TEFLGT. It could also be mentioned that only NTTP have rated ERQ item 4 almost as *very important*. Not a big difference is observable among most of the groups’ mean scores (except items 1 & 8). However, in both items while TEFLGT shows the highest mean score, NTTP in item 3 and TEFLGS in item 8 show lower mean scores.

![Figure 1 Performance on ERQ items](chart.png)

Nine One-way ANOVAs run on mean scores of ERQ items 1 to 9 show significant differences among groups in items 3 and 8, $F (2, 58) = 3.77, p = .02$ and $F (2, 58) = 4.08, p = .02$, and Scheffe tests show the mean differences between NTTP and TEFLGT in item 3, and between TEFLGS and TEFLGT in item 8 are significant (- 0.8 & - 0.84 respectively).
Performance on EPQ items

The results of descriptive analysis of performance of TEFLGS, NTTP, and TEFLGT on EPQ items are displayed in Table 1, left part, and Figure 2.

Generally, all items are given importance between 3.29 and 4.80. Also, no big difference is observed between the groups in all items except items 1 and 8. More specifically, items 2, 3, 5, 6, 7, and 9 are rated between moderately important and important. Item 1 is rated between moderately important and important by TEFLGT, but between important and very important by NTTP and TEFLGT. Item 4 is rated as important by all three groups. Item 8 is rated as important by TEFLGT and between moderately important and important by other two groups. Only EPQ item 1 is rated almost as very important by TEFLGT. All three groups in each item of EPQ have more or less the same mean scores except item 1. In other words, not a big difference is observable among the groups’ mean scores in the aforementioned items except item 1. In item 1, while TEFLGT shows the highest mean score (4.8), NTTP and TEFLGS show lower mean scores (4.45 & 3.95 respectively).

![Figure 2 Performance on EPQ items](image)

Nine One-way ANOVAs run on the mean scores show a significant difference among groups in item 1, \( F(2, 58) = 6.15, p = .004 \), and Scheffe test shows that the mean difference between TEFLGS and TEFLGT is significant (- 0.84); TEFLGT’s mean score is higher than TEFLGS’s.
Performance on PRQ items

The results of descriptive analysis of performance of TEFLGS, PLGS, TEFLGT, and PLGT on each item of PRQ are displayed in Table 1, right part, and Figure 3.

<table>
<thead>
<tr>
<th>Group</th>
<th>PRQ Items</th>
</tr>
</thead>
<tbody>
<tr>
<td>TEFL graduate</td>
<td>Persian reception</td>
</tr>
<tr>
<td>students</td>
<td>questionnaire 1</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>questionnaire 2</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>questionnaire 3</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>questionnaire 4</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>questionnaire 5</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>questionnaire 6</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>questionnaire 7</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>questionnaire 8</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>questionnaire 9</td>
</tr>
</tbody>
</table>

Generally, all items are given importance between 3.19 and 4.77. Also, no big difference is observed between the groups in all items except items 1, 2, 4, 6, and 8. More specifically, while PRQ items 3, 5, 6, 7 and 9 are rated between moderately important and important, item 4 is rated between important and very important. Even it could be declared that item 4 is rated as almost very important by PLGS and PLGT. Item 1 is rated between moderately important and important by TEFLGS; however, it is rated as more than important but not very important by other groups. Regarding item 2, TEFLGS and TEFLGT have rated that between moderately important and important, while it is rated by PLGS and PLGT as more than important but not very important. Regarding item 8, while it is rated between moderately important and important by TEFLGS and PLGS, it is rated as important by TEFLGT and PLGT. All four groups in each item of PRQ (except item 2) have more or less the same mean scores. In other words, only item 2 shows the biggest difference among mean scores of PLGT (4.58) and TEFLGS (3.28).
Nine One-way ANOVAs run on these four group mean scores show significant
differences among groups in items 1, 2, 4, 6 and 8, \( F (3, 92) = 3.97, \ p = .01; \ F (3, 92) = 
16.86, \ p = .00; \ F (3, 92) = 4.26, \ p = .00; \ F (3, 92) = 2.92, \ p = .03; \) and \( F (3, 92) = 3.15, \ p 
= .02, \) and Scheffe tests show that the mean differences between TEFLGS and PLGS, and 
TEFLGS and PLGT in item 1 (- 0.68 & - 0.66); TEFLGS and PLGS, TEFLGT and PLGS, 
TEFLGT and PLGT and TEFLGS and PLGT in it em 2 (- 1.13, - 1.04, - 1.16 & - 1.004 
respectively) are significant; the former groups have lower mean scores.

**Performance on PPQ items**
The results of descriptive analysis of performance of TEFLGS, PLGS, TEFLGT, and PLGT 
on PPQ items are displayed in Table 1, right part, and Figure 4.

![Figure 4 Performance on PPQ](chart.png)

Generally, all items are given importance between 3.2 and 4.42. Also, no big
difference is observed between the groups in all items except items 2, 4, 6, and 9. Item 1 is
considered as important by all groups. More specifically, PPQ item 1 is rated as more than
important but not very important, while items 3, 5, 7 and 8 are rated between moderately
important and important by all groups. Item 2 is rated between moderately important and
important by TEFLGS and TEFLGT and as more than important but not very important by PLGS and PLGT. In addition, while item 4 is rated between moderately important and important by TEFLGS, it is rated as more than important but not very important by the other three groups. However, item 9 is rated as more than important but not very important by TEFLGT and between moderately important and important by the other three groups. Finally, item 6 is rated between moderately important and important by TEFLGS and PLGT, and as more than important but not very important by PLGS and TEFLGT. All four groups in each item of PPQ (except item 2) have more or less the same mean scores, and no noticeable difference is observable. Item 2 shows the biggest difference among mean scores of the groups. In item 2, while PLGT shows the highest mean score (4.3), TEFLGT shows the lowest (3.4).

Nine One-way ANOVAs run on these four group mean scores show significant differences among groups in items 2, 4 and 6, $F(3, 92) = 6.47, p = .001$; $F(3, 92) = 2.58, p = .04$; $F(3, 92) = 3.17, p = .02$, and Scheffe tests show that the mean differences between TEFLGS and PLGT, TEFLGT and PLGS, and TEFLGT and PLGT in item 2 are significant (-0.74, -0.78 & -0.89 respectively); the former groups have lower mean scores.

**Four-way ANOVA among all the variables of the study**

First, the individual items of the questionnaires were assigned to different parts including: form (items 1-3), meaning (items 4-6), and use (items 7-9). Then, to have a better understanding of the whole data analysis, the descriptive statistics of all four variables (group, questionnaire part, mode and questionnaire type) were calculated and a four-way ANOVA was run to see whether there were main effects of all variables as well as any possible interaction among them. The following table shows the results.
Table 2 Restuls of Three-Way ANOVA Among Variables

<table>
<thead>
<tr>
<th>Source</th>
<th>F</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>GROUP</td>
<td>14.63</td>
<td>0.00</td>
</tr>
<tr>
<td>PARTS</td>
<td>5.78</td>
<td>0.00</td>
</tr>
<tr>
<td>RECPRO</td>
<td>0.23</td>
<td>0.63</td>
</tr>
<tr>
<td>ENGPERS</td>
<td>2.66</td>
<td>0.10</td>
</tr>
<tr>
<td>GROUP * PARTS</td>
<td>2.09</td>
<td>0.03</td>
</tr>
<tr>
<td>GROUP * RECPRO</td>
<td>0.61</td>
<td>0.66</td>
</tr>
<tr>
<td>PARTS * RECPRO</td>
<td>1.77</td>
<td>0.17</td>
</tr>
<tr>
<td>GROUP * ENGPERS</td>
<td>0.17</td>
<td>0.99</td>
</tr>
<tr>
<td>PARTS * ENGPERS</td>
<td>0.75</td>
<td>0.39</td>
</tr>
<tr>
<td>GROUP * PARTS * ENGPERS</td>
<td>0.61</td>
<td>0.55</td>
</tr>
<tr>
<td>RECPRO * ENGPERS</td>
<td>1.58</td>
<td>0.21</td>
</tr>
<tr>
<td>GROUP * RECPRO * ENGPERS</td>
<td>0.00</td>
<td>0.98</td>
</tr>
<tr>
<td>PARTS * RECPRO * ENGPERS</td>
<td>0.45</td>
<td>0.64</td>
</tr>
<tr>
<td>GROUP * PARTS * RECPRO * ENGPERS</td>
<td>1.12</td>
<td>0.33</td>
</tr>
</tbody>
</table>

As Table 2 shows, there is a main effect of group (1 to 5), $F(4,900) = 14.62, p = .00$, a main effect of parts (form, meaning and use), $F(2, 900) = 5.78, p = .00$, and an interaction between group and parts, $F(8, 900) = 2.09, p = 0.03$. Scheffé tests were performed for each of these three items. The mean differences between TEFLGS and PLGS, TEFLGS and TEFLGT, TEFLGS and PLGT, and NTTP and PLGT are found significant (-0.96, -0.93, -1.27 & -0.76 respectively); the former groups have lower mean scores. Moreover, the mean difference between form and use is significant (0.56); the mean score of the form part is higher than the use part. Figure 5 illustrates the same information.
Performance on RT

The descriptive statistics of the performance of TEFLGS and TEFLGT on nine items of RT (A to I) show the best performance of the participants on item 2 of RT and the worst on item 9. Pearson correlation tests run between the nine items of RT and their counterpart questions fail to show any relationship between the two.

Performance on PT

The descriptive statistics of the performance of TEFLGS and TEFLGT on nine items of PT show the best performance of the participants on item 2 of PT and the worst on item 9. No significant correlation between the nine items of PT and their equivalent questions in EPQ was found.

Summary of results

Regarding the four variables of the study (group, questionnaire part, questionnaire mode, and questionnaire type), significant effect of group and part and their interaction on different items of the questionnaires was observed. However, no significant effect was observed for mode and questionnaire type. In other words, the participants’ performance on reception questionnaires was generally the same as production questionnaires, and the participants’ performance on L1 questionnaire was generally the same as L2 questionnaire, which means the aspects of vocabulary knowledge were evaluated almost the same in L1 and L2.
Regarding the English questionnaire, generally all 18 items of the questionnaire were rated between *moderately important* and *very important*. In ERQ, TEFLGT rated recognition of the word parts (suffix, stem, prefix) and recognition of the words which are used with this word (collocations) significantly as more important than NTTP and TEFLGS respectively. Regarding EPQ items, TEFLGT rated the *way the word is pronounced* significantly as more important than TEFLGS.

Regarding Persian questionnaire, all the 18 items of the questionnaire were rated between *moderately important* and *very important*. Regarding items, in PRQ, PLGS and PLGT rated the *way the word sounds like* significantly as more important than TEFLGS, and PLGT and PLGS rated the *way the word looks like* significantly as more important than TEFLGS and TEFLGT. Finally, in PPQ, PLGT rated the *way the word is spelled and written* significantly as more important than TEFLGS and TEFLGT, and PLGS rated the same item significantly as more important than TEFLGT.

Regarding the tests, the participants did not perform on the tests based on what they rated as important in ERQ and EPQ.

**Discussion**

Regarding the first three research questions, the participants performed differently (between *moderately important* and *very important*) on some of the 18 items of the questionnaires. TEFLGT ranked three items of the English questionnaire significantly as more important than TEFLGS and NTTP. Moreover, PLGS and PLGT ranked two items of Persian questionnaire significantly as more important than TEFLGS and TEFLGT.

Regarding L2 (English questionnaire), TEFLGT gave more importance to the pronunciation of the words, word parts, and the collocations of the words than TEFLGS and NTTP. Generally, TEFLGT gave more importance to the productive form (pronunciation, spelling and written form, and affixes of the words) than TEFLGS and NTTP. In other words, regarding L2, TEFLGT considered some aspects as more important. It seems that Iranian English teachers consider pronunciation, spelling and written form, word parts, and collocations as being more necessary and possibly as the underlying patterns of words which make a suitable basis for practicing the newly learned words. This in fact is in congruence with what Nation believes.
Nation (2001) believes that teachers should have the ability to estimate the amount of effort required to learn a word. In this way, he continues that teachers can focus on aspects that will need attention and will reveal underlying patterns. Nation (1994) also believes that teachers should give their students the chance to observe the newly learned words in new contexts, which in fact provide new collocations and associations. However, it seems that these factors may not be considered that important by native TESL teachers. The reason for this difference might be what Woods (1996, as cited in Shehni, 2008) declares that teachers’ priorities depend on teachers’ own assumptions about language, learning and teaching. As a result, each teacher’s value system differs from the others, and as Richards (1994) says, differences in teacher beliefs resulted from the amount of teaching experience they have and the issue of whether they subscribed to a functional or grammar based orientation to teaching. Moreover, native TESL teachers might take the aforementioned factors for granted and simply do not consider those factors important in an ESL context. Perhaps pronunciation, spelling and affixes are factors, which are learnt in an English speaking country automatically out of class. More research on opinions may reveal if and to what extent they differ among various academic settings.

