TESOL International Journal

Teaching English to Speakers of Other Languages

TESOL International Journal

Foreword from the editor

This issue of TESOL International Journal contains six papers. In the first paper, **Gui Bao** conducted a study to examining EFL learners' initial vocabulary learning through reading sentences and performing tasks related to target words. It was found that out of four tasks, two involving input (matching and choice) and two output (definition and combining), the definition task outperformed the others for EFL vocabulary knowledge acquisition.

In the second paper, **Pino Cutrone** studied 23 Japanese EFL learners' listenership behaviors by examining the effects of variables such as L2 proficiency, extraversion, and willingness to communicate. Data from video recordings of intercultural dyadic conversations in English between a Japanese EFL learner and a native English speaker with questionnaires and language proficiency tests revealed a great deal of individual variation regarding listenership behaviours and factors such as the specific contexts of each conversation, the personality of the participants, and other peripheral variables such as participants' moods at the time of the conversations seemed to play a part in their oral output.

Zhenjie Weng, Jingyi Zhu, and **Grace J. Y. Kim** systematically reviewed classroom-based empirical studies on language teacher agency within ESL/EFL/bilingual contexts. Chosen studies were scrutinized for theoretical frameworks, contexts, methodologies, major findings, methodological and ethical issues, as well as implications. They found that overall teacher agency still remains under-examined in ESL/EFL/bilingual language education contexts and suggested that data collection methods, such as playback sessions, surveys, mapping, collecting artifacts, and focus groups can be used to triangulate data analysis from different angles in future studies.

In the fourth paper, with data from stimulated recall activities, written observations, and interviews, **Maria Eleftheriou** examined tutors' and tutees' experiences of the effectiveness of directive and non-directive strategies of instruction in a writing center context at a Middle East university. It was found that while tutors and tutees find directive approaches useful for lower order concerns, both tutors and

tutees prefer non-directive approaches when addressing higher order concerns.

Mohammad Nabi Karimi, Ebrahim Zangani, and **Nahid Fallah** investigated the allocation of attention to meaning and form simultaneously in reading comprehension in the foreign language among monolingual and bilingual learners of English in Iran. The study did not find significant difference between bilingual and monolingual learners regarding the level of processing of the targeted lexical form as far as the subject of attention to meaning and form is concerned. However, the results indicated that deeper processing is associated with better reading comprehension.

Enisa Mede and **Şenel Yalçın** explored self-reported beliefs of experienced and novice instructors about using adaptation strategies in intermediate English classes of a preparatory program in Turkey. Data collected from reflective essays, lesson plans, and semi-structured interviews showed that both novice and experienced instructors shared highly positive beliefs about the implementation of adaptation strategies in their courses. It was also found that the adaptive decisions of the participating instructors were closely related to their students, tasks, context, time, and their own beliefs.

Such is the brief overview of the six papers in this issue. If you are actively engaged in research or have done research related to English language education, please do not hesitate to contact us about the possibility of publishing with TESOL International Journal. Apart from individual paper submissions, we also welcome proposals for special issues.

Xinghua (Kevin) Liu

School of Foreign Languages, Shanghai Jiao Tong University, China

Email: liuxinghua@sjtu.edu.cn

Contents

Comparing Input and Output Tasks in EFL Learners' Vocabulary Acquisition

Gui Bao[1]
Nanjing Tech University, China

Abstract

Task-based vocabulary learning has recently attracted lots of attention in the field of second language (L2) vocabulary acquisition. This article compares the effects of input and output tasks on English as a Foreign Language (EFL) learners' acquisition of vocabulary knowledge. Four intact classes of EFL learners were randomly assigned to one of four tasks of learning 18 target words through sentence reading exercises, i.e., matching, definition, choice, and combining. The definition task was found to be more effective than the other tasks in EFL vocabulary knowledge, irrespective of EFL proficiency. At each level of EFL proficiency, the matching and combining tasks performed equally well. The choice task had a small advantage over the matching and combining tasks at the learners' low EFL proficiency level, but not at the intermediate or high EFL proficiency level. Differential processing and division of attention may well affect task effectiveness.

Keywords: input task, output task, EFL vocabulary knowledge, differential processing, division of attention

Introduction

It is generally acknowledged that second language (L2) reading with related word-focused tasks (i.e., reading plus) is more conducive to vocabulary acquisition than L2 reading without such tasks (i.e., reading only) (Laufer & Girsai, 2008; Paribakht & Wesche, 1997; Wesche & Paribakt, 2000). Word-focused tasks through reading arouse the learners' attention to new words, thereby increasing the chances that the words will be retained. Such attention may not be necessarily evoked during a "reading only" task, whose purpose is to gain an overall understanding of the text (Hill & Laufer, 2003, p. 90). L2 researchers and teachers have developed a variety of tasks to ascertain whether some tasks are more effective than others in improving vocabulary knowledge (Barcroft, 2002; Golonka et al., 2015; Huang & Lin, 2014; Joe, 1998). One important question arising from this line of research is how to theorize task effectiveness. Of most recent theoretical interest is the Involvement Load Hypothesis (ILH) (Laufer & Hulstijn, 2001).

The ILH accrues from Laufer and Hulstijn's (2001) concern for the failure of the levels of processing (LOP) framework to provide clear definitions of notions like "depth of processing" and "degree of elaboration" (Craik & Lockhart, 1972; Craik & Tulving, 1975), and for the possible effects of human and social-cultural factors like motivation on information processing. According to the ILH, the construct of involvement is defined as consisting of three concrete task-related components, i.e., need (motivational dimension), search and evaluation (cognitive or information processing dimensions). The involvement load is thus determined by the presence and strength of each component. The ILH predicts that, the greater the task-induced involvement loads, the more likely the word will be learned. A number of researchers (e.g., Folse, 2006; Hulstijn & Laufer, 2001; Keating, 2008) designed various L2 vocabulary learning tasks to test the ILH. Nevertheless, the hypothesis was supported only partially, suggesting that task effectiveness is not always revealed by the involvement load a task generates.

One of the crucial issues is the dependence of task effectiveness on the input- or output-orientation. The ILH hypothesizes that word retention is contingent upon nothing but

E-mail: boggy2008@126.com. School of Foreign Languages and Literature, Nanjing Tech University, No. 30, South Puzhu Road, Pukou District, Nanjing 211816, China.

the involvement load a task generates, be it input- or output-oriented. This suggests that input and output tasks would be equally effective if they generated identical involvement loads. Although numerous studies have addressed the ILH, no due attention has been paid to this suggestion. Accordingly, the current study seeks to examine the effects of several input and output tasks with the same involvement loads on English as a Foreign Language (EFL) vocabulary learning.

Background Literature

The ILH focuses essentially on semantic spread (i.e., the degree to which semantic information is enriched in processing new words), ascribing the learner's retention of hitherto unfamiliar L2 words (or vocabulary items) to the synergism of the three components of task-induced involvement, i.e., need (N), search (S), and evaluation (E). According to Laufer and Hulstijn (2001), "need" is a drive to meet the task demands; it may be absent (-N) when the task is not relevant to the new words, or may have a moderate presence (+N) if it is imposed by an external agent, for instance, if L2 reading comprehension questions are relevant to the new words glossed in the text, or even a strong presence (++N) if it is intrinsically motivated by the learner per se, for instance, in a composition where the learner decides to consult a bilingual dictionary for the unknown equivalents of certain L1 concepts. "Search" is an attempt to find the form or meaning of an unknown word; it may be either absent (-S) if this attempt is not made, or present (+S) if it is. "Evaluation" involves a decision about the meaning or usage of a new word. It may be absent (-E) when the task is not relevant to the new words, or may have a moderate presence (+E) when the task entails recognition of differences among words (as in a fill-in task), or among several senses of a word in the specific context, or even a strong presence (++E) when the task entails the use of a new word in an original text. The involvement load is thus operationalized as the involvement index, with the absence of a component marked as 0, moderate presence as 1, and strong presence as 2. The ILH assumes that each component of involvement carries equal weight in relation to word retention, and that the involvement indexes can be simply added to represent the degree of overall involvement.

Most empirical research regarding the ILH has been conducted through learning words in reading passages or single sentences. In both sizes of context, however, mixed evidence is found for the ILH.

The first mixed evidence was derived from Hulstijn and Laufer (2001), who compared three tasks through passage reading: reading [+N, -S, -E], reading plus fill-in [+N, -S, +E], and composition writing [+N, -S, ++E]. The writing task was found to produce significantly better retention of the target word meanings than both the reading and reading plus fill-in tasks in two experiments, whereas the reading plus fill-in task resulted in better retention than the reading task in one experiment but not in the other, thus largely but not fully supporting the ILH. The majority of follow-up studies also lent partial support to the ILH (in passage contexts: Eckerth & Tavakoli, 2012; Keating, 2008; Kim, 2008; Laufer & Rozovski-Roitblat, 2011; in sentence contexts: Bao, 2015; Folse, 2006; Webb, 2005; Webb & Kagimoto, 2009), with those in full support of the ILH being in the minority (e.g., in passage contexts: Laufer & Girsai, 2008; Min, 2008; in sentence contexts and on overall word recall: Pichette, de Serres, & Lafontaine, 2012). Among the factors affecting word retention are time and type of vocabulary knowledge measurement (e.g., Keating, 2008; Kim, 2008; Webb, 2005), word encounter frequency or retrieval frequency (e.g., Eckerth & Tavakoli, 2012; Folse, 2006; Laufer & Rozovski-Roitblat, 2011), task design (e.g., Eckerth & Tavakoli, 2012; Keating, 2008; Kim, 2008), and word characteristics (e.g., Pichette, de Serres, & Lafontaine, 2012).

To summarize, previous empirical studies regarding the ILH either involved comparing input and output tasks with different involvement loads (e.g., Hulstijn & Laufer, 2001; Webb, 2005), or output tasks inducing identical or different loads (e.g., Bao, 2015; Folse, 2006; Kim, 2008). An interesting question is, would the input- and output-oriented tasks be equally effective for L2 vocabulary learning if the task-induced involvement loads were identical? Although the ILH predicts that higher task-induced involvement will lead to better retention of an unknown word, regardless of whether the task is input- or output-oriented (Laufer & Hulstijn, 2001, p. 20), it is possible that input and output tasks, which entail two qualitatively different modes of

processing, may contribute to L2 vocabulary learning in ways unpredicted by the ILH. Therefore, the problem of concern is how the input and output tasks compare in L2 vocabulary learning if the task-induced involvement loads are held constant.