In the same vein, the aforementioned factors are not allocated that degree of importance by Iranian English students as well. One possible reason could be a discrepancy between what teachers consider as important and what they actually practice in class. In other words, it could be mentioned that despite teachers’ beliefs about the importance of these factors, they are not given enough emphasis and practice in class. As a result, students do not pay much attention to those aspects as well. This class practice is addressed by some scholars like Channell (1998, as cited in Richards & Renandya, 2002), who believes that language learners should do more than just observing the words when being taught unfamiliar vocabulary. In fact, they need to hear the pronunciation and practice saying the word aloud. This finding is in congruence with what Duffy and Anderson (1986, as cited in Zacharias, 2005) found in their study. In a study of eight reading teachers they found that only four out of eight put their beliefs into practices.

Similarly, Yim (1993, as cited in Zacharias, 2005) found that some TESL teachers in Singapore did not pay that attention to the role of grammar in their classes they used to consider as significant. In other words, the connection and interaction between teacher beliefs and their classroom practice does not always develop. Zacharias (2005) considers different
teaching situations in different contexts as the reason why some teacher beliefs have an impact on their teaching in some cases but not in others.

In addition, this disagreement between teachers and students is in congruence with what has already been manifested in some research studies. For example, recent research on learner beliefs about language learning has uncovered a gap between teacher and learner beliefs (Dakin, 1973; Allwright, 1984; Malamah-Thomas, 1987; Horwitz, 1987; Block, 1994; Kern, 1995; Peacock, 1998, as cited in Huang, 2006). Some researchers, based on their empirical or theoretical research, argue that the reasons for this gap are the mismatch between the pedagogical agenda of the teacher and that of the learner (Allwright, 1987; Nunan, 1995, as cited in Huang, 2006). Some others suggest other factors such as culture, teaching styles, learning strategies and personality. Still some researchers propose different ways of filling this gap by teacher intervention (Horwitz, 1988; Kern, 1995; Mantle-Bromley, 1995; Peacock, 1998, as cited in Huang, 2006) and learner-centeredness (Nunan, 1995, as cited in Huang, 2006).

Regarding L1 (Persian questionnaire), PLGS and PLGT rated the way the word sounds like as more important than TEFLGS. PLGS rated the way the word is spelled and written as more important than TEFLGS and TEFLGT. The same as what was mentioned about native TESL teachers regarding their own native language, it seems that TEFLGT and, especially TEFLGS, took the aforementioned aspect of word knowledge in their own mother tongue for granted and did not rate it as important as the way it is rated by PLGS and PLGT. Special attention of PLGT to spelling and pronunciation of words might result from the fact that L1 literacy is the main reason children are sent to school.

Regarding the fourth research question, it should be mentioned that TEFL graduate students and teachers did not perform on the reception and production tests based on what they rated as important regarding aspects of word knowledge in the reception and production questionnaires. In other words, they gave almost the same importance to all 18 items of the questionnaires while their performance on items 17 and 18 was significantly worse than the others. These items are related to the “use” of language (Nation, 2001). The finding is given significance, for we theoretically define beliefs as the underlying learner’s behaviors, their learning strategies, their learning confidence and anxiety, and consequently their linguistic and non-linguistics outcomes (Ellis, 1994; Gardner & MacIntyre, 1992, 1993, as cited in Tanaka & Ellis, 2003). Consequently, a positive relationship between the two (ratings and
scores) is expected. However, the relationship between these variables seems not to be that simple. In other words more mediator variables like proficiency and motivation might be at work. Such findings are reported by Tanaka and Ellis (2003) as well. They examined a 15-week- study-abroad program for changes in student beliefs about language learning (measured by means of a questionnaire) and in their English proficiency (measured by means of TOEFL). They report that Pearson product moment correlations between the students’ responses to the Belief Questionnaire and their TOEFL scores both before and after the study-abroad period were weak and generally statistically non-significant.

Generally, these findings about students’ beliefs and their performance had been to some extent attributable to the fact that students are always busy learning, experiencing, developing ideas, putting into practice their ideas, testing and retesting their own thoughts and opinions. However, the finding of the present study about teachers who had at least 4 years of teaching were almost unpredictable since their beliefs and their performance were supposed to be congruent with each other or at least to have slight differences or incongruence after 4 years of teaching. However, such fundamental incongruence was very much unpredictable.

In line with other studies done on variable factors involved in knowing a word, the findings of this research reveal which aspects of word knowledge are given importance in L1 and L2. The results contribute to the existing literature regarding depth of vocabulary knowledge, beliefs about how important each aspect is in L1 and L2, similarities and differences between the two, agreement or disagreement between students and teachers involved in the field, and any possible relationship between the participants’ performance and their opinions on different items of the Nation’s table of word knowledge.

Regarding the last question, the findings of the study show that generally the participants rated items of Nation’s table between moderately important and very important, and the findings were the same in L1 and L2. As a result, Nation's table was an important (regarding the 5-point Likert) criterion for what involves in knowing a word in L1 and L2 for TEFL graduate students and teachers, in L1 for Persian Literature graduate students and teachers, and in L2 for native TESL teachers or professors. No study has investigated the opinion of teachers and students on the importance of word knowledge aspects; however, the multidimensionality of depth of vocabulary knowledge has been observed by a lot of scholars. For example, Qian (1999, as cited in Shen, 2008) declared that the depth dimension should cover such important components as pronunciation, spelling, meaning, register, frequency,
and morphological, syntactic, and collocational properties. Schmitt (2000) indicated that orthographical, phonological, lexical, and grammatical knowledge as well as frequency and collocation were important factors to both vocabulary knowledge and language processing.

Last but not least, Henrikson (1995) believed that precise relation of the word to other words within the same semantic fields, other linguistic contexts such as collocations the words can be part of, the particular stylistic level of the word, and whether the word can be used figuratively or be part of particular idiomatic expressions as well as the meaning of the word were as important factors involved in knowing a word. Therefore, one whose major is teaching or learning a language, no matter L1 or L2, does not seem to be able to ignore the existence or importance of the aforementioned aspects of word knowledge.

**Implications of the study**

As already mentioned, what it means to know a word has always engaged psychologists, scholars, teachers, and language learners. Knowledge of the meaning of a word as the sole factor in knowing that word is the most important misunderstanding in which some language learners and probably some language teachers still believe. However, the findings of this research surely help those who are ignorant of some other factors involved in knowing a word. Language students face those aspects they may not be aware of, and they might even evaluate themselves to see whether their productive and receptive word knowledge performance is based on what they consider as important in knowing a word.

Language teachers might be reminded of those aspects of word knowledge which they have been ignorant of, and they can become aware of their native colleagues’ ideas and beliefs in this regard. Similarly, they might use the same tests of depth of vocabulary knowledge to see whether they have been successful in teaching different aspects of word knowledge to their students. Moreover, Native TESL teachers can get some ideas about and familiarity with EFL environment, EFL students’ and TEFL teachers’ beliefs about word knowledge which are surely useful for their coming research studies and future teaching opportunity they might be offered.

Generally and hopefully, the findings of this research study increase the general understanding of depth of vocabulary knowledge among L2 learners and teachers as well as L1 learners and teachers, help ESL and EFL learners and teachers to be aware of each other beliefs, and help teachers to have more successful teaching and assessment of different aspects of word knowledge.
References


### Appendix A

Nation’s table of word knowledge

Source: learning vocabulary in another language by Nation (2001), page 27.

<table>
<thead>
<tr>
<th>Form</th>
<th>R</th>
<th>1. What does the word sound like?</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>P</td>
<td>2. How is the word pronounced?</td>
</tr>
<tr>
<td>Written</td>
<td>R</td>
<td>3. What does the word look like?</td>
</tr>
<tr>
<td></td>
<td>P</td>
<td>4. How is the word written and spelled?</td>
</tr>
<tr>
<td>Word parts</td>
<td>R</td>
<td>5. What parts are recognizable in this word?</td>
</tr>
<tr>
<td></td>
<td>P</td>
<td>6. What word parts are needed to express the meaning?</td>
</tr>
<tr>
<td>Form and Meaning</td>
<td>R</td>
<td>7. What meaning does this word form signal?</td>
</tr>
<tr>
<td></td>
<td>P</td>
<td>8. What form can be used to express this meaning?</td>
</tr>
<tr>
<td>Concepts and referents</td>
<td>R</td>
<td>9. What is included in the concept?</td>
</tr>
<tr>
<td></td>
<td>P</td>
<td>10. What items can the concept refer to?</td>
</tr>
<tr>
<td>Associations</td>
<td>R</td>
<td>11. What other words does this make us think of?</td>
</tr>
<tr>
<td></td>
<td>P</td>
<td>12. What other words could be used instead of this one?</td>
</tr>
<tr>
<td>Grammatical functions</td>
<td>R</td>
<td>13. In what patterns does the word occur?</td>
</tr>
<tr>
<td></td>
<td>P</td>
<td>14. In what patterns must we use this word?</td>
</tr>
<tr>
<td>collocations</td>
<td>R</td>
<td>15. What words or type of words occur with this one?</td>
</tr>
<tr>
<td></td>
<td>P</td>
<td>16. What words or types of words must we use with this one?</td>
</tr>
<tr>
<td>Constraints on use</td>
<td>R</td>
<td>17. Where, when and how often would we expect to meet this word?</td>
</tr>
<tr>
<td>(Register, frequency..)</td>
<td>P</td>
<td>18. Where, when, and how often can we use this word?</td>
</tr>
</tbody>
</table>
Questionnaire for TESL Teachers and Professors (Appendix B)

Your delicate, precise and kind answers are very much appreciated in advance.

Level of Education:

Major:

Native Language:

Years of Teaching Experience: Language(s) You Teach:

Familiarity with Any Other Foreign Languages:

Which Language(s): How Familiar You Are:

Word knowledge in the second language might have different aspects. In your opinion, which of the following aspects is important in knowing a word? To what extent is it important? Rate the level of its importance by putting a tick in the table based on the example.

18 items

<table>
<thead>
<tr>
<th>items</th>
<th>Very important</th>
<th>important</th>
<th>Moderately important</th>
<th>Less important</th>
<th>Not important</th>
</tr>
</thead>
<tbody>
<tr>
<td>example</td>
<td>✔</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>The way the word sounds like (auditory word recognition)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>The way the word is pronounced</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>The way the word looks like (with respect to spelling)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>The way the word is spelled and written</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Recognition of the word parts (suffix, stem, prefix) which make the word</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>Knowing and using necessary word parts (suffix &amp; prefix) to express the desired meaning</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>The meaning of the word</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>Knowing and using the appropriate word to express the desired meaning in the second language (For example, we use the word “X” to express the meaning of the word “Y” in the second language)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>Recognition of the components included in the concept which is represented by a word (For example, “flash, lens and button” are components included in the</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
|   | Production of the items the concept can refer to  
(For example, we use the words “book, desk and study” to refer to the concept of the word “library”) |
|---|---|
| 10 | Recognition of other words this word makes us think of  
(For example, synonyms, antonyms and any other words which come to our mind because of this word) |
| 11 | Production of other words this word makes us think of  
(For example, synonyms, antonyms or any other words this word makes us think of) |
| 12 | Recognition of the grammatical patterns in which the word occurs  
(Recognizing which part of speech of the word should happen in an appropriate place in a sentence) |
| 13 | Production of the grammatical patterns in which the word occurs  
(Production of the suitable parts of speech of a word such as verb, noun, adj and adv which completes a sentence) |
| 14 | Recognition of the words which are used with this word  
(collocations) |
| 15 | Production of the words that must be used with this word  
(collocations) |
| 16 | Where, when and how often this word is expected to be met |
| 17 | Where, when and how often the word is used |
| 18 |   |
Sample Reception Test (Appendix C)

Name: ________________________  Professor: __________________________

Term: ________________________  Course: __________________________

A. Listen and underline the word you hear.  Estimated time: 1 min.

Example: a. acquire  b. require  c. replier  d. retire

1. a. price  b. prize  c. pride  d. prime
2. a. seat  b. set  c. seed  d. sit
3. a. unable  b. enable  c. liable  d. disable

B. Underline the correct spelling of each word.  Estimated time: 2.5 min

Example: Analasis                      analysis                      analisis                      analyses
1. celebrate                      celebrate                      celebrait                      celebrat
2. approach                      aproch                        approach                      approache
3. necesary                      nasessary                      necessary                      neseccary

C. The following words are broken into their component parts. Underline the correct item.  Estimated time: 2.5 min

Example: variation

a. vari + ation  
b. varia + tion  
c. va + ria + tion  
d. vary + ation

1. irresponsible
   a. irres + pons + able
   b. ir + response + able
   c. ir + responsible
   d. irresponse + able
2. disagreement
   a. dis + agree + ment
   b. disagree + ment
   c. dis + agreement
   d. disag + reement
3. redefine
   a. rede + fine
   b. re + define
   c. re + de + fine
   d. redef + ine

D. Choose the appropriate Persian translation for the following words.  Estimated time: 2.5 min

Example: Numerous

a. مسئول  b. مسئولی  c. مشوق  d. مشوقی

1. Abuse
   a. تبخیرشدن
   b. سوء استفاده کردن
   c. داشتن انتعاش
   d. تاریخی
2. Nightmare
   a. غم
   b. روتا
   c. کبوس
   d. تاریخی
3. Deliberately
   a. هم‌نیاگر
   b. "عدم"یکدیر
   c. زمان یک در
   d. نا اگاهانه

E. Which of the following items do the words include? Choose the most appropriate group.  Estimated time: 5 min

Example: "camera"

a. lens, flash, button  b. bag, fabric, plastic
c. pipe, plastic, electricity  
d. button, picture, money

1. “race”
   a. winner, prize, compete  
b. happiness, game, environment  
c. hope, future, cure  
d. horse, attempt, doubt

2. “commerce”
   a. book, relationship, far  
b. sell, business, fabric  
c. buy, iron, business  
d. buy, sell, business

3. “college”
   a. university, desk, entertainment  
b. course, board, professor  
c. desk, board, traffic  
d. book, course, government

F. Choose the most appropriate group of words which you can think of in response to each word.
   Estimated time: 5 min

   Example: “sympathy”
   a. interest, friend, money  
b. kindness, mother, love  
c. whiteness, silence, pain  
d. understanding, enemy, heal

   1. “cancer”
      a. cough, pill, treatment  
b. life, cure, upset  
c. hope, future, cure  
d. death, pain, hospital

   2. “storm”
      a. movement, pleasure, coldness  
b. anger, shout, revolution  
c. shake, atmosphere, flood  
d. fear, shout, flight

   3. “battle”
      a. escape, joy, run  
b. fight, violence, agreement  
c. gun, hate, soldier  
d. hate, anger, understanding

G. Select the correct answer. Estimated time: 5 min
   Example: I have very fond .......... of my time in Spain.
   a. memorize  b. memorial  c. memorizing  d. memories (noun)

1. My company has ............ a new approach to staff meetings.
   a. adopting  b. adopted  c. adoption  d. adoptive

2. Unfortunately, scientists have been unable to find a .......... for this complaint.
   a. curable  b. cures  c. cured  d. cure

3. Eating fatty foods can damage your ............... 
   a. healthy  b. health  c. healthful  d. unhealthy

H. Choose the correct collocation for the underlined words. Estimated time: 5 min
   Example: Life in the past was much simpler as many people work to .......... their basic needs.
Today, for many of us, our job is not just a way of ________ a living.
a. having b. doing c. meeting d. making

For many of us, work ________ an important role in our everyday life and gives us strong sense of personal fulfillment.
a. does b. plays c. makes d. earns

Even in our leisure time we have to ________ so many choices about what to do or watch.
a. make b. take c. get d. set

There are 10 pairs of words in the following section. They have the same meaning but they are used in different contexts and situations (for example, different level of formality, different English speaking countries, different places, etc). Compare them and choose where, when and how often we encounter the first one which is bold. In other words, choose the appropriate and probable place, time and frequency of use of the first words. Estimated time: 5 min.

Example: “truck; lorry”
   a. in America, when two drivers are talking together, sometimes met
   b. in America, when a taxi driver is transferring a passenger, usually met
   c. in England, when two drivers are talking together, often met
   d. in England, when a taxi driver is transferring a passenger, usually met

1. “hell; damn”
   a. in an interview, when the interviewer is asking some questions, usually met
   b. at a conference, when the members are lecturing, sometimes met
   c. at a university, when the professors are teaching, often met
   d. on the phone, when two friends are talking angrily, sometimes met

2. “poll; voting”
   a. in the offices, everyday, always met
   b. in daily conversations, when the society is experiencing an economic crisis, usually met
   c. in newspapers, during an election time, rarely met
   d. in newspapers, during an election time, usually met

3. “blast; explosion”
   a. in a fire station, when the fire fighters learn how to fight the fire, always met
   b. in a fire station, when the fire fighters learn how to fight the fire, usually met
   c. in newspapers, when a huge explosion is reported, always met
   d. in newspapers, when a huge explosion is reported, rarely met
Sample Production Test (Appendix C, Cont.)

A. Read the following words.
   1. considerable
   2. accommodation
   3. succeed

B. Listen to the following words and write them down. Estimated time : 1 min
   1.
   2.
   3.

C. Add an appropriate affix (prefix, suffix) to each of the following words to express the meaning mentioned in parentheses. Estimated time : 5 min
   Example: Associate (organization) : associate + ion = association = organization
   1. Legal (against the law) :
   2. Child (stage of being a child):
   3. Suit (appropriate) :

D. Translate the following words into English. Estimated time : 2.5
   Example: Nوجوان = teenager
   1. نوجوان
   2. دھکاربودن
   3. رفتار

E. What items can the following words as concepts refer to? In other words, which concepts the following words include? Mention three items for each. Be aware that synonyms are not acceptable. Estimated time : 10 min
   Examples: "Music" : exciting, classical, live
            “library” : book, desk, study
            1. career
            2. collapse
            3. citizen

F. Write three appropriate associations or synonyms for each word. Estimated time : 10 min
   Example: abroad foreign country, travel, plane ticket
            Or Outside, overseas, foreign country
            1. authority
            2. purchase
            3. charity
G. Decide on the appropriate form of the word and write it in the spaces provided. Time: 7.5 min

Example: Children will always …**compete**… for their parents’ attention. *(compete) verb*

1. Even a small …**suggest**… can leave a piece of memory in the brain.
2. The rocket took off into the air and the crowd …**scream**… in a wild manner.
3. Why don’t you drop off my mother first and then come back and …**pick**… us up.

H. Fill in the blanks with the appropriate word (collocation). Estimated time: 7.5 min

Example: John made a mistake, and he should …**accept**… its consequences.

1. Women …**far less crime**… than men.
2. She tried to …**attention**… to what he was saying.
3. These chemicals have been found to …**serious environmental damage**….

I. There are 10 pairs of words in the following section. They have the same meaning but they are used in different contexts and situations (for example, different level of formality, different English speaking countries, different places, etc). Compare them and write where, when and how often we encounter the first one which is bold. In other words, write the appropriate and probable place, time and frequency of use of the first words. Estimated time: 10 min.

Example: “**bike**; bicycle”: in an informal conversation, when two friends are talking about their bicycles, sometimes met

1. “**thus**; so”
2. “**etc**; blah blah”
3. “**cool !**; great !”
Issues important in RT and PT design (Appendix D)

a. All items in RT were multiple-choice questions, and in PT, the participants were required to produce words or explanations. Item discrimination and item facility were run and necessary replacements were done.

b. Regarding item numbers 1, 2, 3, and 4, the target words with different syllables were selected. Furthermore, in RT (item 1), the words selected as distractors were harmonious.

c. Regarding item numbers 5 and 6 in both RT and PT, words with different affixes were selected. However, since words with affixes have low frequency, only adverbs with high frequency stems (3 or more frequency bands) were chosen. In PT (item number 6), the meaning or synonym of each target word which was supposed to be produced was written in parentheses. However, their Persian equivalents were added when the definitions did not seem understandable.

d. Regarding item numbers 7 and 8 in both RT and PT, target words with different parts of speech were selected. In PT, any probable correct translation was accepted. Only adverbs with high frequency stems were selected. Moreover, when meaning was the focus, almost always the first meaning of the words was considered.

e. Regarding item numbers 9 and 10, what was tested was what a word included. In other words, the features of a concept were considered as the answers rather than the synonyms. In these parts, the participants had to decide whether or not the written features existed in the concept. For example, button and flash exist in camera, and excitement as well as being alive exist in music. Therefore, the participants were expected to select such features and not to provide synonyms for the words camera and music. Moreover, while in RT the best possible and most appropriate group of answers was considered to be chosen, in PT, any possible answer was accepted.

f. Regarding item numbers 11 and 12, what was tested was word association, i.e. the words which come to our mind because of a special word. For example, the associations of the word camera are party, celebration and scenery. However, based on what Nation believes (communicated via e-mail), synonyms can also be considered as associations of the words. As a result, synonyms mentioned in PT were also accepted. In RT, the participants had to choose the most appropriate group of words.

g. Regarding item numbers 13 and 14, grammatical patterns (parts of speech) of the words were tested. The participants needed grammatical knowledge to find the correct answer. Therefore, the frequency of the words making up the sentences was not controlled, although simple sentences were chosen. The correct answers were selected from among four important parts of speech (adjectives, verbs, nouns and adverbs).

h. Regarding item numbers 17 and 18, the equivalence of the target words was added beside them, and the participants were asked to compare each pair and choose (receptively) or write (productively) where, when and how often they encounter the target words, which were bold and italicized.
“A Small Clean Warm Plastic”: Understanding EFL Autobiographical Writing and Identity

Shizhou Yang

Yunnan Nationalities University, China

Bio Data
Shizhou Yang received his PhD from La Trobe University, Australia. His present research interests include identity, literacies, autobiographical writing and narrative practices in minority educational contexts. He currently lectures at Yunnan University of Nationalities, Kunming, China.

Abstract
In the past two decades, autobiographical writing has grown into a powerful way of informing Second Language or L2 learners’ identities, but not yet in the EFL context. Focusing on Anne, the writer of “An unforgettable event in childhood” to be quoted partially below, this case study—as guided by a postmodern framework through a multi-storied approach to narrative analysis—explores the impact of EFL autobiographical writing on identity. Anne was an ethnic Bai English major student from a poor family in the countryside of southwest China. Prior to the study, passing examinations was her primary reason to write in English. As a member of an autobiographically-oriented extracurricular writing group, she wrote, rewrote and shared her stories, especially of poverty and family, with other group members. In this process, Anne re-authored her relationship with her mother by reinventing as creative art her painful childhood poverty experiences in which her mother had played a focal role; and became able to come to terms with / accept her mother and her childhood self. Through autobiographical writing, she also became a more confident and expressive writer in English. These findings suggest that EFL autobiographical writing in an appreciative writing group may position struggling writers as knowers and capable learners and contribute to rich understandings of their social identities.

Keywords: autobiographical writing, identity; EFL, narrative, re-authoring

Introduction

An unforgettable event in childhood

I walked out, my mother took out a clear piece of plastic quickly under her coat and gave me. I received it and ran into my seat without any words. I tucked it into my desk very angrily. After class, I didn’t go until I made sure everyone had left. It still rained heavily. I had to wear the plastic. I ran home as quickly as possible. Tears ran out of my eyes. (Anne 1, May 9, 2008)
EFL autobiographical writing and identity

In the past two decades, identity issues of L2 learners have become a frequently featured research area in the fields of Applied Linguistics, Second Language Acquisition (SLA) and L2 literacy, particularly in studies conducted in the ESL context (Burke, 2011; Hirvela & Belcher, 2001; Ivanič & Camps, 2001; Lam, 2000; Menard-Warwick, 2006; Norton, 2000; Ouellette, 2008; Pavlenko, 1998, 2001, 2007; Shin, 2010; Solé, 2007; Thompson, 2011; Yi, 2009). Adopting a sociocultural approach, these studies have contributed to a rich understanding of L2 learners as complex social beings who learn a L2 purposefully to construct desirable social identities. As such, L2 learning should no longer be treated as a process of acquiring linguistic knowledge alone. Instead, it also involves L2 learners’ constant reconfiguration of their sense of self through their use of the L2. However, because most of the afore mentioned studies were situated in ESL context, whether their findings about L2 learning applies to the EFL context remains a topic of debate. Some scholars (e.g., Block, 2007, 2008) seem to suggest that the EFL context is a futile ground for L2 learners to reconstruct their identities through L2. Others (e.g., Gao, 2007, 2009; Gu, 2010; Norton Peirce & Kamal, 2003; Peirce & Kamal, 2003), in contrast, seem to hold that identity issues are relevant to EFL learners as well. Addressing this issue is important because it may help EFL professionals around the world to understand the nature of their profession: whether just as a skill training activity or as a more complex identity-moulding business.