Another question of pedagogical interest is whether task effectiveness regarding vocabulary learning is related to L2 learner factors like L2 proficiency. If the positive effect of a task as compared to another task persists across L2 proficiency levels, this task would apply to the normal classroom setting where a regular class consists of L2 learners with mixed L2 proficiency levels. In one of the two experiments on English as a Second Language (ESL) learners with two levels of ESL proficiency, Kim (2008) compared two tasks, which were hypothesized to have the same level of task-induced involvement (composition writing (+N, -S, ++E), sentence writing (+N, -S, ++E)). In initial word learning, there was neither a significant main effect for either task type or ESL proficiency nor a significant two-way interaction. However, it remains to be found whether the independence of task effectiveness from language proficiency is generalizable to tasks other than composition writing and sentence writing.

Thus, this study addressed the following two questions:
1. How does task type (i.e., matching, choice, definition, combining) affect EFL learners' vocabulary knowledge?
2. Do the task type effects on EFL learners' vocabulary knowledge vary with EFL learners' proficiency?

Method

Research Design

This study was an examination of how word-focused tasks would affect EFL learners' initial word learning, employing a pretest-posttest experimental design. Task type, a between-subjects factor, had four levels, consisting of two input tasks (i.e., matching and choice) and two output tasks (i.e., definition and combining). These tasks were assumed to have identical involvement loads, but differed in the input-output orientation, presence of a target word, or both (see the Input and output tasks section). Initial word learning was measured by an immediate posttest of EFL learners' passive recall and use of the newly learned target words. Unlike the posttest, however, the pretest was a test of EFL proficiency. The whole experiment was conducted in class. During the experiment, all classes were asked to do the same sentence reading exercises involving the target words, but to complete a different word-focused task assigned to them within the allocated time limits. A short time after task completion, all the classes were asked to take the test of the target words.

Allocating different time limits to different tasks was due to the relationship between a task and task time. That is, all else being equal, tasks of different types tend to require different lengths of time to complete. Like a number of previous studies (Folse, 2006; Hulstijn & Laufer, 2001; Keating, 2008), this study considered task time to be an internal part of a task.

Participants

The participants were 167 first year intermediate EFL learners from different non-English specialties at one Chinese university. They are mostly female (151 females vs. 16 males), and their ages ranged from 16 to 22 years old (M = 18.51, SD = 0.86). These participants had learned English at school for at least six years before they went to university. They had to continue to learn English as a compulsory course at university for the first two years. They were taught English in one of the four randomly assigned intact classes for 4 hr per week.

Each intact class was randomly assigned to one of the four vocabulary learning tasks. One month before the experiment, all four classes of EFL learners took an EFL proficiency test with a full score of 100. The descriptive statistics for their EFL proficiency are displayed in Table 1, including sample size (n), mean (M), standard deviation (SD) and Shapiro-Wilk Test.

Table 1
Descriptive statistics for EFL proficiency across task type

	n	M	SD	Shapiro-Wilk Test	
				W	p
Matching	42	67.39	12.14	0.93*	0.018
Choice	44	67.14	12.09	0.92**	0.004
Definition	44	66.67	7.96	0.91**	0.002
Combining	37	67.91	7.93	0.98	0.701

* and ** indicate significance at the 0.05 and 0.01 levels, respectively.

As Table 1 indicates, all four groups had similar EFL proficiency scores when they took the test one month previously. For the combining task, the data are approximately normally distributed, $w = 0.98$, $p > 0.05$, but for the other tasks, the assumption of normality was violated, $p < 0.05$ or 0.01. Levene's test found that the assumption of homogeneity of variances among the groups was violated, $F(3,163) = 2.97$, $p = 0.033 < 0.05$. A robust ANOVA using 20% trimmed means (X_t) and 20% Winsorized variances (s_w^2) found no statistically significant difference, $F_t (3, 55) = 0.15$, $p = 0.928 > 0.05$, so the four groups were considered equivalent in EFL proficiency.

Target Words and Sentence Reading Exercises

The participants in this study were at the same EFL proficiency level as those in Bao (2015), so the same 18 target words used in his study were also used in this study. These target words consisted of an equal number of nouns, verbs and adjectives. The nouns were *accessory, bristle, cabaret, fracture, gimmick, palette,* the verbs *allege, haunt, mumble, scrub, shudder, strangle,* and the adjectives *candid, cavalier, devious, egalitarian, erratic, malign.*

EFL reading exercises, serving as language input, comprised 18 semantically disconnected reading sentences, each of which involved a target word. These sentences were the same as in Bao (2015). The reading sentences were randomly divided into three sets of six sentences each. Immediately after each reading sentence, the gloss of a target word was given in the brackets, including its L1 (Mandarin Chinese) translations, part of speech and inflection. Below each reading sentence was a 4-point self-report scale (1= *not understood,* 2 = *partially understood,* 3 = *largely understood,* 4 = *totally understood).* The purpose of the scale was to encourage the participants to read each sentence carefully lest they should jump to the input or output task assigned to them.

Input and Output Tasks

In both the matching and definition tasks, several definitions/descriptions as one of the three sets were presented on the right, and on the left were the target words or the blanks to be filled in with the target words. The matching task was input-oriented, while the definition task output-oriented. The participants on the matching task were instructed to draw a line linking each target word to one definition/description. The purpose of giving more definitions/descriptions than the target words was to reduce reliance on guessing. The definition task omitted the distracters used in the matching task, and replaced the target words with blanks. The participants on this task were asked to fill in each blank with the target word matching each definition/description.

For both the choice and combining tasks, the 18 sentences were evenly divided into three sets of six sentences each. Each sentence was segmented into five word strings. In both tasks, the participants were asked to rearrange the word strings into a grammatically correct sentence. The choice task was input-oriented while the combining task output-oriented. In the choice task, the word strings, numbered 1 to 5, were randomly presented in each item stem, followed by four sequences numbered A to D for the participants to choose from, but only one sequence was grammatically correct. In the combining task, however, the numbers 1 to 5 in each item stem were omitted, and the adjacent word strings were separated with semicolons. The participants were asked to write down the correct sentence order in the line below the item.

All the tasks induced moderate need (+N), since the need was imposed by the task instructions. They demanded no search for the target words (-S), since the glosses of these words were given. Both the matching and definition tasks induced moderate evaluation (+), as evaluation "entails recognizing differences between words" (Hulstijn & Laufer, 2001, p. 544). Both the choice and combining tasks also induced moderate evaluation (+), as evaluation entailed comparing word strings and deciding how the target word or word string combined with others into a given sentence or text rather than requiring a decision as to "how additional words will combine with the new word in an original (as opposed to given) sentence or text" (ibid., p. 544). Therefore, all the tasks in this study were assigned an involvement load index of 2.

Vocabulary Knowledge Test Instrument

The vocabulary knowledge test instrument was an immediate posttest of the target words. It was adapted from Min's (2008) 4-point version of the five-point Vocabulary Knowledge Scale (VKS), which was developed by Paribakht and Wesche (1997). Min (2008) used the unknown/known word dichotomy to distinguish the first two statements (the unknown word category; Categories I and II) from the second two statements (the known word category; Categories III and IV). According to Paribakht and Wesche (1997), the first two categories were focused on form and meaning recognition (in a self-report form), whereas the last two were intended to measure meaning recall and comprehension (in a production task). This study employed Categories III and IV to measure the leaner's demonstrable vocabulary knowledge of each target word.

For the participants' understanding, the test instructions were given in Mandarin Chinese, their native language. The participants were presented with a list of the target words and instructed to indicate their levels of knowledge for each. For Category III, the participants were asked to write down the English definition or Chinese equivalent of each target word. Those who could complete Category III should proceed to Category IV, where they had to write a meaningful sentence with each target word.

Procedures

One week before the experiment, the researcher trained four EFL teachers on how to handle the materials, and answered questions about the instructions. The experiment was carried out during the regular class periods. Since the teaching schedules were not exactly the same, each teacher was allowed to administer the experiment in their own class sessions on a separate day of the same week.

Different time limits were set for each task (i.e., each vocabulary learning task plus sentence reading exercises). Following Bao (2015), the time limits of 15 and 30 min were set for the definition and combining tasks, respectively. Since the matching and definition tasks were similar in design format, and so were the choice and combining tasks, the matching and choice tasks were given the same time limits as the definition and combining tasks, respectively.

At the beginning of the experiment, all the participants were instructed to do the reading exercises first, followed by the assigned word-focused tasks. While performing the vocabulary learning tasks, the participants were allowed to refer to the reading materials for better understanding of the target words. 10 min after the teachers collected all the materials, the participants were unexpectedly given a 20-min posttest of the target words. All the test papers were collected by the teachers when the time was up.

Scoring and Data Analyses

This study dichotomized each participant's responses to Categories III and IV. For Category III, a score of 1 was awarded for a correct synonym or translation of the target word, but a score of 0 for no attempted response or an incorrect synonym or translation. No penalty was given for the wrong Chinese characters in the translation or replacing the correct Chinese characters with Chinese phonetic symbols as long as the correct meaning could be reasonably guessed. Regarding Category IV, a score of 1 was awarded if both the meaning and grammatical usage of the target word were correct in the sentence, regardless of errors elsewhere in the sentence, but a score of 0 was given otherwise.

Two experienced EFL teachers were trained to rate the participants' Categories III and IV for each target word independently. The inter-rater agreement was 100%, as all disagreements were discussed between the two teachers until consensus was reached. Each participant's vocabulary knowledge was represented by the cumulative scores of Categories III and IV.

To address the first research question, one robust one-way ANOVA was conducted on EFL vocabulary knowledge with task type as the between-subjects factor, followed by six post hoc linear contrasts (Hochberg's method was used to control the family-wise error rate). To address the second research question, one robust ANCOVA [3] was conducted on EFL vocabulary knowledge to test the task type effects at each of three design points of EFL proficiency (covariate). R 3.5.1 was run for all data analyses. All functions for robust estimates came from Wilcox (2017). The statistical significance level was set at $\alpha = 0.05$. One robust explanatory measure of effect size, called d_R, was used. Under normality and homogeneity, $d = 0.20, 0.50$ and 0.80 represent small, medium and large effect sizes, respectively (Cohen, 1988). In this study, this criterion for the magnitudes of effect sizes was followed, and a medium effect size or beyond was deemed important.