In this paper, I attempt to participate in the debate by focusing on the following question: What effects does Anne’s active engagement with EFL autobiographical writing in an extracurricular writing group appear to have on her identities? Compared to previous studies on L2 literacy and identity, the present study has three innovative features. First, the study’s main data type – an EFL student writer Anne’s educational autobiographical writing – makes the study particularly relevant to EFL literacy education. Previous autobiographical studies mainly used published L2 memoires from the ESL context (see Casanave, 2005; Li, 2007; Pavlenko, 1998, 2001, 2007; Steinman, 2005) as data. Although they show powerfully that L2 writers often reconfigure their identities when they write in English as an additional language, as H. Nicholas (personal communication, September 27, 2010) suggests, such autobiographical writing is not educationally oriented and has only limited capacity to inform L2 learning in educational contexts. Moreover, like Shen’s (1989) autobiographical account critiqued by Stapleton (2002, pp. 180-181), they were tinted with the socio-political
ideologies of a particular era, often pre-2000s. Therefore, such studies are far from offering a clear window into the actual processes of language acquisition and use by EFL learners today. Second, the present study also fills in a gap of knowledge about L2 autobiographical writing and identity, two lines of inquiry that rarely intersect. On the one hand, most researchers tend to use EFL learners’ autobiographical writing for non-identity-related research purposes. Often, EFL autobiographical writing is used as a tool to evaluate particular pedagogies, such as story grammar (Alroe, 2011; Amer, 1992; Cheng, 2008; El-Koumy, 1999; Ellis, Sheen, Murakami, & Takashima, 2008; Nilforooshan & Afghari, 2007) or as linguistic data to demonstrate certain textual features, e.g., the use of reference, typical to a particular group of L2 learners (Genc & Bada, 2006; Kaminura & Oi, 2001; Kang, 2004, 2005, 2006, 2009; Lee, 2003; Lingley, 2005), to name just a few. On the other hand, L2 research on student writers’ identities often draws on academic writing instead of autobiographical writing (e.g., Burke, 2011; Helms-Park & Stapleton, 2003; Matsuda & Tardy, 2007; Ramanathan & Atkinson, 1999; Thompson, 2011). As a result, it remains unclear whether L2 learners may engage in significant identity work through EFL autobiographical writing. Third, the present study offers a rich description of an EFL student writer’s identity (re)construction process. It involves the student’s non-test-oriented autobiographical writing for over nine months, yielding a multi-storied understanding of her identities. In addition, the study draws on multiple data types such as drawings and interviews. Together, these features make the study potentially interesting to those who are engaged in using narrative-oriented approaches to L2 literacy research (e.g., Casanave, 2003, 2005; Coffey & Street, 2008; Solé, 2007) and those seeking new possibilities for L2 literacy education, both in the ESL and EFL contexts (e.g., Cummins, 2005; Leki, 2001; Norton, 2009; Pavlenko, 2001; Tran, 2007).

Language, learning and identity

Similar to Pavlenko (2001), I take ‘postmodern’ perspectives on language, learning and identity. I see language as a mediating tool (Vygotsky, 1978), a form of symbolic capital (Bourdieu, 2004, 1991), and learners learning a language as an ‘investment’ in their desired social identities (Kanno & Norton, 2003; Pavlanko & Norton, 2007; Peirce, 1995). Learning, including learning non-primary languages, is always sociocultural, involving movements from doing with others (intermental) to doing by oneself (intramental), and always with some kind of cultural tools such as speaking, drawing, writing etc. (Lantolf, 2000; Vygotsky, 1978).
For these reasons, learning is best approached as a situated activity (Lave & Wenger, 1991; Wenger, 1999), a mediated action (Wertsch, 1995, p. 64) involving participation in ‘communities of practice’ (Wenger, 1999) with consequences on the learners’ identities (Hall, 1997; Wenger, 2006).

According to Pavlenko (2001, p. 319), these postmodern ideas of language, learning and identity have three main implications for SLA. First, second language acquisition is a socialisation process in which L2 learners, understood as agents with multiple and adaptable identities, construct their new social identities. Second, L2 learning and learners’ identities constitute each other. On the one hand, learners’ previous identities ‘mediate their access to linguistic resources and interactional opportunities.’ On the other hand, identity possibilities in L2 shape language learners’ ‘agencies and investments in language learning and use.’ Third, L2 learners’ self- and / or other-perceived sense of self may be reconfigured by the learners’ participation in new communities of practice mediated by the L2.

I regard identity as a dialogical notion (e.g., Hermans, 2001; Pietikäinen & Dufva, 2006; Silseth & Arnseth, 2011). That is, besides the multiplicities and changes of subjectivities, as espoused by poststructuralist scholars (e.g., Block, 2007; Davis, 2011; Momenia, 2011; Norton, 2010; Park, 2009; Peirce, 1995; Weedon, 1987), identity also entails continuity of meanings through narration (Bruner, 1991; Holquist, 2002; Menard-Warwick, 2005). After all, identities, to the degree they are storied and made coherent, have real effects on people’s lives (Eakin, 1999; White, 2007), as can be observed ‘from the experiences of those who are alienated and oppressed, and those suffering mental illness’ (Mackenzie, 2008, p. 16).

Continuity of identity, i.e., a person maintaining the same position in different social domains, is important for L2 teaching practice. Several studies (Cummins, 2005; Heath, 1993; Iddings & Katz, 2007; Kibler, 2010; McKay & Wong, 1996; Norton, 2009) have shown the pedagogical effectiveness of allowing L2 learners to continue their cherished identities, whether in relation to dance, drama or their heritage languages. Norton’s own study has also shown the reverse to be true: if L2 teaching practice was not congruent with L2 learners’ social identities, non-participation may arise (Norton, 2001).

Two perspectives may be helpful in understanding the continuity of identity. The first is a historical perspective. According to Ivanič (1998, pp. 23-24, 181-186), one dimension of writer identity is autobiographical self, i.e., the writer’s evolving life history of encounters
with others up to the moment of writing. As such, a writer’s identity may contain certain enduring facts. The other is a narrative perspective, which points to the essential role played by narrative or story in identity making. Stories contain two landscapes, one of action and one of consciousness (Bruner, 1986) or identity (White, 2005). The landscape of action refers to terrains of sequentially linked events and experiences in one’s life, which are emplotted. The landscape of consciousness refers to, at least partially, identity conclusions people draw in relation to themselves and others by mapping culturally specific categories to the landscape of action (White, 2005, p. 61). People use stories to construct meanings, to represent parts of their lives, and once performed, i.e., with others as audiences, stories may reconstitute people’s lives and relationships (Menard-Warwick, 2007; White & Epston, 1990, pp. 9-11). People also use stories to bring coherence to their past, present and future selves (Eakin, 1999) and values at different times (Holquist, 2002, p. 37). Thus, it is fitting to see, as Bakhtin did, self as ‘constituted as a story, through which happenings in specific places and at specific times are made coherent’ (Pavlenko & Lantolf, 2000, p. 158).

Materials and Methods
This study employed narrative analysis (Chase, 2008; Clandinin & Connelly, 2000; Josselson, 2011; Pavlenko, 2007; Polkinghorne, 1995a), particularly a multi-storied analytical approach (Grant, 1997, 2001), assisted by ethnographic and case study methods (Heath, Street, & Mills, 2008; Stake, 2003).

Group composition, meetings and the case study participant
In mid-March 2008, 19 student volunteers and I formed an extracurricular writing group at Lakeview University in southwest China. The student members represented three year levels—one, two and three—and were all English majors at the University’s Foreign Languages College. Among them, fourteen were Bai and seventeen were female. Fourteen student participants were ethnic Bai, two Yi ethnic students and three Han Chinese. Most of the student members had encountered frustrating or failing experiences in either or both Chinese and English literacy before joining the writing group.

The writing group had a strong autobiographical orientation. Half of the group time–16 weeks from March to July 2008, was devoted to autobiographical writing. During the second half of the group time, 16 weeks from August to December 2008, although the focus shifted to argumentative writing, the members were sometimes encouraged to re-write their
stories into short poems. Moreover, I frequently shared my own autobiographical writing with the members.

The group met once a week, each meeting usually lasting about one and half hours. On average, nine members attended each meeting (based on attendance records). At the meetings, the student members and I often sat in a circle in an empty classroom of the Lakeview University to share each others’ writing samples and discuss writing-related topics. One featured activity of the writing group was “oral publication,” i.e., the group members reading what they regarded as their best writing samples to the group and receiving comments from others. Oral publication sessions, three in total, were held at my home, attended by visitors and accompanied by food. Some visitors were foreigners like my American wife Cathryn and my Australian supervisor / advisor Audrey.

The writing group arrangement played a pivotal role in my understandings of EFL autobiographical writing and identity. First, the arrangement allowed me to work primarily as a group leader, member and co-writer rather than a teacher. Second, the arrangement allowed the students to participate in the research on purely voluntary basis. Third, the arrangement allowed me to generate non-test-oriented writing samples. Fourth, the arrangement allowed me to know the student writers more intimately than would be possible in a regular size class of 50 students.

Data were collected after the student members had given me explicit written consent. They include EFL narratives, argumentative essays, poems, letters, emails, short-text messages, drawings, transcripts of life story interviews, text-related interviews, and video / audio recordings of meeting discussions etc.

Anne, the case study participant in this study, was one of the ten first-year English major students in the group. She spoke Bai at home with her parents and Chinese or English at school (1 I, April 25, 2008). Both of Anne’s parents were “文盲 (illiterate)” farmers (30W, 2 I, May 30, 2008). Like other case study participants, Anne could not write the Bai script, nor had she had any opportunity to learn it. Anne was chosen as a case study participant for her ethnic Bai background, frustrating experiences in both Chinese and English literacy, and active participation in the writing group activities, i.e., attending 24 of the extracurricular group meetings, frequent involvement in group discussions, and producing more than ten English autobiographical narratives.
Analytical approach

I adopted a multi-storied approach (Grant, 1997, 2001), akin to Pavlenko’s (2007) macro- and micro- analysis of multilinguals’ narratives, to analyze Anne’s autobiographical narratives and identities. This approach is based on two broad propositions: that humans are interpretive beings (Geertz, 1973; Holquist, 2002) and that narrative is one of the essential tools to construct coherent meanings out of life experiences (e.g., Bruner, 1986, 2004; Polkinghorne, 1995b). In other words, a multi-storied approach treats ‘life as narrative, persons as the storytellers’ and narrative analysis as ‘the interpretive readings’ of ‘narratives and identities’ from multiple perspectives of context (Grant, 1997, p. 34). More specifically, the personal context foregrounds learners as persons with histories, from particular localities and endowed with local knowledge; the social context foregrounds learners as persons constructed by society; the global context foregrounds learners as active contributors to one shared world; and the transformative context foregrounds futures of new and better possibilities (Grant, 2001, pp. 12-13).

By valuing the personal context, a multi-storied approach builds in itself a defence against critical analyses that leave individuals’ agency out of the scene, as discourse analysis sometimes tends to do (see Grant, 1997; Thesen, 1997 for critiques). By suggesting a transformative context, which integrates the personal, social and global contexts, a multi-storied approach treats the four layers of context as they are: interpenetrating and working simultaneously.

Analytical procedure

My analysis of Anne’s data included four main phases. The first phase (March – December 2008) was the identification of EFL autobiographical writing samples. During this process, I looked for textual features such as the use of “I”, “my”, “mine”, and “me” (e.g., “My Father”) to determine whether a particular piece of writing met the minimal definition of ‘autobiographical writing’, i.e., any writing about any aspect of one’s own life experiences. There is an inevitable cultural bias in relying on such textual features alone. For cultural, sociopolitical and or ideological reasons, writers in oriental societies may be less comfortable with the frequent use of self-referent singular pronouns (Kashima & Kashima, 1998; Matlene, 1985; Morling & Lamoreaux, 2008; Na & Choi, 2009; Ramanathan & Atkinson, 1999; Shen, 1989) than those in western societies where individualism seems to be the norm (Atkinson, 2001). However, in the context of this particular writing group, the use of first person...
referents is a valid indicator because it has been frequently modelled. For instance, some of my writing samples that I shared with the group members as an example (e.g., “My First Picture”, “When I Am Serious, God Is Romantic”) showed to the members that they could and were encouraged to write autobiographically by using “I”, “my”, “mine”, and “me.” The result of such an analysis provided textual evidence to show the extent to which Anne engaged in EFL autobiographical writing in the group.

The second phase, which parallels with the first phase, also involved the preliminary content analysis. During this phase, I mainly coded Anne’s EFL autobiographical writing samples ‘according to emerging themes, trends, patterns, or conceptual categories’ (Pavlenko, 2007, p. 166). The purpose was to examine if Anne engaged in any identity work through EFL autobiographical writing. Following postmodernist theories of identity, language and learning, I assumed that EFL student writers’ identities are relational, multiple and evolving and that the writers play an agentive role in the changes and continuities of their own identities (Eakin, 1999; Menard-Warwick, 2005; Peirce, 1995; Thesen, 1997). I also assumed that EFL student writers’ identities are informed by culturally framed categories and descriptions (Bruner, 1986). Accordingly, while reading and giving written comments to the group members on their writing samples, I attended closely to Anne’s descriptions of those with whom they were closely related (e.g., “I…even felt her ugly and disgusting). I also attended closely to different descriptions Anne gave to the same or similar experiences in different writing samples (e.g., “But I now learn to taste her voice and cherish it”). In identifying and selecting other relevant evidence, I was guided by my previous understandings of the key concepts such as voice, i.e., uniqueness in content, expression and structure (Yang, 2007), discourse, i.e., communal norms (Gee, 1990, p. 143) and literacy events, i.e., text-mediated interactions and interpretations (Heath et al., 2008, p. 102). I adopted two tools for the preliminary data analysis. The first tool was a reflective journal, which I kept throughout the writing group activities (March – December 2008). The second tool was the NVivo 8 software. Its “Annotation”, “Coding” and “Memo” functions allowed me to code the digitalised EFL autobiographical writing samples.

The third phase (February – December 2009) was the contextual analysis. Its main purpose was to contextualize Anne’s investment in EFL autobiographical writing, the roles played by the writing group in both her EFL autobiographical writing and identity work and the nature of autobiographical writing in the writing group. The primary data informing this
contextual analysis were Anne’s interview transcripts, the selected transcripts of video recordings from the group’s weekly meetings and some of Anne’s own contextual descriptions elicited through interviews. A multi-storied approach to narrative analysis (Grant, 1997, 2001) guided this analytical process. In understanding the personal context, I attended closely to what Anne said she was doing through her writing samples. In understanding the social context, I paid attention to how social others, particularly those in the writing group, shaped the writing and revision of Anne’s particular writing samples (e.g., the relationship between Anne’s title “God has tricked me” and my title “When I Was Serious, God Is Romantic”). For this purpose, I paid close attention to the linguistic features of Anne’s texts (e.g., the use of personal pronoun, verbs and modal verbs) and possible dialogical relations between these texts and other texts. I then interviewed Anne to confirm such relations. In understanding the global context, I attended closely to cultural elements that joined the local with the global (e.g., Anne’s writing sample about the ever growing gap between the countryside and the city). In understanding the transformative context, I attended closely to the interconnections between the personal, social and global contexts and imagined possibilities rooted in these interconnections.

The fourth phase was narrative analysis. The purpose was twofold: to locate if there were any movements in Anne’s identities and to holistically represent these identities as entailing both ‘continuity and change’ (Menard-Warwick, 2005, p. 253). During this phase (April –2009 – October 2010), I wrote Anne’s data into stories in line with Polkinghorne’s (1995b, p. 15) notion of narrative analysis, i.e., ‘the procedure through which the researcher organizes the data elements into a coherent developmental account.’ As such, my analytical focus shifted from the interpretation of discrete pieces of data to ‘the configuration of the data into a coherent whole’ (Polkinghorne, 1995b, p. 15). One strategy I found helpful was starting Anne’s story with a storyline and organising most of my analysis chronologically. In drafting Anne’s storylines, especially in storying her movements in her identity as a writer, I found her drawings, labels and explanations at the end of the writing group activities particularly useful as a reference.

Influenced by a sociocultural theory of learning (Kutuzova, 2010; Vygotsky, 1978), I adopted three strategies to ensure that I have produced the most plausible interpretations of my data. First, I shared my writing with others. My expanded audience mainly included Anne, some classmates at La Trobe University, with whom I met weekly for a year, my three
supervisors, my wife Cathryn, Dr. Bernie Neville’s monthly research group, several conference audiences, and anonymous reviewers of three academic journals. Like scaffolds, these interactions helped me to see the real complexities of my data, their alternative interpretations and take positions based on evidence. Second, I tried to be open to critiques. Connelly and Clandinin (1990) encouraged narrative inquirers to take every critique seriously. I engaged in such a process especially when I was comparing Anne’s data with other case study participants’ data – resulting in a chapter of my PhD thesis. Knowing that I was neither an expert in narrative analysis, nor one in narrative therapy, I frequently consulted my supervisors regarding the use of narrative terms. Third, I rewrote my drafts frequently. This process helped me to build on my earlier drafts and interactions with others. In particular, it challenged me to be clear about my arguments and necessary evidence, which often resulted in re-analysis or expanded analysis of my data. Taking into consideration my interactions with others in one way or another, I regard my data analysis as a scaffolded learning process.

Results and discussion

Previous literacy experiences

During the writing group activities in 2008, I interviewed Anne thirteen times, covering topics on her life, literacy experiences and processes of writing specific samples. Below is a summary of the interviews most relevant to this article.

<table>
<thead>
<tr>
<th>Code</th>
<th>Main topics</th>
<th>Date</th>
<th>Main language</th>
</tr>
</thead>
<tbody>
<tr>
<td>Interview I  (1 I)</td>
<td>Life experience themes</td>
<td>April 28</td>
<td>English</td>
</tr>
<tr>
<td>Interview II (2 I)</td>
<td>Literacy experiences</td>
<td>May 30</td>
<td>Chinese</td>
</tr>
<tr>
<td>Interview IV (4 I)</td>
<td>Beliefs</td>
<td>September 3</td>
<td>Chinese</td>
</tr>
<tr>
<td>Interview V (5 I)</td>
<td>“优秀论文” (Model Essay)</td>
<td>September 3</td>
<td>English</td>
</tr>
<tr>
<td>Interview XI (11 I)</td>
<td>Mentor’s assignments</td>
<td>December 9</td>
<td>Chinese</td>
</tr>
</tbody>
</table>

Interview II with Anne revealed her test-oriented learning history concerning English writing. She said that soon after she started learning English in middle school –

all my writing in English was done for examinations. All the teachers said–gave you a fixed format, such as you should, should, you’d better write in three parts: beginning, conclusion and middle.²
When asked about what she remembered most about her English classes from middle school and high school, Anne said it was copying keys to tests, including model essays: “Often after each [English] test, the teacher would hand out the papers, and then ask us to copy them.” As an English major student, Anne regarded writing as important “because TEM-4 mainly depends on listening comprehension and writing.”

Anne’s test-oriented learning history was most vividly shown by her drawing during the last writing group meeting—a “small and weak” tree with “a few fruits” (32V, December 20, 2008)–to represent herself as a writer before joining the writing group (see Figure 2). She explained in her caption “most time I write just for examinations:”

![Figure 2 Anne’s first tree](image)

Anne’s words and drawing give a glimpse into her ‘autobiographical self,’ which, according to Ivanič (1998), is as stable as habitus (p. 24). That is, “writ[ing] just for examinations” had been Anne’s habitual way of relating to English writing.

**Anne’s autobiographical writing samples**

Table 1 gives a summary of EFL narratives written by Anne when she was a writing group member. She read some of them to the writing group and revised some others. All titles were chosen by her.

Table 1: Anne’s autobiographical writing samples

<table>
<thead>
<tr>
<th>Event(s) or experience(s)</th>
<th>Others involved</th>
<th>Title</th>
<th>Context</th>
<th>Code</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>God</td>
<td><em>Life river and chapter</em></td>
<td>During the 3rd meeting, drawing with</td>
<td>3D/W W</td>
</tr>
</tbody>
</table>
These writing samples show that Anne engaged with her family and poverty experiences extensively. My analysis below mainly focuses on 4W_W, 11W_W, 9W_W and 14W_W. The reasons are twofold. First, these samples share a general theme of the poverty experiences that had filled Anne’s childhood, high school (1 I) and university life (4 I). As such, they may illustrate two important dimensions of Anne’s identities: her evolving understandings of her life and a sense of continuity (Giddens, 1991; Holquist, 2002; White & Epston, 1990, pp. 9-10). Second, these samples have rich contextual information. The ‘literacy events’ (Heath et al., 2008) around these samples, e.g., “oral publications” at group meetings, email correspondence, and interviews, allow me to probe both their dialogical relations (Bakhtin, 1981) and what Anne says she was doing (Thesen, 1997, p. 504) through them.

“An unforgettable event in childhood” (4W_W)
Before she started “An unforgettable event in childhood” about the plastic sheet event, briefly cited at the beginning of this paper, Anne wrote (see Extract 1):

Extract 1: “An unforgettable event in childhood” (4W_W)

At the age of eight, I was a carefree and invinsible girl. I was very boisterous... and
neve listened carefully to my mother, sometimes even felt her ugly and disgusting, because she didn’t meet my needs. For example, I had no lovely clothes, no toys, no candies and all. (¶2, emphasis added)

According to the text, poverty experiences shaped Anne’s relationship with her mother and consequently her identity as a daughter.

Anne’s representation of her mother should be read in reference to three social contexts: the writing group, her family and the Foreign Languages College. At my invitation, Anne shared her story “An unforgettable event in childhood” with the writing group on May 23 at the 11th meeting (attended by eight student members). Her description about her mother as “ugly” immediately led to a heated discussion. She had to give additional reasons for using such a taboo word about her mother: “Because she’s very short, small, I always think my mother [is] not like other mother.” Unconvinced, another group member, Peter said laughingly, “But mother [is] still mother!” Another member, Mary joined in, suggesting that Anne should change “ugly” to something like “not beautiful.” In one way or another, others suggested that Anne drop “ugly.” There were only two exceptions. Hope regarded it as normal for an eight-year-old to see her mother as “ugly.” Abby commented that “ugly” does not describe a person’s appearance, but character (11V). A week later, when I asked Anne how she felt about other members’ comments on her word choice, she maintained her stance, adding that “in countryside most people … are not pretty or beautiful than the people in city. Because they are very, very tired” (2 II). Anne’s story (4W_W) only hinted at her present position towards her mother, which was “I feel grateful for my mother,” as she said in response to Matt’s question.

The group members’ interactions showed that while multiple ways of representing Anne’s mother existed, the socioculturally dominant ways seemed to favour the nice way. In addition, they revealed Anne’s other associations of her mother as a member of the “short, small” and “very tired” country women farmers, which informed Anne’s representation of her.

The second social context Anne’s family was revealed at the beginning of her Chinese report “一道永不愈合的伤口” (A wound never to be healed) (see Extract 2):

Extract 2: “一道永不愈合的伤口” (A wound never to be healed)
During my childhood, Younger Sister and I often nagged Mother to tell us stories. But Mother was illiterate, unable to tell any exciting stories. So she told us about her childhood life.

She said life used to be very hard. People often had to eat wild plant--now used to feed pigs only--and steamed buns made from rice chaff. The best food they could have was steamed corn buns. If anyone had any rice, washing and cooking it had to be totally in secret. Some people died of constipation after eating rice chaff. Some of them starved to death. Families had little clothing but many children. So, clothes were like treasure, worn first by the older children until they were outgrown and then handed down to the younger ones. She said although life was hard, people were diligent and capable rather than lazy. Mother said when she was young, she had to both study and do housework, such as washing clothes, cooking, farming, and gathering firewood in the mountain. She was indeed “grabbing with both hands”!

Earlier ethnographic studies (e.g., Heath, 1982; Heath, 1983) showed that children are socialised into their parents’ ways of using language and telling stories. In “An unforgettable event in childhood” (4W_W), Anne foregrounded her poverty experiences as did her mother during her childhood. More importantly, like the writing group, Anne’s family also channelled to her sociocultural values of the wider society in China, e.g., values associated with loyalty to one’s own family. Note that Anne’s mother’s stories were followed by a description of herself as one of the people from the hard times, who “were diligent and capable and not lazy.” The housework her mother did, e.g., “washing clothes, cooking, farming and gathering firewood in the mountain,” exemplified her contributions to her family. Anne’s life revealed similar evidence. She started doing housework early, “就一年级吧……还是一年级都不到的时候我就已经学会什么做饭啊,洗衣服那些” (I learned to do things like cooking and washing clothes when I was in the first grade, or not even the first grade, 1 I). When I asked her about her biggest challenges in learning English, she responded, “我就感觉吃苦我不怕” (I’m not afraid of working hard, 1 I). At university, she paid her tuition and board with support from her younger sister, a migrant worker in the capital city of her province, and applied for a study loan (4 I). She also worked in her university cafeteria for a while to finance her studies, but she saved most of her money for her parents. “反正回去那种
我一般都是给她们钱的，自己只留一小点” (Every time I go home I usually give them some money, leaving only a small amount for myself, 4 I). Therefore, in ways of using language, telling story, as well as valuing, Anne was as the proverb says, ‘Like mother, like daughter.’

The third social context is Anne’s university department, best understood through its policies and practices. Anne joined the writing group when her department had implemented a “导师制”–mentoring system–for the second year. The system required that each teacher from the Foreign Languages College should guide about five students in the department in how to study English. The mentors were supposed to meet their students regularly and ensure, among other policies, that by the time their students graduated, they would have written “50 passages” in English, which was one of the “五个五十” (Five by Fifty) requirements (based on a departmental public notice, October 17, 2008). Although mentoring was a major responsibility for the teachers, its practice seemed to have been compromised by many factors: the extent to which the teachers took it seriously, the pressure for them to engage in research and get published on top of the usual heavy teaching loads for all the English teachers from the Foreign Languages College. For instance, one young teacher, about thirty years old, complained during an interview that he had virtually become “a teaching machine” (Interview on April 7, 2008). From Anne’s perspective, her mentor was a poor example of mentoring. Under this mentoring system, Anne wrote several passages, including “God made a joke with me” (1W_M), a thematic predecessor of “An unforgettable event in childhood” (4W_W). But Anne said her mentor was “不怎么负责” (not so responsible), because “写几篇作文，做什么都以我们为主” (It’s all up to us as to how many compositions to write and what to do) (11 I). By the end of the writing group, Anne had written ten passages for her mentor but had received no comments and assumed none of what she had written had been read (11 I).