Results

Task Type and EFL Vocabulary Knowledge

This section graphically compares the input and output tasks in vocabulary learning. Figure 1 displays the data patterns of EFL vocabulary knowledge for each task.

Figure 1
Violin plots of vocabulary knowledge vs. task type
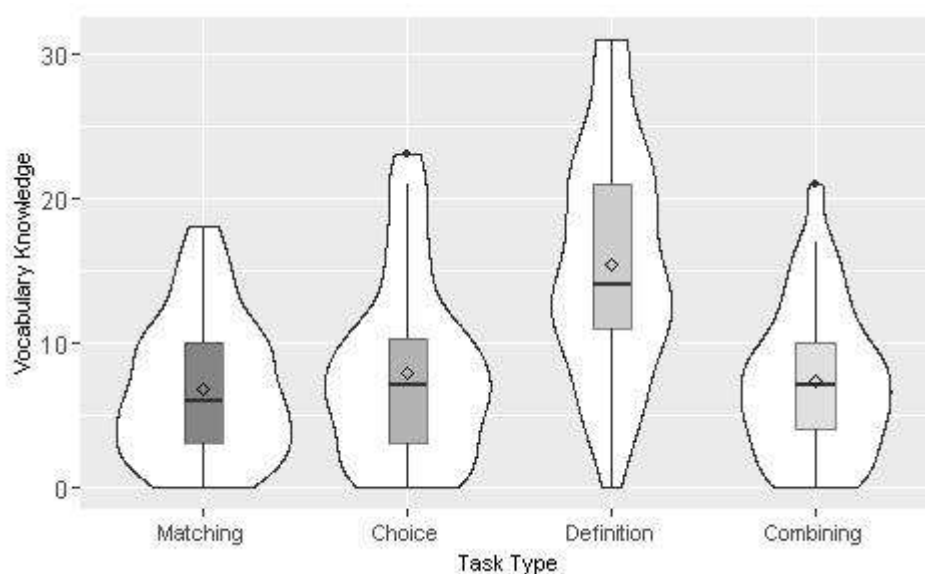
Figure 1 consitsts of four violin plots, i.e., kernel density plots superimposed in a mirror-image fashion over boxplots. Here, the boxes range from the lower to the upper quartile, the solid black line and the diamond in each box represent the median and mean respectively, and the black dots are outliers. Several features are discernible in this figure. To start with, the data distribution for each task is basically unimodal and right-skewed to a different extent. Furthermore, the data distribution for each task is somewhat platykurtic, especially for the definition task. Thirdly, one outlier is present for both the choice and combining tasks, but not for the other tasks. Finally, a comparison of the means and medians shows that the definition task perfroms remarkably better that the other three taks, which perform almost equally well.

Effects of Task Type on EFL Vocabulary Knowledge

Table 2 reports the descriptive statistics for the vocabulary knowledge data, including sample size (n), 20% trimmed mean ($\bar{X}_t$), 20% Winsorized variance (s^2_w), and 95% CI.

Table 2
Descriptive statistics for EFL vocabulary knowledge across task type

	n	$\bar{x}$	s^2_w	95% CI
Matching	42	6.31	10.65	4.58–8.04
Choice	44	7	14.37	5.05–8.95
Definition	44	15.18	29.03	12.40–17.96
Combining	37	6.83	13.93	4.71–8.95

As Table 2 shows, the definition task fares best in vocabulary learning, and its 95% CI of the trimmed mean does not overlap with any other task's, suggesting a significant difference between the definition and any of the other tasks. Much overlap in 95% CIs suggests no significant difference among the matching, choice and combining tasks.

A one-way robust ANOVA found that task type had a statistically significant effect on EFL vocabulary knowledge, $F_t(3, 55.32) = 12.37$, $p < 0.001$. Table 3 reports the results of each linear contrast.

Table 3
Tests of task differences in EFL vocabulary knowledge

	$\hat{\psi}$	p	$p.crit$	d_R
Matching-Choice	-0.69	0.579	0.017	0.13
Matching-Definition	-8.87+	< 0.001	0.013	1.27
Matching-Combining	-0.52	0.697	0.025	0.10
Choice-Definition	-8.18+	< 0.001	0.010	1.13
Choice-Combining	0.17	0.887	0.050	0.03
Definition-Combining	8.35+	< 0.001	0.008	1.14

$\hat{\psi}$ indicates a trimmed mean difference; *p.crit* refers to the critical value for a test of significance; + indicates significance at the specified critical *p* value.

As Table 3 indicates, all the differences between the definition task and any of the other tasks reach statistical significance, with very large effect sizes. All the differences among the matching, choice and combining tasks fail to reach statistical significance, with very small effect sizes.

Effects of Task Type on EFL Vocabulary Knowledge at Different Levels of EFL Proficiency
The previous section did not examine whether the task type effects were independent of EFL proficiency. In this section, a robust ANCOVA was conducted to examine the task type effects at each of the three EFL proficiency design points. These design points were 61.75, 69.25 and 74.71, representing the low, intermediate and high EFL proficiency level, respectively. Table 4 presents the results of the tests of task differences in EFL vocabulary knowledge at each design point.

Table 4
Tests of Task Differences in EFL Vocabulary Knowledge at Each EFL Proficiency Design Point

	Design point	n_i vs. n_j	$\hat{\psi}$	p	$p.crit$	d_R
Matching-Choice	61.75	23 vs.19	-2.85	0.197	0.025	0.41
	69.25	30 vs.33	-0.70	0.661	0.025	0.12
	74.71	29 vs.30	0.90	0.497	0.017	0.19
Matching-Definition	61.75	23 vs.10	-9.63+	0.000	0.010	1.55
	69.25	30 vs.30	-8.28+	< 0.001	0.01	1.36
	74.71	29 vs.22	-9.41+	< 0.001	0.013	1.41
Matching-Combining	61.75	23 vs. 21	0.46	0.792	0.050	0.08
	69.25	30 vs.27	0.09	0.918	0.050	0.02
	74.71	29 vs.22	0.88	0.550	0.025	0.17

	61.75	19 vs. 10	-6.78	0.029	0.013	0.82
Choice-Definition	69.25	33 vs.30	-7.58	0.0005	0.013	1.07
	74.71	30 vs.22	-10.31	< 0.001	0.010	1.55
	61.75	19 vs. 21	3.31	0.138	0.017	0.46
Choice-Combining	69.25	33 vs.27	0.79	0.618	0.017	0.12
	74.71	30 vs.22	-0.02	0.985	0.050	0.01
	61.75	10 vs. 21	10.09	0.000	0.008	1.55
Definition-Combinin g	69.25	30 vs.27	8.37	< 0.001	0.008	1.25
	74.71	22 vs.22	10.29	< 0.001	0.008	1.44

$\psi^{\wedge}$ indicates a trimmed mean difference; *p.crit* refers to the critical value for a test of significance; + indicates significance at the specified critical *p* value.

As Table 4 indicates, at each EFL proficiency level, the definition task outperforms all the other tasks significantly, with a large or very large effect size. No significant difference was found among the matching, choice and combining tasks at each EFL proficiency level. Nevertheless, at the low EFL proficiency level (design point = 61.75), given the observed low to medium-sized and medium-sized effect sizes, the small advantage of the choice task over the matching and combining tasks should be taken note of.

Discussion

This study investigated the task type effects on EFL learners' vocabulary knowledge, followed by examination of whether such effects were independent of EFL proficiency. Of the four types of task investigated, the definition task was found to outperform all the other tasks in promoting EFL vocabulary knowledge, and the matching, choice and combining tasks performed equally well. When EFL proficiency was considered, the definition task was also found to outperform all the other tasks at each level of EFL proficiency, and the matching and combining tasks showed no difference, irrespective of EFL proficiency. The choice task had an advantage over the matching and combining tasks at the low level of EFL proficiency, but not at the intermediate or high level of EFL proficiency. The task type effects are largely independent of EFL proficiency, a finding similar to Kim (2008). It seems that, once low-proficiency EFL learners have reached a level of proficiency sufficient to allow them to complete common vocabulary-focused tasks, it is possible for them to benefit from the tasks as equally or nearly equally as high-proficiency EFL learners. It should be noted, however, that the EFL learners in this study were at the same grade level, and their differences in EFL proficiency might not be as large as they appeared. Thus, it would be of great interest to examine whether the present findings would still hold across a wider range of proficiency levels.

With regard to the ILH, only partial support was provided, since all the four tasks were not equally effective in promoting vocabulary knowledge. The task type effects allow for explanations other than the ILH. In what follows, the effects are explained mainly in terms of differential processing and division of attention.

Differential Processing

The word exposure frequency effect, evidenced by some studies (e.g., Eckerth & Tavakoli, 2012; Laufer & Rozovski-Roitblat, 2011) may partly account for the present findings. Take for example the comparison between the definition and combining tasks. The definition task increased the frequency at which each target word was encountered and evaluated, thus reinforcing the form-meaning connection in the learner's mental lexicon. Such frequent encounters of the target words could not be expected to occur in the combining task, where target words were given individually, and no comparisons between them were needed. The word exposure frequency effect might also account for the small advantage of the choice task over the matching or combining task when the learners were compared at the low EFL proficiency level. The choice task, where the target word appeared in each of the four options, might offer the learners multiple exposures to the target word. Compared to the learners at the intermediate or high EFL proficiency level, those at the low EFL proficiency level were more likely to repeatedly compare the four options concerning the target word in order to make a

correct choice. Consequently, multiple exposures to the target word might have increased the learners' chances of knowing the word.