Among these three social contexts, the writing group played a significant role in shaping Anne’s writer identity by its “literacy events” (Heath et al., 2008). Anne’s essay was a ‘performed story’ (cf. White & Epston, 1990), i.e., a story shared with an audience in the writing group. After I read it, I commented in an email to Anne (May 22. 2008): “I admire your courage to tell the story,” it is “powerful,” “painfully beautiful,” showing the “demoralising power of poverty.” Anne’s email response (see Extract 3) two days later, May 24, the day after she read her story to the writing group, was filled with joy for being “prised².”
I take “prised” to mean “praised.” During Interview II on May 30, 2008, Anne explained to me about this email, saying that she was “happy” because it was “the first time my composition is being praise, prized, praised” (2 I). Her explanation suggests that she was in the process of correcting herself.

Writing and sharing “An unforgettable event in childhood” with the writing group was a landmark event for Anne. It was the first time that she had written about this poverty experience, the first time she’d read her writing to a group of people, and the first time of “being praise[d]” for her English composition. A week later, at the time of interview, she regarded this story her best (2 II).

About five weeks later, Anne shared another best piece, “The first time to hear her voice in phone” (9W_W), with the writing group.

“My first time to hear her voice in phone” (9W_W)
The focal event described in Anne’s next story took place when Anne was in high school. She did not know that her father had been hospitalized until her uncle called her on the phone. When her uncle asked her to speak to her mother on the phone, she wanted to refuse, “But too late, she got the phone.” Thus began Anne’s first time to hear her mother’s voice on the phone. Due to her mother’s hearing problem (“Both Dad and Mom have poor hearing and it’s hard to explain on the phone. Even at home when we talk we have to yell,” 1 I), the telephone call was a one-way conversation featuring her mother’s “penetrat[ing]” voice and Anne’s silence. However, as revealed in her story, the event proved crucial in the evolution of Anne’s relationship with her mother.

Three major events contextualised Anne’s title. One was reading a text at the 14th meeting on June 13, My First Time to Say “I Love You!” about a Chinese university student’s confession of love for his sweetheart on the phone. At the meeting, we first guessed the contents of the story based on the title, and then took turns to read the story and make
comments on it (14Vii). Anne’s comments show that she was impressed by the story’s “title,” “organisation,” and use of “conversation,” which she regarded as an example of “voice” (14V). She said, “Here the writer used conversation ((looking up at me)) (I: uh huh, uh huh, conversation) ((looking at the handout again)) (I: uh huh) and I think, this is one of his voice in this article ((laughing)))” (14V).

Another event contextualising Anne’s text was reading my essay “Unleash Your Writing Voice,” which was intended to encourage the members to write with uniqueness. This was the second time I discussed the notion of voice as uniqueness with the group. To make sure that the members understood what I meant by “voice,” I first read my essay in English, and then explained it in Chinese (14th meeting plan, 14V). In my essay, I used Anne’s “An unforgettable event in childhood” (4W_W) to illustrate using personal experiences as unique content.

The unique content is part of the voice. Remember Anne’s story in which she recalls how she experienced poverty as an eight-year-old…. Many, including me, have experienced poverty. But Anne’s account, revealing the imprints made on her by the storm, her classmates’ ‘lively little umbrellas and water shoots’³, her mother’s appearance with two pieces of plastic, has breathed her personal voice—in terms of unique content—into her poverty story. Thus, a writer’s voice is heard where the writer notes specifics of events: both the happening and the personal interpretation.

A causal relationship did not exist between reading my essay on voice at a group meeting and Anne’s title “My first time to hear her voice on phone.” Nonetheless, it seems fair to suggest that the group context encouraged Anne to keep drawing on her personal experiences, as she had done in “An unforgettable event in childhood” (4W_W).

The third event was the group being invited to write with uniqueness. During the 14th meeting, after we had finished reading and commenting on My First Time to Say “I Love You!,” I asked the members to write “with uniqueness” on the topic “my first time to…” (14V). In response, Anne first wrote—in red—about seeing her father’s tears. After reading other members’ samples, she first used green ink to make some minor revisions to her original text and then wrote in green at the bottom what became the title of her full piece (see ”).

Extract 4), “My first time to hear her voice in phone.”
Two cues from this draft and other texts are important. First, Anne did not cross out the original title about seeing her father’s tears. Second, according to “My Father” (1W), “the day when I saw my father’s tears for the first time” was “when I was in Grade Two in high school.” Hearing her mother’s voice for the first time on the phone took place one year earlier, when Anne “was a freshman in senior high school,” as the completed story will show. The two cues suggest that Anne was actually re-starting with a new title.

Two weeks passed. We were now approaching the end of the first stage, the 16th meeting on June 28. To celebrate, I encouraged the members “to orally publish a piece of writing you are most satisfied with” at the meeting (handout of the 14th meeting). In response, Anne finished “My first time to hear her voice in phone [on the phone]” and emailed it to me, asking for “help before I show it off” (see 錯誤！書籤的自我参照不正確。5).

Extract 5: Anne’s email on June 25

This week I wrote an article, but I need your help before I show it off. I hope you can give me some advice.

I started my comments by appreciating her story (“Such a turning point… is worth exploring and writing about,” 26W). I then responded from a reader’s perspective (e.g., “Your
first paragraph successfully keeps my attention with you. But the second paragraph somewhat interrupts that flow. What if you continue with your uncle’s asking you to speak with your mother and your no-choice situation instead?”). I also made it “optional” for Anne to consult a handout on sentence structures and consult the usage of “image vs. imagine, hung.” Anne integrated some feedback, such as changing ‘image’ to ‘imagine’ and added (though it could have been made more explicit) why she kept silent, “because I knew no matter how hard I shouted, she never could hear my voice.”

The meeting where the stories were to be shared orally was joined in by nine members and two foreign guests—my American wife, Cathryn and her Malaysian friend, Chingyee. Both of them had met the group before and had become an important part of the group. At the first meeting on March 14, 2008, they talked in a friendly manner to the group members and helped me entertain the members with games and foods. Cathryn also attended the 13th meeting, wrote with the group on the topic “Study Abroad” and shared her sample with them (13V).

The group and the guests sat in a circle in our well-lit living room in my temporary home. When it was Anne’s turn to read her story (9W_W), she gave a photocopy to each attendee and said with a smile, “first I will read this article and after I read it, maybe you have some questions” (see Extract 6, place names omitted).

Extract 6: “My first time to hear her voice in phone” (9W_W)

What her voice will be like? I constantly asked myself since I entered senior high school. In elementary school and middle school, I felt extremely disgusted for her voice. I never listened to her voice carefully. I didn’t image her voice until that day.

When I was a freshman in senior high school, father suffered a serious illness. He needed a further check, so he had to go to [ ] with mother, and uncle would wait for them at [ ]. One day when I was doing my homework in dormitory alone, all of a sudden, the phone rang. I was dreaded by the ring. I answered the phone, it was uncle. He told me mother and father had arrived at [ ], I was confused. So I asked why they went there. He told me the situation of father, I felt sad and anxious. He said if I wanted to talk to mother. I should have refused, for I knew she never picked up a phone, and I knew clearly she couldn’t hear my voice at all. But too late, she got the phone.
“Anne, are you OK at school?” mother said, her voice was so loud that could penetrate my ear, but very clear.

“Mom, I’m very well. Why do you go to [ ]?” I said anxiously.

“At school, you must take care of yourself, no hungry, no cold, and eat well. Do you understand that?” mother continuously said.

“Mom, mom……” I broke our crying.

“Er…and your father only need a physical check, not too serious. We will be back tomorrow, if possible, we can meet each other at the county. Do you have anything else to say?” mother said, never gave me a chance to say, her loud voice was still going on.

I kept silent, because I knew no matter how hard I shouted, she never could hear my voice.

“Since you have nothing to say, I will give the phone to your uncle” mother finally said.

I said few words with uncle, totally controlling my feelings. Then I hung up the phone, could not help my tear at tall.

In fact, her voice was so loud and clear, maybe beautiful. It was a pity that I didn’t find it before. But now I learn to taste her voice and cherish it. It will be always around me no matter where I go.

When Anne had finished reading her story, applause followed. Anne smiled again. Questions, answers and comments continued for about seven minutes. Anne explained to my wife, Cathryn about her mother’s hearing problem: “her ear not … so good.” She told Matt that she had since learned to “cherish” her mother’s voice “because I think my mother is important–to me. And, and this time is my, is my, for the first time for my mother to pick the phone and also it was my first time to talk [to] her in [on the] phone.” Cathryn appreciated Anne’s expression in her story, “I really like that “I LEARN TO TASTE HER VOICE AND CHERISH IT…a very interesting image…a very unique way of expressing that idea.” Lastly, Anne explained to Peter she “felt angry and anxious” because “they never…let me know everything bad happened in my family.” These interactions at the meeting positioned Anne as the owner of yet another story (after “An unforgettable event in childhood,” 4W_W), one which she shared via writing in English at the 14th meeting with fellow learners and some
‘outsider witnesses’ (White, 2007). She was in effect thickening the storyline (White, 2007) of accepting her mother. This storyline first emerged in “An unforgettable event in childhood” (4W_W): “I…neve listened carefully to my mother, sometimes even felt her ugly and disgusting…I didn’t change my opinion until that day”). In “My first time to hear her voice in phone” (9W_W), her mother’s voice became the focus (see Extract 7):

Extract 7: “My first time to hear her voice in phone” (9W_W)

In elementary school and middle school, I felt extremely disgusted for her voice.
I never listened to her voice carefully. I didn’t image [really hear] her voice until that day. (¶1)

In “An unforgettable event in childhood” (4W_W), Anne’s position as a daughter was unclear. It was in answering another member Matt’s question that her position became clear.

M …So you felt ((looking up and smiling at Anne)) very angry. Yes. At that time. ((Spoken in a falling tone.))
A Why I felt angry?
M Yeah, you have said, you, um
A You mean why I felt angry at that time? ((speaking in a falling tone))
M En, yes.
A Because I also want to have a lovely umbrella like my classmates.
M And now, what’s your feeling about that? (I: en)
A Um, I feel grateful for [to] my mother ((laughing)). ((I nod my head.)) (11V)

Later, Anne showed that she had accepted her mother in her revised conclusion to “My first time to hear her voice in phone” (9W_W, see Extract 8).

Extract 8: “My first time to hear her voice in phone” (9W_W)

In fact, her voice was so loud and clear, maybe beautiful. It was a pity that I didn’t find it before. But now I learn to taste her voice and cherish it. It will always be around be no matter where I go.

Here, Anne was storying a new self, an accepting daughter, more clearly than she had done in “An unforgettable event in childhood” (4W_W).

So far, Anne’s autobiographical writing concerned a wound, one primarily associated with her relationship with her mother as mediated by her experiences of poverty, and its
process of healing at a family and relational level. “A never healed wound” (11W, August 29, 2008), to be presented next, goes further, revealing how Anne’s previous poverty experiences became a lens through which she approached and understood new experiences.

“A never healed wound” (11W_W)

The story recounted Anne’s recent experience of encountering a city girl from another province, who had taken a university-organized learning trip to the countryside in Anne’s hometown during the summer holiday (July-August, 2008). The city girl was “pretty,” “very kind and excellent.” Anne and she “soon became good friends and slept in one bed.” Yet a gap began to surface when the city girl revealed her childhood experiences (see Extract 9):

Extract 9: “A never healed wound” (11W_W)

When she was a little child, her family traveled [to] so many places. She also told me how her parents taught her elder sister and her.

…an unspoken feeling occurred to me. I only felt how large the distance was, even though we were so near. …I saw the distance between us, and it was a never healed wound. But I also very clearly know it was not only the distance between us, but also the countryside and city’s.

Anne’s story created a paradoxical image: two in one bed with an immense gap in between. What divided them were not only the different childhood stories, but also the different learning opportunities they each could access, simply because of where they had lived in China.

Three aspects about this story are intriguing to me. Anne first mentioned the city-girl in her introduction to a Chinese report “一道不可愈合的伤口” (A Wound Never to Be Healed) (30W, August 24), which had won her—as one of the three from her class of 48 students—an “Excellent Essay” award from her university. It was the first time Anne’s Chinese composition had been “praised” (5 I, September 3, 2008).