Perhaps more importantly, the superiority of the definition task over the other tasks was due to more retrieval or mental effort. The learners doing the definition task had to infer a conceptually familiar L1 word equivalent corresponding to each definition/description, and then return to the reading exercises to find the target word whose actual L1 equivalent was the same as or similar to the inferred one. In order to ensure the form-meaning correspondence, the learners might have compared and evaluated the target words and the definitions/descriptions repeatedly. The forced output facilitated the definition task learners' access to the target word forms. The mental effort demanded by inferring and forced production might have greatly strengthened the form-meaning connection. This also explains Wesche and Paribakht's (2000) finding that interpretation of form-meaning relationships in the definition task induced more mental effort than recognition of form-meaning relationships in the matching task.

In the final analysis, differential processing could probably account for task effectiveness for vocabulary learning. Unlike the other tasks, the definition task could induce both elaborate structural (i.e., orthographic) and elaborate semantic processing of the target words. Although the matching task provided the same rich contextual cues as did the definition task, the learners on the matching task might have found no need to infer the target words, which had already been given. Even though the matching task learners might compare or evaluate the different target words and their corresponding definitions/descriptions, such comparisons or evaluations did not warrant the same semantic elaboration as did the definition task. A similar case is made for the choice and combining tasks. Like the matching task, these two tasks did not require effortful retrieval but recognition of the target words. The comparison between the definition and the other tasks reveals the importance of both structural and semantic processing in vocabulary learning.

Division of Attention

One may still wonder why the combining task was not more effective for vocabulary learning than the choice task. Intuitively, it appears reasonable to anticipate that the combining task would perform better than the choice task, not only because the former took longer than the latter, but also because the former was a recognition task, whereas the latter was a production one. More time on task does not necessarily lead to better retention of the target words, however. For instance, Craik and Tulving (1975) compared reaction time among three levels of processing, i.e., questions concerning type-script (structural level), rhyme questions (phonemic level), and sentence questions (semantic level), finding that slow responses were recognized little better than fast responses at each level of processing. In the same vein, Hill and Laufer (2003) found that task effectiveness could be attributed to task type rather than time on task. In the current study, the choice and combining tasks proved to be more complex than the matching task or perhaps even the definition task, because the learners doing these tasks, especially those doing the combining task, had to understand the meanings of all word strings and analyze the syntactic relations among them so as to rearrange the word strings into a proper sentence. However, like time on task, task complexity does not necessarily contribute to better word retention, either. For example, Joe (1998) found that an experimental task, where the adult L2 learners received explicit instruction on generative tactics and retold the passage without the aid of the text, did not perform better in word retention than a comparison task, where the learners did not receive explicit instruction but had the text available to them while retelling, although the experimental task was expected to outperform the comparison task. The increased task demands and higher learning burden imposed on the experimental group may have led to their failure to outscore the comparison group (Joe, 1998, p. 373).

In this study, although processing the target word was intended to be the primary activity, and unscrambling word strings into a proper sentence the secondary activity, the choice and combining tasks, especially the combining task, were so demanding (e.g., in terms of syntactic knowledge) that the learners might have diverted their attention to the secondary activity, resulting in shallower encoding of the target word processed in the primary activity. Still possibly,

various requirements dictated by the combining task depleted the learners' attentional resources, and this in turn led to faint memories of the new words which could possibly have been processed elaborately. This may explain why the combining task had no advantage over the matching or choice task. It should be noted that, though, the participants in this study were not advanced EFL learners, and their syntactic competence was not developed fully. Equipped with more syntactic competence, EFL learners doing the combining task would have been less likely to divert much attention to those requirements not directly related to the target words, and thus might have memorized the target words better, since the sentence context could have consolidated the form-meaning connection, or even the word usage, a claim that remains to be verified.

Conclusion

This study examined EFL learners' initial vocabulary learning through reading sentences and performing tasks related to the target words. Four tasks, involving two input (matching and choice) and two output tasks (definition and combining), were compared in EFL vocabulary knowledge acquisition. The definition task outperformed all the other tasks across EFL proficiency levels. The matching and combining tasks performed equally well, regardless of EFL proficiency. The choice task gained a small advantage over the matching and combining tasks when the learners were at the low level of EFL proficiency, but this advantage disappeared when the learners arrived at the intermediate or high EFL proficiency level.

This study contributes to a better understanding of what characteristics of a word-focused task determine its effectiveness for word retention. Differential processing may well account for the superiority of the definition task over the other tasks in retention of new words, suggesting the importance of both structural and semantic processing. Thus, in order for the ILH to better explain or predict task effectiveness, the notion of evaluation may need to be extended to cover both structural and semantic elaboration instead of the latter only. Division of attention was another probably important factor affecting word retention, since the task demands could direct or divert the learners' attention to the new words. Therefore, although all vocabulary learning tasks can be labeled as "word-focused", it does not follow that they would induce the same amount of learners' attention to new words. Divided attention, which depends on the task requirements and complexity, would probably reduce word retention even if the new words were once processed more deeply or elaborately.

These findings have potential implications for EFL vocabulary instruction. EFL teachers should keep in mind that neither more task time nor a more complex task is necessarily beneficial for vocabulary learning, since word retention depends much more on what kind of processing the task elicits and how much of the learners' attention the task directs to the new words. They are advised to design tasks which can induce access to both word form and meaning to help the learners consolidate the form-meaning connection. EFL teachers should also be aware that, at the initial stages of EFL vocabulary learning, attention to the structural properties of a new word may be of vital importance, since EFL vocabulary acquisition "typically does not involve learning new concepts while learning new word forms", as is true of L1 vocabulary acquisition (Barcroft, 2002, p. 356). Specifically, in designing vocabulary learning tasks with identical involvement loads, EFL teachers are encouraged to employ tasks like the definition task more frequently than those like the matching task if the primary pedagogical goal is to improve vocabulary learning. When the learners are just beginning to encode new EFL words, tasks like the combining task is not recommended, for those tasks may exhaust learners' processing and attentional resources which could otherwise be employed to process new word forms. If improving EFL learners' syntactic ability is the main pedagogical goal, however, tasks like the choice or combining task might be in order, with acquisition of new words being a by-product. For learners with low EFL proficiency, the choice task seems preferable to the combining task in contributing to vocabulary learning.

Notes

1 The 20% trimmed mean is computed by first removing 20% of the smallest and largest sample values and then averaging what remains.

2 The 20% Winsorized variance is the variance of the 20% Winsorized values, which are derived by pulling the smallest 20% of the sample observations up to the smallest value not trimmed, and the largest 20% of the sample observations down to the largest value not trimmed.

3 A robust ANCOVA picks points (design points) on x (covariate), and then compares the 20% trimmed y (dependent variable) means for all independent groups, based on the y values for each group corresponding to the x values in the neighborhood of each design x point. A robust ANCOVA allows nonnormality, heteroscedasticity, and even curved regression lines. The Benjamini-Hochberg method was used to control the family-wise error rate at each design point. See Wilcox (2017) for more information.

Acknowledgments

References

Bao, G. (2015). Task type effects on English as a Foreign Language learners' acquisition of receptive and productive vocabulary knowledge. *System, 53*, 84-95.

Barcroft, J. (2002). Semantic and structural elaboration in L2 lexical acquisition. *Language Learning, 52*(2), 323-363.

Bird, S. (2012). Expert knowledge, distinctiveness, and levels of processing in language learning. *Applied Psycholinguistics, 33*(4), 665-689.

Brown, T. S., & Perry, F. L. Jr. (1991). A comparison of three learning strategies for ESL vocabulary acquisition. *TESOL Quarterly, 25*(4), 655-670.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.

Craik, F. I. M., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior, 11*, 671-684.

Craik, F. I. M., & Tulving, E. (1975). Depth of processing and the retention of words in episodic memory. *Journal of Experimental Psychology: General, 104*(3), 268-294.

Eckerth, J., & Tavakoli, P. (2012). The effects of word exposure frequency and elaboration of word processing on incidental L2 vocabulary acquisition through reading. *Language Teaching Research, 16*(2), 227-252.

Folse, K. S. (2006). The effect of type of written exercise on L2 word retention. *TESOL Quarterly, 40*(2), 273-293.

Golonka, E., Bowles, A., Silbert, N., Kramasz, D., Blake, C., & Buckwalter, T. (2015). The role of context and cognitive effort in vocabulary learning: A study of intermediate-level learners of Arabic. *The Modern Language Journal, 99*(1), 19-39.

Huang, L.-L., &. Lin, C.- C. (2014). Three approaches to glossing and their effects on vocabulary learning. *System, 44*, 127-136.

Hill, M., & Laufer, B. (2003). Type of task, time-on-task and electronic dictionaries in incidental vocabulary acquisition. *IRAL, 41*, 87-106.

Hulstijn, J., & Laufer, B. (2001). Some empirical evidence for the involvement load hypothesis in vocabulary acquisition. *Language Learning, 51*(3), 539-558.

Joe, A. (1998). What effects do text–based tasks promoting generation have on incidental vocabulary acquisition? *Applied Linguistics, 19*(3), 357-377.

Keating, G. (2008). Task effectiveness and word learning in a second language: The involvement load hypothesis on trial. *Language Teaching Research, 12*(3), 365-386.

Kim, Y. (2008). The role of task–induced involvement and learner proficiency in L2 vocabulary acquisition. *Language Learning, 58*(2), 285-325.

Laufer, B., & Girsai, N. (2008). Form-focused instruction in second language vocabulary learning: A case for contrastive analysis and translation. *Applied Linguistics, 29*(4), 694-716.

Laufer, B., & Hulstijn, J. (2001). Incidental vocabulary acquisition in a second language: The construct of task–induced involvement. *Applied Linguistics, 22*(1), 1-26.

Laufer, B., & Rozovski-Roitblat, B. (2011). Incidental vocabulary acquisition: The effects of task type, word occurrence and their combination. *Language Teaching Research, 15*(4), 391-411.

Min, H. (2008). EFL vocabulary acquisition and retention: Reading plus vocabulary enhancement activities and narrow reading. *Language Learning, 58*(1), 73-115.

Paribakht, T. S., & Wesche, M. B. (1997). Vocabulary enhancement activities and reading for meaning in second language vocabulary acquisition. In J. Coady & T. Huckin (Eds.), *Second language vocabulary acquisition: A rationale for pedagogy* (pp. 174-200). Cambridge: Cambridge University Press.

Pichette, F., de Serres, L., & Lafontaine, M. (2012). Sentence reading and writing for second language vocabulary acquisition. *Applied Linguistics, 33*(1), 66-82.

Sagarra, N., & Alba, M. (2006). The key is in the keyword: L2 vocabulary learning methods with beginning learners of Spanish. *The Modern Language Journal, 90*(2), 228-243.

Webb, S. (2005). Receptive and productive vocabulary learning: The effects of reading and writing on word knowledge. *Studies in Second Language Acquisition, 27*(1), 33-52.

Webb, S., & Kagimoto, E. (2009). The effects of vocabulary learning on collocation and meaning. *TESOL Quarterly, 43*(1), 55-77.

Wesche, M. B., & Paribakt, T. S. (2000). Reading–based exercises in second language vocabulary learning: An introspective study. *The Modern Language Journal, 84*(2), 196–213.

Wilcox, R. R. (2017). *Introduction to robust estimation and hypothesis testing* (4th ed.). San Diego, CA: Elsevier.

About the Author

Gui Bao, PhD, is a Professor of Applied Linguistics at Nanjing Tech University, Nanjing, China. His research interests cover second language research methodology, second language vocabulary acquisition and applied statistics for linguistics.

Profiling Performances of L2 Listenership: Examining The Effects of Individual Differences in The Japanese EFL Context

Pino Cutrone[a]
Nagasaki University, Japan

Abstract

This paper describes a study designed to profile performances of L2 listenership. The writer examines the listenership behavior of 23 Japanese EFL learners, who were all freshmen students at a national university in Japan (16 females and 5 males) at the time the study was conducted, in an attempt to identify some of the features associated with different levels of performance concerning listenership behavior. Specifically, this study sought to identify some of the common characteristics of Japanese EFL participants who exhibited competent backchannel behavior compared to those who did not. Assessments involved having each student participate in an intercultural conversation, complete a questionnaire, and be interviewed. Quantitative and qualitative data analysis methods were used to investigate the relationships between variables (i.e., whether various performances in sub-categories of listenership are interrelated, as well as how individual performances in sub-categories of listenership may be related to L2 proficiency, personality dimensions, willingness to communicate, etc.). Besides helping to provide researchers with a more detailed description of the dynamics of listenership/backchannel behavior, the results of this study will have practical implications for Japanese EFL practitioners.

Key Words: listenership, backchannel behavior, Japanese EFL context, individual differences, pragmatics

Introduction

This paper attempts to piece together a profile of successful versus non-successful learners where listenership/backchannel behavior is concerned. The first step is to provide a clear and concise definition of what a backchannel is. While several different definitions of the term exist in the research literature (see Fujimoto, 2007 for a list of 24), backchannels can be understood in general terms as "the brief verbal and nonverbal responses and/or reactions that a listener gives to the primary speaker when the primary speaker is speaking" (Cutrone, 2011, p. 53). To understand what this means, it is necessary for readers to also be familiar with the notion of turn taking in conversations. Thus, when one person is taking a turn at speaking in the conversation, they are considered the primary speaker, and their talk is the main channel of communication. The listener is then considered the non-primary speaker and their utterances during the primary speaker's turn are backchannels, which in turn serve to provide short cues to notify the primary speaker that the non-primary speaker is listening. In other words, the primary speaker is the one that is carrying conversation and driving it forward (i.e., has the floor and on topic), while a non-primary speaker (i.e., listener) is the one that is reacting to what the primary speaker is saying. Example I illustrates this difference:

> I. Carrie: In some high schools in America, they offer Japanese.
> Akie: Uhum.

Nonverbal and non-word vocalizations such as head nods and laughter respectively can be considered backchannels if they serve a listening function. For instance, in Examples II and III, it is clear that Akie's head nods (shown by the symbol ∧) and laughter respectively are reactions to Carrie's statements.

> II. Carrie: New York's China town is very huge.
> Akie: ∧∧

EMAIL: k-pino@nagasaki-u.ac.jp, School of Global Humanities and Social Sciences, 1-14 Bunkyo-machi, Nagasaki City, Nagasaki Prefecture, Japan, 852-8521

III. Carrie: That is so cute because it looks like a grade schooler.
Akie: (*Laughter*)

Examples used in this paper, such as I, II, and III shown above, have been taken from authentic conversations produced in this study, as shown in Appendix A. In some cases, the examples have been modified and/or various aspects of the transcription conventions have been omitted in order to make them easier to understand. Such examples are used simply to provide models of backchannel behavior occurring in naturally occurring speech, and, thus, more in-depth analysis of the issues involved in creating and deciphering of conversational transcriptions are needed.

Differentiating a Backchannel from a Turn

One of the most difficult issues in identifying a backchannel seems to be in determining whether a behavior constitutes a backchannel or a separate turn (i.e., a sub-issue is whether to include longer reactive utterances as backchannels or not). Ergo, it is necessary to be able to understand and identify specifically what constitutes a turn in this study. In their seminal work, Sacks, Schegloff, and Jefferson (1974) proposed a model for the organisation of turn-taking in conversations in which they describe a turn to consist of one or more *turn-constructional units* (TCUs). According to their model, TCUs can range in size from a single word to clauses filled with many embedded clauses. Each TCU ends at a *transition-relevant-place* (TRP)*,* which is identified as a moment in the conversation at which an exchange of turn is appropriate. TRPs are signalled by the conversation's participants to each other through various contextual cues such as *silence* or the *end of a question.* TRPs are commonly observed in similar conversational contexts as backchannels (Clancy, Thompson, Suzuki, & Hongyin, 1996; Cutrone, 2014; Maynard, 1997; White, 1989).

Although Sacks, Schegloff, and Jefferson's (1974) model is useful for understanding the general set of rules that govern the turn-taking system, it may not be the most suitable for identifying backchannels in this study, as it does not account for the concept of *having the floor* (Edelsky, 1981; Hayashi, 1988). While the definition of a TCU is primarily grammatical, the concept of having the floor is based on participants' sense of who has the floor and is on topic, as well as the quantity and frequency of their speech. The concept of having the floor does not seem to fit within the framework of CA in terms of identifying backchannels since a speaker could continue to hold the floor while non-floor holders ask questions and/or make comments to drive the floor holder into new directions of conversations. Such questions and comments would constitute full turns in the field of CA, whereas they would not necessarily do so in terms of having the floor.

In this study, the writer approaches the observation of listening behavior not only from a research perspective that relies mainly on providing further descriptions of this phenomenon; but also, in the context of this study, the writer is concerned with how listening behavior is used in, and affects, real-world intercultural communication (IC). Hence, as Fujimoto (2007) and Thonus (2007) have suggested, it may be more practical from such a perspective to consider backchannels as any listener response that reacts to what the primary speaker has said. Following O'Keeffe and Adolphs (2008), the term 'listener response' is used as an umbrella term to describe any response which reacts to something that the primary speaker has said (p. 74). In the context of this study, backchannels/listener responses would extend beyond what is meant by the term backchannel in many other studies to also encompass longer utterances which also act in response to an interlocutor's utterance. The rationale for this becomes clear below, where the framework for assessing listening behavior is detailed. Within this framework, the writer employs Markel's (1975) definition of turn to analyze listener responses in this study:

> A speaking turn begins when one interlocutor starts solo talking. For every
> speaking turn there is a concurrent listening turn, which is the behaviour of
> one or more nontalking interlocutors present. (p. 190)

Hence, in this turn-taking system, the only time that a change in speaking turn occurs is

when the non-primary speaker begins solo speaking, which is recognized here as some point or utterance made which serves to actually advance the conversation (i.e., this does not include short backchannel utterances such as *uhuh, mmm* and/or *I see,* which seem only to serve a listening and reactive function). In instances where simultaneous speech occurs, the primary speaker continues to *have the turn* if the primary speaker continues to solo speak after the simultaneous speech. However, if the non-primary speaker begins solo speaking after the simultaneous speech, then a change of primary speaker turns would have occurred. Within this framework, brief questions such as *Really?* or *Is that right?*, which are formed in terms of requests for clarification, are considered backchannels, as they are thought to primarily serve a listening function. In contrast, a question such as *Why did she move?* is considered a full speaking turn because it serves a speaking function in terms of driving the conversation in a new direction.

Thus, responses to questions are considered full speaking turns and not backchannels. That is because, unlike responses to questions, backchannels are optional and not required (Ward & Tsukuhara, 2000). Further, responses to questions, even when they are quite brief (often due to ellipsis), would also seem to provide new information that helps steer the conversation forward constituting a change of primary speakership. Finally, researchers must decide how to deal with utterances found between turns at talk, i.e., would such statements be recognized as listening reactions or part of a turn at talk? Following the writer's previous analyses (Cutrone, 2005, 2014), utterances were considered listener responses in this study only when they occurred immediately after the primary speaker stopped talking (within one second) and were followed by a substantial pause before the next turn at talk started (exceeding one second). This decision was made because it was felt that such listener responses were produced in response to the primary speaker's utterance, and they occurred before a substantial turn transitional period started.

Types of Backchannels

Listener responses are recognized to occur as verbal backchannels and/or nonverbal ones. According to Tottie's (1991) oft-used classification, verbal backchannels in this study are grouped according to three types: simple, compound, and complex. To illustrate this distinction, it is useful to also understand the difference between a backchannel and a backchannel item. A simple backchannel such as *uhuh* is one which has only one backchannel item. A compound backchannel such as *yeah yeah yeah* is one in which one backchannel item exists but is repeated more than once. A complex backchannel such as *yeah, I know* consists of multiple and varied backchannel items. Nonverbal backchannels, which can occur both simultaneously and independently of the three verbal types above, fall within the following categories: simple accompanied by a head nod(s), compound accompanied by a head nod(s), complex accompanied by a head nod (s), isolated head nod, multiple head nods, smile, laughter, raised eyebrows, and two or more nonverbal backchannels occurring simultaneously.

A broader categorical distinction involving listener responses is presented in Stubbe's (1998) *feedback continuum* (p. 259). At one end of the continuum is listener feedback, which is brief and minimally supportive, while at the other end is lengthier feedback which conveys a higher degree of involvement in the conversation. Following this framework, minimal responses are defined as any simple verbal backchannel (including non-word vocalizations such *uhuh* or *mm*) and/or nonverbal backchannel occurring in isolation. Extended responses, in contrast, are defined as the lengthier, verbal listener feedback consisting of multiple and varied words as characterized by complex backchannels, irrespective of nonverbal backchannel accompaniment.