Secondly, several reasons may explain why Anne re-wrote the story in English for the writing group. Maybe her encounter with the city girl had had a strong impact on her, or she was encouraged by the award and expected a similar affirmation from the writing group, or her previous experiences of sharing her English writing in the group had been positive, or
English was the written language of the group. Presumably, it would have been less likely for Anne to have done so for her “irresponsible” mentor, who did not seem to read her writing.

Thirdly, Anne concluded her texts differently in Chinese and in English. In her Chinese report, after giving various reasons, Anne concluded sadly: “农村和城市间的这大距离是道用不可愈合的伤口” (The big gap between countryside and city is a wound never to be healed). In contrast, she wrote in English (see Extract 10):

Extract 10: “A never healed wound” (11W_W)

On the way home, a dumb idea accidentally appeared to my mind: I wanted to change the situation of whole countryside. But now I’m not sure if it can come true. Maybe it is a daydream. I’m still looking forward to it any way.

This comparison pointed to new possibilities associated with writing in English as an additional language (e.g., Kramsch & Lam, 1999; McKay, 1993; Pavlenko, 2001; Steinman, 2005; You, 2008). In Anne’s case, writing in English seems to allow her to pursue hope and see herself as an agent heralding preferred futures.

In “A never healed wound” (11W_W), Anne moved into a sociocultural context to understand the larger world of which she was a part. Anne’s engagement with this context not only showed the growing gap between countryside and city in China, but also imagined alternative possibilities, in which she would be an active agent. But even this interpretation of Anne’s did not exhaust the possible meanings of her experience. In time, Anne would revisit her previous poverty experience and find alternative meanings, as is the case with “A Wound in my Soul” (14W_W).

“A Wound in my Soul” (14W_W)

During the second stage, our writing group was to focus on argumentative writing, but several times the members were encouraged to turn their experiences into 50-word poems as a way of learning to be concise. For the last meeting during the 32nd week, to share with the group her best writing, Anne wrote an autobiographical poem about the same plastic sheet event recounted in “An unforgettable event in childhood” (4W_W). Eleven members and three foreign guests—Cathryn, Chingyee, and my Australian supervisor / advisor Audrey—joined in the meeting. When it was Anne’s turn, she produced a small piece of lined paper, and, after a brief introduction, read her poem “A Wound in My Soul” (14W_W, see Extract 11).
Extract 11: “A Wound in my Soul” (14W_W)

A Wound in my Soul

I was waiting, waiting,
she appeared wearing a plastic
but no lovely umbrella, rainboot
merely a small clean warm plastic.
I ran into rain, cried,
so did the God.
A small figure gradually disappeared in the cold rain.
How could I forget it easily?
The Wound in my soul.

After finishing reading her poem, Anne translated it into Chinese. Applause followed. I was pleasantly surprised when Anne said, “This poem is also a gift to Nicholas” (32V). English writing was now serving an additional social function for Anne.

Anne’s poem marked yet another journey she had travelled in making sense of her poverty experiences. The plastic sheet, described only as “clear” before, was now “small clean warm.” Her mother, who used to be represented as “ugly,” now, without being described as such, became anonymously represented in “she gradually disappeared in the cold rain.” “I cried” revealed a shift of her feeling from mainly shame, as in “An unforgettable event in childhood” (4W_W), to “I was moved also I feel ashamed,” as she explained to Peter (32V). The multiple ‘potential meanings’ (Holquist, 2002; Robinson, 1999, pp. 200-201) of the plastic event were now grafted into alternative interpretations. Moreover, reading the line “A small figure gradually disappeared in the cold rain” (14W_W), I had a sense of Anne feeling loss for her mother. In the end, the event was the same, but the meaning lived by Anne now was not. It went deeper than the emotional hurt to the level of “a wound in my soul.”

Anne’s second tree
At the last writing group meeting during which she shared her poem, Anne drew another tree (see Figure 3) to represent herself as a writer after joining the writing group. It was “tall and strong,” with “so many fruits” (32V).

Anne’s “tree” images and captions showed her self-perceived changes in her subjectivity (e.g., from “knowing little” to “knowing more about writing” such as “the importance of a title”3) and her relationship with English writing (e.g., from “I write just for examinations, I hate it” to “I can express my feelings freely in my biography”). Anne’s special moments in the writing group, listed above her first tree, remind us of the sociocultural nature of her learning process, through which her changes took place.

Conclusions
Two major shifts took place in Anne’s poverty stories analysed above. First is the shift of context. At first, as in “An unforgettable event in childhood” (4W_W), Anne was writing mainly personal and relational stories about her and her mother. “A never healed wound” (11W_W) touched upon a broader sociocultural context of the underprivileged countryside Anne grew up in China. “A Wound in my Soul” (14W_W), written for the last writing group meeting of the year, went further, revealing, in an expressive way, a deeply felt hurt from her
poverty experience. “I cried, so did the God.” Second is the thickening of an alternative storyline, which plotted Anne’s increasing acceptance of her mother and her childhood self. When it first appeared in “An unforgettable event in childhood” (4W_W), this alternative storyline was indistinct. That’s why group members such as Matt asked Anne how she felt about her mother now. In “My first time to hear her voice in phone” (9W_W) read at the 16th group meeting, in contrast, Anne’s position towards her mother was clear. “But now I learn to taste her voice and cherish it.” Her mother was no longer represented as “ugly” in “A Wound in my Soul” (14W_W), but as a caring and suffering figure. Particularly worth noting is the shifted meanings of the plastic sheet, which now stood for motherly love and a mother of aspirations. As Anne explained about the “small clean warm plastic” (see Extract 12),

"Small", because I was just a child. My mother certainly took a small piece of plastic to me."Clean", because my mother loved me, and she was a very careful woman (家里虽然穷但她很有志气的). "Warm", because the plastic was hidden under my mother’s coat. (Email on May 8, 2010)

Extract 12: Anne’s email on May 8, 2010

It is impossible to tease out exactly to what extent autobiographical writing, English, and the writing group each contributed to the two shifts. After all, these features were intertwined in the kind of writing Anne engaged in and should be understood as such. Nonetheless, evidence suggests that they all made distinctive contributions to the shifts Anne made. Firstly, autobiographical writing repositioned Anne as a knower, especially of her own poverty experiences, which were “carved in my mind and soul” and “unforgettable” (8W). Had Anne started with argumentative writing, the usual type of writing in exams, she would have struggled for “powerful evidence,” which, according to her, included “famous people’s story … their biography, and their saying, and maybe some true things happened in reality, in daily life.”

Secondly, writing in English positioned Anne as a confident and capable writer. Anne’s past experience of writing in Chinese was one of failure and fear. “All the time I just (could not finish?) my composition at tests. So I developed a fear for writing.” Furthermore, not having a rich vocabulary in Chinese, a must-have according to Anne, also restricted her as a Chinese writer (2 II). Even with her award-winning Chinese report, Anne said, “I think it’s
not good. Because I can’t use some beautiful words to express my feelings” (5 I, September 3, 2008). That language choice mattered to Anne as a writer was most clear in her statement, “You can write a lot in English using a few simple words, because–which is impossible in Chinese."

Thirdly, the writing group repositioned Anne as a writer with a real responsive audience. As shown earlier, as a title, “My first time to hear her voice in phone” (9W_W) was born in the context of her reading others’ texts, followed by the members’ attempts to write differently from each other. When Anne completed her story, she intended to “show it off” at a group meeting. It was in the writing group that Anne was “praised” for her writing in English for the first time in her life. Even with her controversial description of her mother as “ugly,” Anne received sympathy. As one member Abby commented, “I have the same experience and same feeling [about my parents] but now when I recall ( ) I feel regret” (11V).

Autobiographical writing, English and a supportive writing group by no means exhaust possible contributors to Anne’s shifts mentioned above. Nevertheless, together, these three dimensions of genre, language, social context form a stark contrast with Anne’s previous test-related experiences of writing in Chinese and English and her ongoing experiences of doing English writing for her mentor.

Let us situate Anne’s stories in a re-stor(y)ing / transformative context (Grant, 2001), asking ourselves, as ESL / EFL literacy educators, what possibilities can be pursued to support learners like Anne. I offer two rough ideas drawing on sociocultural theories of learning. First, writing autobiographically in English in an appreciative writing group seems to reposition struggling writers as knowers and capable learners. Suppose their demonstrated level of competence is naming an event significant to them, their potential level may be developing the event into a story or several stories, with appreciation and assistance from others. As such, a rich learning space may be opened for learners to learn multiple meanings of their experiences, conventions of writing in English, and see themselves through the eyes of others, such as their immediate audience of writing peers.

Second, writing autobiographically in English in an encouraging writing group seems to provide a space for learning writers to develop a ‘multi-voiced sense of self’ (White, 2005, p. 13), i.e., the understanding of oneself as an agent in relation to one’s multiple experiences. In Anne’s case, she re-authored her identity from a writer “just for examinations” to a writer of multiple stories, with multiple meanings of poverty, her mother and herself. A space like
this appears promising to redeem literacy education away from irrelevance and to introduce alternative possibilities.

What Anne has written about in English is far more than “merely a small clean warm plastic.” Her stories show that identity is as relevant to EFL autobiographical writers as it is to L2 published memoirists in the ESL context, at least for Anne from this particular writing group.

Limitations and future research directions
The present study has three major limitations. First is interviewing Anne in Chinese and English, Anne’s two additional languages, instead of Bai, her home language. If I could speak Bai and had interviewed Anne also in Bai, I might have obtained more revealing information about her identity work. Second is the limited sample size. However, this has been partly compensated by the in-depth description of this study and three divergent case studies presented in my PhD thesis. Third, focusing on EFL autobiographical writing and identity, the study has not examined whether or not Anne improved her English writing as a result of her active engagement in EFL autobiographical writing.

To extend the present study, future researchers may consider using fluency as an additional construct in evaluating the pedagogical usefulness of EFL autobiographical writing. Further SLA writing research may also consider three practice-oriented categories that focus on the use of autobiographical writing by L2 learners in different parts of the world. One is situated in various kinds of writing groups. The other is situated in regular classrooms, which integrate certain writing group elements. The third is identity work as mediated by autobiographical writing in non-English languages such as Chinese. Such research does not resolve major challenges such as oversized classes and marginal status of writing particularly in the EFL context (Leki, 2001; You, 2004). Nonetheless, applied research of this kind seems particularly beneficial to SLA as ‘a theory of practice’ (Hall, 1997) and to L2 learners, whose identities cannot be told by any single story. Anne has used EFL autobiographical writing in powerful ways to express, continue and re-configure her multiple identities. May such also be the case for other L2 learners as more SLA scholars adopt alternative approaches to L2 literacy research and education.

Acknowledgements
My thanks go to Endeavour Asia Award (2007) from the Australian government and David Myers Scholarship of La Trobe University, which respectively financed my theoretical explorations of the proposed study and my fieldwork in China. My sincere gratitude also goes to Dr. Howard Nicholas, Dr. Audrey Grant, Dr. Peta Heywood, my wife Dr. Cathryn Yang, and the anonymous reviewers who provided valuable comments on the earlier drafts of this paper. All remaining errors and inconsistencies are mine.

References


Park, G. (2009). "I listened to Korean society. I always heard that women should be this way ...": The negotiation and construction of gendered identities in claiming a dominant language and race in the United States. *Journal of Language, Identity, and Education, 8*(2/3).


Here, EFL refers to contexts where English is not used in people’s daily lives (Leki, 2001, p. 23) and is mainly associated with language learning in classrooms (Block, 2007).

Spelling, grammar, case and punctuation as in the original. To retain the flavour of the original, the same applies to other excerpts from Anne’s writing samples.

All student names in this article are pseudonyms. Other first names are used with permission.

In my PhD dissertation completed in 2011, I also told stories of three other case study participants, including another first-year student and two third-year students.

Italicized texts in this article indicate my translation from Chinese.

TEM-4 stands for Test of English for English-Majors, Band-4, taken by English-major students in their second year. Each student can only repeat the test once.

This indicates Anne’s first writing sample for her “导师” (Mentor).

This is an allusion to a Chinese government policy, with one hand managing “物质文明” (material civilisation) and the other dealing with “精神文明” (spiritual civilisation).

A. Grant (personal communication, 2010) suggested Anne might have meant “prized” here, meaning “valued, treasured by.”

The story was not written by any writing group member. It was a sample included in A Peking University Coursebook on English Exposition Writing (Coffman, 2003).

I use 14V to refer to the transcripts from the 14th meeting video recording.

The two words ‘umbrellars’ and ‘shoots’ are kept as original. They should be ‘umbrellas’ and ‘boots’ instead.

The idea occurred to me when I attended a presentation by Hassall (2007) on Extremely Short Stories Competition (ESSC).

It is interesting that what Anne described here was mainly what she learned during the autobiographical writing stage, even though several months had passed.