Functions of Backchannels

The most common function of a backchannel, to allow the primary speaker to continue speaking, is deeply embedded in navigating the turn-taking system and specifically on the non-primary speaker forsaking the opportunity to take a primary speaking turn (Schegloff, 1982). This clearly demonstrates the apparent link between how much (or little) a person speaks with how frequently (or infrequently) they provide backchannels. Several intercultural

analyses involving Japanese EFL speakers have shown a relative lack of primary speaker incipiency in tandem with the frequent use of backchannels, which seems to negatively affect perceptions across cultures. In addition to the continuer function described above, Maynard (1997) identifies a few more prominent backchannel functions, such as to show understanding, agreement, support and empathy, emotion, as well as to include minor additions (see further explanations and examples of these functions in Cutrone, 2005).

Listenership across Cultures and Targets for Listenership Behavior

The followings sub-sections will identify some of the areas of listenership and put forward some general targets for Japanese EFL/ESL speakers (JEFLs hereafter) to adhere to in their intercultural encounters in L2 English. The establishment of such targets was based on two goals: (1) trying to approximate the conversational patterns and behaviors of fully proficient speakers of English, and (2) dealing with the issues that Japanese EFL/ESL speakers have been known to have where listenership behavior is concerned. Therefore, more specifically, this involves having JEFLs provide minimal backchannels less frequently (especially while one's interlocutor is speaking), with greater variability (but at context-appropriate moments), while asking questions and taking the primary speakership in the conversation more often, and initiating conversational repair strategies when they do not understand and/or disagree rather than feign understanding and agreement. With these targets in mind, it should be noted that there can exist a great deal of individual differences in listenership behavior within any given culture or group. Thus, since listenership behaviors are often individualistic and context-driven (and contain considerable overlap between sub-categories), it does not seem wise to prescribe quantifiable targets in precise terms. Rather, based on the recorded observations of fully proficient speakers of English in the literature, which were limited to native English speakers (NESs hereafter) where listenership behavior was concerned, the targets in the following list provide practitioners (i.e., teachers and users of the language) with general directions for assessing various aspects of backchannel behavior.

Target 1: Approximating the Listenership Behavior of Proficient Speakers of English Overall Frequency

Several studies have reported JEFLs uttering backchannels significantly more than NESs (Clancy et al., 1996; Crawford, 2003; Cutrone, 2005, 2014; Ike, 2010; Maynard, 1986, 1990, 1997; White, 1989). Depending on the study, JEFLs have been observed to send anywhere from two to four times as many backchannels as NESs. Although the context of the conversation will always be the overriding factor as to when they should send backchannels, in general terms, one goal for JEFLs is to backchannel less and eliminate many of the superfluous, empathy-building backchannels that they provide in English.

Variability

One of the findings in Cutrone's (2005, 2014) previous analyses is that the JEFLs in the researcher's studies tended to rely mainly on minimal backchannels (i.e., short, brief and repetitive non-word vocalizations, and head nods) in their listener feedback and this was perceived negatively in ELF (English as a Lingua Franca) conversations across cultures. The NESs in Cutrone's studies, comparatively, tended to balance minimal backchannels more evenly with extended backchannels (i.e., longer backchannels consisting of content words, phrases, and expressions). Hence, in general terms, another goal is for JEFLs to develop a more diverse repertoire of backchannels to use in their intercultural encounters in English. In other words, JEFLs should work towards increasing the number of extended backchannels and decreasing the number of minimal backchannels they produce. The former objective operates in tandem with increasing WTC (willingness to communicate), which will be discussed in Target 2 below.

Discourse Contexts Favoring Backchannels

This category encompasses a term coined by Maynard (1986) used to describe the locales in the primary speaker's speech where backchannels are commonly found and includes primary

speaker behaviors that seem to attract listener feedback. In several studies, grammatical completion points (i.e., phrasal and clausal boundaries) and pauses (especially occurring simultaneously) have been identified as primary discourse contexts favoring backchannels in English (Cutrone, 2005, 2014; Maynard 1986, 1990, 1997; White, 1989). Several other discourse contexts including self-adaptors and gesticulation (Duncan & Fiske, 1977), gaze (Kendon, 1977), and prosodic features (Ward &Tsukuhara, 2000) have also been suggested in the research literature. This is a difficult category to offer precise targets since backchannels that are sent in locales other than the ones mentioned above are not necessarily considered incorrect and their adequacy is largely dependent on the context of the conversation and the function that the non-primary speaker desires to communicate. Nonetheless, as a general method of measuring performances, the researcher compares the percentage of backchannels the participants employed in the opportunities they were provided in primary discourse contexts such as grammatical completion points and/or pauses.

Simultaneous Talk

At rates similar to those of overall frequency, several studies have shown that, when compared to NESs, JEFLs tended to send backchannels that co-occur with the primary speaker's speech creating simultaneous talk much more frequently (Cutrone, 2005, 2014; Hayashi, 1998; Lebra, 1976; Maynard, 1997; Mizutani, 1982). In tandem with the general targets of discourse contexts favoring backchannels presented above (i.e., to send backchannels at grammatical completion points and/or pauses), the goal for JEFLs here is to generally try to avoid sending backchannels while their interlocutor is speaking.

Form and Function

Choosing suitable and appropriate linguistic forms to correspond with specific (and desired) functions of backchannels may have the greatest impact on one's communicative effectiveness. With this in mind, success in this area will be extremely difficult to measure because it is highly context driven and largely dependent on the individual intentions and feelings of the person providing the backchannels. Further, success will be very difficult to measure quantitatively, as there is considerable overlap between forms and functions (see Cutrone, 2010 for a sample inventory of backchannel forms corresponding to the functional categories presented above). Here, the writer addresses some of the JEFLs' unconventional uses of backchannels in English such as their tendency to employ continuer, understanding, agreement and/or support, and empathy type backchannels in situations when they did not understand what their interlocutor was saying. In tandem with certain aspects of Target 3 below (i.e., initiating conversational repair strategies), the main goal here is for JEFLs to convey their feelings with appropriate backchannel forms, and for these intentions to be recognized accordingly by their interlocutors.

Target 2: WTC and Conversational Involvement

In various intercultural analyses, JEFLs' reticence and minimal responses have been cited by NESs as reasons negatively affecting IC (Anderson, 1993; Cutrone, 2005, 2014; Sato, 2008). Demonstrating how backchannel categories are highly interconnected, this is yet another category that overlaps with others. That is, if JEFLs make a concerted effort to initiate conversation more (which is the goal here) they will, in turn, backchannel less (which was one of the goals stipulated in Target 1 above). Similarly, the goal of providing more extended backchannels over minimal ones (as stipulated in the *variability* sub-category of Target 1) would seem to fit in well with the goal here of speaking more. Lastly, in the same way, the goal of employing full speaking turns (as touched upon in the *form and function* sub-category of Target 1 and to be discussed again in Target 3 from a conversational management perspective) instead of backchannels to get over certain obstacles that come up in a conversation is also in line with the goal of more conversational involvement. The JEFLs' involvement and WTC in conversations will be measured in three ways in this study: (1) WTC scores (using the widely used WTC scale designed by McCroskey, 1992, see Appendix B), (2) how much they spoke in the conversations, and (3) the number of questions they asked their interlocutor.

Target 3: Conversational Micro-Skills

As stated above, one of the goals concerning more effective listenership behavior is for JEFLs to exhibit a higher degree of WTC and conversational involvement. To this end, JEFLs would be well advised to make use of conversational management techniques, which refer to the ability to effectively incorporate the following in conversations: appropriate usage of discourse markers and listener responses, evaluative comments, return questions, follow-up questions, new topic initiation, expansion techniques, the ability to ensure comprehension on the part of the listener, and the ability to initiate repair when there is a potential breakdown. Concerning the latter (which was introduced in the *form and function* aspect of Target 1), JEFLs must deal specifically with their issue of sending unconventional backchannel types (such as continuer, agreement, and understanding listener responses) when they do not understand what the interlocutor is saying.

Target 4: Intercultural Communicative Competence (ICC)

The final criterion for JEFLs to demonstrate effective listenership behavior involves exhibiting a certain degree of ICC. According to Spitzberg's (2000) well-known model of ICC, the optimal conditions for successful ICC are provided when knowledge, skills, and motivation are aligned with meeting the other person's expectations regarding appropriateness and effectiveness. Thus, any instrument seeking to measure appropriate and effective listenership behavior in intercultural conversations must consider perceptions across cultures. From the perspective of JEFLs, meeting the expectations of the global ELF community is paramount to achieving success in this area. To measure this aspect of listenership, this study will use Hecht's (1978) Interpersonal Communication Satisfaction Inventory (see Appendix C). This is useful in assessing interlocutors' listenership behavior, conversational satisfaction, and perceptions of one another after conversing.

Research Questions

Research into the area of listenership behavior, particularly concerning individual differences, is in its infancy and, thus, much remains unknown. Most studies to date have focused primarily on detailing the patterns of the backchannel output of various groups in terms of frequency and discourse contexts favoring backchannels, and to a somewhat lesser extent the variability and backchannels creating simultaneous speech. Presently, very little is known about the characteristics of both successful and unsuccessful communicators where listenership is concerned. As the writer mentioned in Target 3 above, the degree of an individual's WTC may affect their conversational performances (and listenership) and is worthy of more in-depth exploration. Another variable that will be investigated in this study is the extraversion/introversion dimension of personality. Extraversion is particularly relevant to this study, as it has traditionally been thought to be at the centre of personality models (Eysenck, 1992), and, similar to the WTC construct, has been shown to affect L2 use (Dewaele & Furnham, 2000). Further, concerning the four targets for listenership provided above, it is not known if success or failure in one area will correspond to success or failure in others. Accordingly, RQ 1 attempts to detail some of the features associated with different levels of performance concerning listenership behavior.

> RQ 1: What are some of the common characteristics pertaining to the JEFLs that demonstrated competent listenership behavior compared to the JEFLs that did not?

Similar to RQ 1, the objective of RQ 2 is also to contribute to the profile of successful and unsuccessful learners with regards to listenership behavior. To this end, RQ 2 investigates L2 proficiency as a moderator variable. Concerning L2 proficiency, researchers such as Heffernan and Jones (2005) have attempted to use individual differences to create profiles of successful JEFLs. Concerning pragmalinguistic features of language such as listenership behavior, it is not yet clear how L2 proficiency affects the learning of such targets. Researchers in the area of L2 listenership such as Thonus (2007) have surmised that instruction on

listenership behavior is best begun at intermediate levels; however, to date, this hypothesis has not been tested. Likewise, there seems to be a general assumption of a strong correlation between L2 proficiency and successful listenership behavior that requires empirical validation. Hence, RQ 2 has been formulated.

> RQ 2: Do student L2 proficiency levels (according to the TOEFL) correlate to their levels in listenership behavior?

In short, this study attempts to piece together profiles of performance associated with listenership behavior by examining the effects of variables such as L2 proficiency, extraversion, and WTC.

Methodology

Participants

The study included 26 participants. The 21 student participants were all first year students at a national university in southern Japan (16 females and 5 males), who were enrolled in a faculty that focuses on the study of global humanities and social sciences and that emphasizes the study of English. When this study began, participants were on average at an intermediate level of English proficiency (as reflected by their TOEFL PBT scores), between 18 and 20 years old, and had studied English for 6.5 years on average (including a collective six years in junior and senior high school). The students had been enrolled at the university for six months and were all taking courses related to English study. Additionally, this study included 5 NES participants to serve in two capacities: two of the participants (1 female and 1 male) served as interlocutors in the intercultural dyadic conversations and three of the participants (2 female and 1 male) were involved in assessing the JEFLs' performances in the video recorded conversations. Participating of their free will and understanding the nature of the study, all participants were given explicit instructions (i.e., verbal and written, in both English and Japanese) regarding this study and their role in it. To protect the identity of the participants, pseudonyms are used in this paper.

Instruments

Three methods of measurement were used in this study: observations, questionnaires, and language proficiency tests. Observations consisted of the researcher video recording (and subsequently having NESs observe and assess) intercultural dyadic conversations between a JEFL and an NES in English. The three questionnaires used in this study were administered respectively: (1) to gauge how extraverted the JEFLs thought they were, (2) to determine the extent to which the JEFLs were willing to communicate across cultures in English, and (3) to assess the levels of intercultural communicative competence (ICC) attained by each JEFL. First, the researcher administered a Ten Item Personality Inventory, i.e., the TIPI-J (developed by Oshio, Abe, & Cutrone, 2012, see Appendix D). Since it was thought that Extraversion/Introversion would have the greatest impact on listenership behavior (and how much individuals speak), this dimension was the focus of this analysis. Second, in administering the WTC questionnaire, the researcher used McCroskey's (1992) well-known WTC scale (see Appendix B). McCroskey's WTC scale is a 20-item, probability-estimate scale. Eight of the items are fillers and 12 are scored as part of the scale. Considering the focus of this study, the sub-scores corresponding to interpersonal communication were used in the assessment of performances. The third type of questionnaire used in this study included a modified version of Hecht's (1978) widely used Interpersonal Communication *Satisfaction* Inventory. Although this inventory may at first appear dated, it is still widely used in linguistic research pertaining to both listening behavior and ICC due to its high degree of reliability and validity when used to measure interactional satisfaction in actual and recalled conversations (Harrington, 1995). Consistent with the expectancy principle in Spitzberg's (2000) model of ICC, Hecht (1978) proposed that communication satisfaction depends on the fulfilment of expectations. With this in mind, the researcher administered this questionnaire to three members of the target community (i.e., proficient users of ELF, which in this study was limited to NESs). The role of

this group was to watch each of the 21 video recorded conversations and subsequently provide impressions of the JEFLs' conversational and listenership behavior by filling out the questionnaire. Lastly, in order to measure overall English proficiency, a paper-based version of the Test of English as a Foreign Language (TOEFL) was administered by the university administration. This test was used because scores were readily available to the researcher, as the JEFLs' university administration uses it monitor students' English progress over time.

Data Collection Procedures

Three methods of data collection were used in this study: observations, questionnaires, and interviews. The observation phase involved the videotaping of intercultural dyadic conversations between JEFLs and NESs. These conversations took place in a private office at the researcher's work place. The video recording equipment used was a Sony digital video camera, which was set up unobtrusively in the corner of the room on a tripod. While the conversation was being video recorded, only the participants were present in the room. Although conversational prompts were provided by the researcher to help stimulate conversation initially, participants were encouraged to talk about anything they liked. Following the methods used in the researcher's earlier studies (Cutrone, 2005, 2014), conversations were video recorded for a period of thirty minutes, of which only the middle three minutes of each conversation were included as data to be transcribed. Moreover, the WTC and Personality questionnaires were given to JEFLs directly prior to their participation in the intercultural conversation. Finally, once all intercultural conversations were completed, the researcher sent digital video copies of the conversations, with corresponding conversational assessment questionnaires, to the three NES assessors in this study. The NESs were instructed to watch each three-minute conversation and to provide their impression as to the adequacy of each JEFL's conversational and listenership behavior by completing the corresponding questionnaire. Further, questionnaires were administered before the intercultural conversations took place, whereas interviews with the JEFLs were conducted directly upon completion of their intercultural conversation. Interviews involved the participants watching a recording of their just completed conversation and answering questions posed by the researcher. The interviews were semi-structured in that the interviewer had a general plan for the interviews but did not use a predetermined set of questions, as some questions were guided by the circumstances in the videotaped conversations and the responses of the interviewee.

Data Analysis

From a macro perspective, the data derived from the observations and questionnaires will be examined to inform judgements within the four assessment categories of listenership behavior described above: approximating listenership behavior of NESs (in the observable areas of backchannel frequency, variability, discourse contexts, and simultaneous speech), conversational involvement (via WTC scores and the number of words and questions uttered), conversational micro-skills (i.e., examining participants reactions in situations of non-understanding), and ICC (NES observer perceptions of the participants based on their conversational performances). This involves both quantitative and qualitative analyses. Concerning the intercultural conversations, the first phase of analyzing data involved a thorough examination of the transcribed conversations to ascertain whether any patterns and/or relationships were evident. Desiring to highlight some of the features associated with different levels of performance concerning listenership behavior, the researcher examines and compares the performances of individual participants across the sub-categories of listenership.

<center>Results</center>

Observations: Approximating the Listenership Behavior of
NESs *Frequency*

As stated above, JEFLs who provide fewer backchannels per interlocutor word are thought to be more in accord with NESs' listenership behavior. The JEFLs to backchannel an average of

less than every 20 interlocutor words were Hanako, Mayumi, Mana, Sae, Keiko, and Kenichi. Although less frequent backchannelling often correlates to more primary speaker words, that was not always the case in this study. While Hanako, Mayumi, Mana, and Sae were among the most talkative JEFLs in this study, producing 97, 113, 165, and 250 words respectively, Keiko and Kenichi uttered only 12 and 47 words, respectively, in their conversations. On the other end of the spectrum, the JEFLS who sent backchannels the most frequently were Runa, Sakura, and Yuka (providing a backchannel every 7.78, 7.2, and 5.88 words respectively). Predictably, Runa and Sakura were among the least talkative (uttering 30 and 4 words respectively); however, Yuka, who produced backchannels the most frequently in this study, managed to utter 92 words as a primary speaker (which was 15 words more than the JEFL average of 77).

Variability

To determine the extent that JEFLs were using diverse and varied responses, the researcher examined the ratio between minimal and extended responses that each participant provided. As no JEFL in this study employed more than one extended response (and only four JEFLs produced one extended backchannel), the results pertaining to this category seem largely negligible. For instance, it is difficult to say whether Yuka, Momo, and Rika's ratios of minimal to extended backchannels of 25:0, 24:1, and 21:1 respectively are better than Hanako (5:0), Mayumi (4:0), and Yoshimi's (2:0) ratios. These results suggest that students, overall, were quite weak in this area. Some of the JEFLs who sent backchannels most frequently were, noticeably, the only ones to produce extended backchannels. Using the Pearson Product Moment Correlation Coefficient test to measure the relationship between variables (see Appendix E), a significant positive correlation was found between the JEFLs' Extraversion scores and the number of extended backchannels they uttered in the conversations $(p<.03)$. That is, the higher a participant's Extraversion scores were, the more they were likely to produce extended backchannels. In addition, a significant negative correlation was found between the JEFLs' WTC scores and the number of minimal backchannels they produced in the conversations $(p<.05)$. Hence, the higher the JEFLs' WTC scores were, the fewer minimal backchannels they sent.

Discourse Contexts

For this category, the researcher examined how often JEFLs produced backchannels at clause final boundaries in their interlocutors' speech. Overall, most of the JEFLs did so between 25 and 50 % of the time such opportunities presented themselves. A few exceptions were Yukari, Yohei, and Taro who produced backchannels at a much higher clip of 71, 63, and 58 % in this discourse context. This result seems to reflect the fact that Yukari, Yohei, and Taro sent backchannels more frequently overall (i.e., sent a backchannel every 5.88, 10, and 10.61 interlocutor words respectively). In other categories such as TOEFL scores and number of words spoken, these JEFLs were near the average. On the other end of the spectrum, one JEFL, Miki, did not provide a backchannel in any of the 13 opportunities she had in this discourse context. However, it should also be noted that Miki did not produce any backchannels during this study. This performance was indicative of Miki's below average TOEFL score (447) and consistent with how she performed in other areas, i.e., Miki was able to produce only 44 words, most of which were not initiated by her but rather coaxed out by her interlocutor's questions.

Simultaneous Speech Backchannels (SSBs)

The average number of SSBs for the JEFL group was 2.2, with SSBs ranging from 0 to 6 among participants. A few JEFLs such as Nami and Haruna uttered 6 and 5 SSBs respectively. A closer look at their profiles demonstrates some key differences between them. For instance, Nami's TOEFL score (560) was among the highest in the group (and 65 points higher than the average score), whereas Haruna's TOEFL score (477) was one of the lower ones (18 points below average). Further, Nami produced roughly twice as many words and backchannels as Haruna (i.e., 91 to 40 words and 20 to 10 backchannels). In contrast, a few JEFLs such as Mana and Miki did not produce any SSBs, but this appeared to be attributed to the fact that they sent less backchannels overall, as Mana only produced 2 backchannels overall and Miki did not produce

any as mentioned above. Unlike Miki, Mana's performances in other areas varied, as Mana's TOEFL score of 487 was only slightly below average, and she managed to produce 165 words in her conversation (which were 89 more words than the average for the group).