Asked about the God in the line “I cried, so did the God,” Anne said the rain reminded her that “好多就形容下雨的时候，就说：‘老天哭了!’” (“Many describe raining as ‘The Old Heaven /God is crying!’” (Interview on December 20, 2008). Therefore, Anne mentioned God more for rhetoric purpose than religious faith.

“My family was poor, but she had aspirations.”

“用英文写拿几个简单的单词就可以写好多好多了，因为用汉语写写不出什么来。”Here, Cathryn reminded me that Hemingway wrote mainly with one-syllable words.

For instance, Anne’s engagement with a different other also played a significant role in shifting her contexts of writing. In “An unforgettable event in childhood,” her mother was the other, familiar and near. In “A never healed wound,” the city girl was the other, unfamiliar and from a distance. In “A Wound in my Soul,” God joined in with her mother as an other, giving additional meaning to the rain.”
Networked Collaborative Learning: Social Interaction and Active Learning (1st ed.)

Reviewed by Weihong Wang
China University of Geosciences
Wuhan, China

Networked Collaborative Learning (NCL) is a critical topic for 21st century education worldwide. Guglielmo Trentin’s *Networked collaborative learning: Social interaction and active learning* has addressed this topic from a social-cognitive perspective and successfully delineated a four-dimensional framework for the implementation of NCL in future education. The insightful perceptions of NCL and the detailed instructions on NCL practice presented in this book are of great significance for educators and practitioners, especially teachers and would-be teachers who attempt to know how to employ NCL in their teaching so as to meet the educational needs in this new era.

The six-chapter book starts with a favorable presentation of numerous potentials of technology-enhanced learning (TEL) through comparison with conventional classroom-based learning. Against this background, NCL is put forward as the cornerstone of future TEL sustainability in this first chapter. It then goes further to explicate how NCL can be best developed from four dimensions, namely pedagogical approaches, e-teacher professionalism, instructional design models, and student performance assessment. The four-dimensional model in developing NCL offered in this chapter serves as a roadmap for understanding the whole book, as the four dimensions are developed into the next four chapters respectively. Chapter 2 explores the philosophy underpinning NCL and the necessary paradigm shifts involved in moving from traditional classroom learning to collaborative networked learning. The presentation of defining pedagogical features of NCL in comparison with traditional learning propels teacher-readers to continually reflect on traditional pedagogies while reading this chapter. This kind of reading experience may gradually lead them to the necessary shifts...
required in moving beyond the passive, content-based pedagogies to embrace the more interactive and constructive ones proposed in NCL.

Chapter 3 investigates the professional development involved in NCL, mainly the abilities required for teachers to implement this e-pedagogy. It first provides the rationale for teacher professionalism development, and then goes on to elaborate the multiple roles teachers in NCL environment have to take, like being a subject expert, instructional designer, classroom teacher, e-teacher, or e-moderator. In this chapter, the author particularly emphasizes that in NCL teacher training goes far beyond training teachers to implement traditional teaching practices with new technical skills. Instead, NCL takes as its aim in teacher education and professional development to enlighten teachers to “be capable of making autonomous and informed decision about what e-teaching strategies will prove most effective for meeting the needs at hand” (p. 54).

Chapter 4 focuses on instructional designing in NCL environment. It offers thorough and detailed instruction on how to design coherent and effective NCL courses. The instruction covers macro-level considerations of aims, objectives, course prerequisites, course flexibility, educational strategies, evaluation criteria and micro-level designing of a variety of e-tivities and e-content.

Chapter 5 is mainly about the evaluation and assessment of NCL. It presents readers a repertoire of evaluating tools and methods so that diverse users can choose from them the most suitable ones according to their own needs. In other words, this chapter turns the intricate details of students’ performance in collaborative work into measurable variables, which enables teachers to make comparatively objective evaluation on individual student’s development.

Chapter 6, the last chapter, re-addresses the importance of teachers’ efforts in the entire four dimensions to guarantee the distinction of NCL and the effectiveness of e-learning.

The topic discussed in this book is a prompt response to the recent advances in e-learning technology and practice. Drawing on the author’s own experience in classroom application of NCL and comparing NCL with conventional learning, this book offers insightful perceptions and useful guidelines for networked learning. However, as it is an advantage of this book to foreground e-pedagogy by strategically comparing it with traditional ones, readers might take a black-and-white view towards the two types of pedagogies and interpret them as absolutely incompatible. The author’s intention here is to
remind teachers of the danger of directly transferring classroom practice into the cyber world but not advise them to abandon everything related to traditional teaching practices.

Generally speaking, this book deserves the attention of educators at all levels. It will definitely benefit teaching practitioners in deepening their understanding of NCL in the midst of transition from classroom teaching to e-pedagogy. It can also guide educational institutions in their decision-making as regards investing e-learning in their own educational contexts.
Uncovering EAP: How to Teach Academic Writing and Reading

Reviewed by Sally Ashton-Hay
Southern Cross University
Tweed Heads, Australia

Uncovering EAP is a handy resource for teachers or lecturers who teach academic skills to students preparing for tertiary study, IELTS examinations, or those already enrolled at university. The focus is on English for Academic Purposes (EAP) reading and writing with useful tips for teaching these crucial study skills. New or experienced teachers will find this Macmillan Books for Teachers’ text full of practical applications for the classroom.

Ten chapters cover vital academic reading and writing skills: (1) writing and reading approaches; (2) understanding task questions and selecting ideas; (3) types of academic assignment; (4) academic writing genres; (5) critical thinking; (6) academic vocabulary; (7) giving feedback and redrafting; (8) reading academic texts; (9) research and referencing; and (10) moving from general EAP to subject-specific teaching.

Each chapter presents a brief theoretical overview, offers good practice case study exemplars, and a bank of learning-centred activities and photocopiable worksheets. The value of the book is enhanced by practical teaching tips geared to developing specific skills in academic writing and reading. Intercultural case studies illustrate common challenges faced by non-English speaking students at university with proven strategies for achieving success. The chapter materials are detailed yet easily adaptable for other contexts and disciplines. With over forty activities in the book, any teacher’s tool kit can be readily extended.

Chapter five is a particular highlight because of the strategies for coaching critical thinking, often considered an unfamiliar or alien concept to many international students. Structured, bottom-up approaches scaffold a developmental continuum and overcome the slippery dip of inexperience. Staged activities encourage recognition of fact versus opinion, analysis of text through verbal reasoning, and production of an opinion prompted by visual
aids. Model questions investigate what students do not know or what they would like to know and extend judgemental and evaluative capacity while temporarily shifting the focus away from grammar. Chapter five concludes with two good practice cases and four purposeful activities linked to building the evaluative and analytical skills discussed. The activities clearly list aims, materials required, appropriate level, time required, and methodological steps to follow as well as an argument mapping worksheet to accompany an advanced pair or small group activity. The case study examples portray international students with little experience in analysis or succinct writing skills and advise specific approaches to improve understanding and achieve more successful outcomes. These case study dilemmas are beneficial and thought-provoking reminders for any teachers facing the challenge of teaching this skill to the uninitiated.

The McCarter and Jakes volume resonates with the voice of experience using tried and tested approaches to develop fundamental academic skills. Theoretical-based skill development, active learning, and good practice principles make this book a practical teaching companion. The only disadvantage is that there is little mention of technology or additional online website resources for extending academic skill development. Despite the lack of technological applications, the book is still a worthwhile contribution to the field of academic skills and recommended for any teachers bridging the double gap of academic English language skills and the challenges of higher education.
In *Language and Learning in the Digital Age*, Gee and Hayes draw attention to the evolving nature of communication, from spoken language to the written word to the emerging role of digital technologies in the dissemination of ideas. They argue that digital media, as a delivery system for communication, can be seen as a means to “power up” (p. 1) oral and written discourse. In an age where learners are largely treated as consumers of information, the authors argue that digital media present all individuals with opportunities to have a say in the production of knowledge. This makes the material broad in its appeal, addressing educators in all fields, particularly those whose focus is on language education or the social sciences.

The book consists of fourteen chapters, which are presented through explanations and examples. The first chapter familiarizes readers with the notions of “hope and fear” (p. 4) concerning the use of digital media. The second and third chapters introduce the definitions of language (e.g. oral communication and literacy) and elaborate on the strengths and weaknesses of spoken and written discourse. Chapter 4 explores the social aspects of language and interaction, focusing on bonding and distancing languages in particular. In the fifth chapter, the authors elaborate on the role of digital media in framing social relations.

The authors then underline the ways that digital media can be used as a means to substitute for experts and institutions, which they refer to as being “in crisis,” in chapter 6. The seventh chapter introduces school-owned literacy and includes school as part of the crisis in question, arguing that the knowledge offered in the context of conventional schooling fails to develop essential skills in problem-solving and advancement. In chapter 8, the authors call for revolutionary changes in education. To this end, the need for out-of-school learning systems is highlighted, and passionate affinity spaces are offered as a viable option. The ninth
chapter provides solid evidence of such spaces, expanding on the application of video games in teaching and learning.

Next, in chapter 10, appropriately titled “Cat, Passion and Expertise,” the authors illustrate their hands-on experience involving a cat, noting how discussion forums, with contributions from various individuals, became instrumental in their problem solving. Chapter 11 encourages readers to question the top-down imposition of knowledge and consider how amateurs could play a decisive role in knowledge construction, while in chapter 12 deals with multimodality and illustrates how words could mean more when combined with images and experiences. The thirteenth chapter is devoted mainly to social formations, citing oral, literal and digital formations respectively. In the final chapter, “Multitasking, Diversity and Commonality,” the authors provide a discussion of the possible advantages and disadvantages of digital media.

The vivid nature of the examples, the clarity of the definitions, and the authors’ determination to leave nothing up in the air result in a work geared towards practice, rather than theory. By viewing each topic from a range of disciplines, including history, linguistics and politics, they make the material attractive and enriching to a wide audience. Moreover, the authors encourage their readers to think about both sides of each issue, presenting the opportunities as well as the drawbacks of digital media, though the former are much more focused.

Despite these positive attributes, the book is not without its shortcomings. The reader at times wonders about the chapter divisions, as the shift from one to another is not smooth. To mitigate this issue, it might have been useful to include information about how each chapter was developed in order to make the progression more visible to the audience. The extensive use of “I” language in chapter two also leaves it unclear who is doing the talking, and the integration of games while in chapter 9 may be of limited appeal to many readers. Yet overall, the book provides reader-friendly and helpful insights into the changing spirit of communication and deserves a place on the bookshelves of educators from every discipline.
The nature of any medical communicative event, bilingual or monolingual, is in itself complex. Intended as a broader review in the theories of cross-cultural communication for medical interpreters, Medical Interpreting and cross-cultural communication offers a clear review of different theories for readers from many different professions.

The book opens with a review of literature which examines the arguments between the roles of language interpreters. In the literature, there is an invisible interpreter and a visible one. The concept of the invisibility model has presumed that there is no interaction between interpreters and speakers and speakers themselves. Visible interpreters, who are the focus of this book, are seen as essential partners in the interpreted communicative event (ICE). The author asserts that interpreters are influenced by social and institution factors during ICE.

In the second chapter, the author asserts the need to provide adequate healthcare services for patients who have limited English competency and explore how the doctor-patient relationship is essential to produce effective communication.

In the third chapter, the author explores three different lenses to understand interpreters’ roles: (1) the lens of society and the institution based on social theory, (2) the lens of the interaction based on sociological theory, and (3) the lens of discourse based on linguistic anthropology.

From the fourth chapter to the seventh, the author utilizes a public hospital California Hope (CH) as a case study to explore the roles and practices for interpreters. The author describes the research design, subjects, and process first. In chapter five, the author collects both quantitative and qualitative data from 392 ICEs at CH by spending twenty-two months studying ten interpreters. After analyzing the data, in chapter six the findings of the linguistic
and communicative strategies used by interpreters are various. They reveal interpreters have different levels of visibility. In chapter seven, the author sorts out the role of interpreters by analyzing their true voices from semi-structured interviews. Based on the analysis, in chapter eight, the author comes out with four metaphors for interpreters: (1) interpreters as detectives, (2) interpreters as multi-purpose bridges, (3) interpreters as diamond connoisseurs, and (4) interpreters as miners. The implications of theory development in integrating all of the complexities of the communicative event and in including the act of interpreting with its entire context are provided. In the conclusion, the author poses concerns and a strong statement for the readers to pay more attention to events of cross-cultural communication and make efforts to change the invisibility of interpreters in ICEs to actively visible.

In general, the book is highly informative. It is thoroughly done in the process of research design with a case study approach, and the findings are important to enrich the current literature on interpreting in a medical setting. The following suggestion might help the readers to form a better understanding for the book. In the fifth chapter, the qualitative data analysis may need more elaborations to conclude the findings. It is needed to present the analysis with a more organized and easy-to-read way. Taking this question into account, readers will no doubt treasure this book as an important case study reference book that illustrates what roles interpreters play in ICEs at a public hospitals.