Conversational Involvement and Willingness to Communicate
WTC

Taro, Nami, and Hanako's WTC scores of 80, 70, and 70 were noticeably higher than the average (42), whereas Nao, Reiko, and Sachi's scores of 3, 0, and 0 were significantly lower. This was reflected to some degree in their word output and Extraversions scores. Regarding word output, Taro, Nami, and Hanako were all well above average (76), uttering 89, 91, and 97 words respectively. Conversely, Nao, Reiko, and Sachi were all well below average, uttering 60, 30, and 4 words respectively. On the Extraversion scale, Taro, Nami, and Hanako were all above average (4.1), scoring 5, 4.5, and 4.5 respectively, while Nao, Reiko, and Sachi were all below average, scoring 4, 2, and 3.5 respectively. TOEFL scores among the participants with high WTC scores varied as Nami and Hanko's scores of 560 and 523 were above average (495) but Taro's score of 463 was well below. TOEFL scores among the participants with lower WTC scores varied as well, as Sachi and Nao's scores of 497 and 483 were near the average, whereas Reiko's score of 443 was well below.

Word Output

Since the number of words a conversational participant utters in a primary speaker role is directly linked with how frequently they are in a listener role and, thus, able to send backchannels, part of the analysis for this sub-category of *word output* has already been presented above in the *frequency of backchannels* sub-category (i.e., in tandem with word output). Nonetheless, to add to what has been presented thus far, the results of the data analysis seem to suggest a strong connection between English language proficiency and how much the JEFLs spoke in the conversations. That is, a significant positive correlation was found between the JEFLs' TOEFL scores and the number of words they uttered in the conversations ($p<.023$). This was especially evident in the case of Sae, who scored 593 on the TOEFL (which was more than 30 points higher than anyone else); Sae uttered 250 words, which was by far the most of any JEFL (85 more words than anyone else).

Number of Questions

This is another area in which JEFLs were weak across the board. In 21 conversations, only 4 questions were posed, and each question was uttered by a different JEFL. Thus, the results here are largely negligible. The implications of these findings will be discussed below.

Conversational Repair Ability
Situations of Non-understanding

As stated above, some unconventional uses of the JEFLs' backchannels were brought to light, namely, their tendency to employ continuer, understanding, agreement, and/or support and empathy type backchannels in situations when they did not understand what their interlocutor was saying. Using retrospective interview techniques and analysis, the researcher examined the number of non-understanding situations experienced by the JEFLs in the conversations and how they reacted to them. First, it is necessary to point out that more proficient JEFLs were less likely to encounter situations of non-understanding, as a significant negative correlation was found between the JEFLs' TOEFL scores and the number of non-understanding situations they experienced in the conversations ($p<.017$). Like the *variability* sub-category, this appears to be an area where students did not perform well across the board. That is, in the 39 times they experienced situations of non-understanding, JEFLs produced unconventional backchannels 38 times (97%). More proficient JEFLs tended to produce less unconventional backchannels, as a significant negative correlation was found between the JEFLs' TOEFL scores and the number of unconventional backchannels they uttered in the conversations ($p<.017$). Additionally, a significant negative correlation was found between the JEFLs' Extraversion scores and the number of unconventional backchannels they uttered in the conversations ($p<.018$). Ergo, the

higher participants' Extraversion scores were, the fewer unconventional backchannels they produced.

Repair Ability

The ability to use conversational repair strategies is directly linked to the aforementioned discussion of the JEFLs' tendency to employ unconventional backchannels when they did not understand what their interlocutor was saying. Rather than feign understanding or agreement in these situations, JEFLs may choose to convey their true feelings in one of two ways: by providing a minimal backchannel expression with a rising intonation, or by employing a longer expression or phrase as a conversational repair strategy. No JEFL in this study was able to employ a conversational repair strategy (minimal or full-turn repair) in situations of non-understanding. In fact, the one time that a student did not produce a nonconventional backchannel in this situation, she just simply remained silent and did not do anything at all. It was only revealed to the researcher post hoc in the playback interviews that she did not understand the gist of what her interlocutor was saying.

Intercultural Communicate Competence (ICC)

As discussed above, a fundamental requirement of ICC is for a foreign language speaker to be seen as a competent conversationalist by members of the target culture. Accordingly, a small group of NES observers were able to watch each of the 21 video recorded conversations and subsequently provide impressions of the JEFLs' conversational and listenership behavior by filling out the revised version of Hecht's (1978) Interpersonal Communication *Satisfaction* Inventory questionnaire (see Appendix C). Subsequently, the researcher calculated (and averaged) the overall scores (as associated with positive perceptions) of NES observers on this questionnaire. The three JEFLs to score the highest were Sae, Rika, and Nao (with scores of 90, 89, and 86). Sae and Rika were also among the most proficient in TOEFL (595 and 540 respectively), while Nao's score of 483 was slightly below average. Their word output fluctuated as well, with Sae uttering 250 words, but Rika and Nao producing only 75 and 60 respectively. In the same way, concerning backchannel frequency, Sae provided a backchannel for every 23 of her interlocutor's words, while Nao and Rika did for every 11 and 10 respectively. On the other end of the spectrum, Miki and Aya's scores of 65 and 61 were well below the average of 78. Both produced a minimal number of words, as Aya uttered 60 and Miki uttered 44. Interestingly, there was notable disparity in their English language proficiency, as Aya's score of 533 was well above average on the TOEFL (495), while Miki's score of 447 was well below.

Discussion

From the findings analyzed above, some general observations can be made. Addressing RQ 1, the JEFLs who performed well in one area of the four assessment criteria of listenership behavior did not necessarily perform well in other areas; however, many of the JEFLs who performed poorly in one area of the four assessment criteria tended to perform in the same way in other areas. Although the significance of individual differences has been well documented, it is also necessary to point out that each sub-skill may have its own unique interface with individual learners, and, thus, the mastery of one, or even many, of the sub-skills involved in listenership behavior does not guarantee success in other areas of this highly complex and multifaceted skill-set. Within single sub-categories, the researcher always discovered exceptions bucking the general trends, which make it impossible to draw any concrete and comprehensive conclusions towards definitive profiles of listenership behavior.

Regarding RQ 2, it is not possible to say that English proficiency predicts success in EFL listenership behavior; however, there appears to be a tenuous link between proficiency and performance in several of the sub-skills of listenership behavior. Most notably, more proficient students were generally able to adopt a primary speaker role more often and had fewer situations of non-understanding, which, in turn, meant producing less unconventional backchannels. Nonetheless, one of the peripheral findings of this study is that students in varying levels of proficiency would benefit from instruction on listenership behavior. Across the board, students collectively were not able to able to produce extended listener responses, pose

adequate questions, expand upon initial utterances, and use conversational repair strategies and management techniques.

All three of the NES observers commented on the inability of the JEFL to drive the conversation forward. The following excerpt provides a typical example of what the NES observers are referring to:

G
6. John; they are very very near same?// what's the main difference//=
　　　　　　　　　　　　　　　　　　　　　　　　　^

=umm

　　　　　　　　　　　　　　　　　　　　　　G　　　　　　　　　　　　　^ ^ ^
^ ^ ^
7. John; i mean price was the same?//　　　price un(.) cost?//　same? // near same?//
　　　　　　　　　　　　　　　　　　　　　　^　　　　　　^ ^ ^　　　　　^ ^ ^
　　　　　　　　　　　　　　　　　　　price　　　　　　　　　　　　　　　　　uh
same

　　　　　　　　　　　　G
8. Taro; (..)umm(.)program (.)umm program difference//=　　　　(..)umm

　　"
　　　　　　　　　　　　　　　　　　　　　　　　　　　=uh　　　　　　huh
okay

---

9. John; do you have some questions for me?//=　　　　　　not really?//
　　　　　　　　　　　　　　　　　　　　　　　　　　　　> > >
　　　　　　　　　　　　　　　　　　　　　　　=umm(.) umm

　　　　　　　　　　　　　　G
10. John; so did you get the/ get a homestay // did you write letter your homestay

---

11. John; family?//　did they reply?//
　　　　　　　^^　　　　　　　　　　> >
　　　　　yes

In the excerpt above, it is clear that John is driving the conversation forward, while Taro is merely reacting to what John is saying. In fact, John poses 11 questions, while Taro does not ask any. As shown on line 9, in an attempt to encourage Taro to take some conversational responsibility, John even resorts to asking Taro if he has any return questions, which is probably not considered a natural thing to do in a casual ELF conversation. Taro's performance may owe a great deal to his somewhat lower proficiency in English (i.e., 463 TOEFL); however, it was clear that Taro, like many of the JEFLs, was extremely hesitant and did not make a great effort to drive the conversation forward. This type of behavior was consistent with what Cutrone (2014) and Sato (2008) reported of the JEFL participants in their intercultural analyses. Although the NES interlocutors in these studies generally expected and accepted that they would have to carry the conversation in their NS-NNS exchanges, they also admitted that this onus detracted from their conversational satisfaction and enjoyment.

Conclusion

In short, what we have confirmed here is that there is a great deal of individual variation where listenership is concerned, and output is often influenced to varying degrees by, among other things, the specific contexts of each conversation, the personality and demeanour of the participants, and the chemistry between the participants in the dyadic conversations, as well as seemingly peripheral variables such as the amount of sleep the participants had the night before and the mood of the participants at the time of the conversations, etc. With this in mind, this

study has helped identify some areas of listenership that EFL teachers and trainers in Japan can target for instruction in their classrooms. Specifically, the researcher advocates the teaching of conversational management techniques, which involves the appropriate usage of discourse markers and listener responses, evaluative comments, return questions, follow-up questions, new topic initiation, expansion techniques, the ability to ensure comprehension on the part of the listener, and the ability to initiate repair when there is a potential breakdown.

Since the so-called *rules of conversation* are quite different in Japanese than they are English, the writer suggests a three-pronged teaching approach that first begins by raising awareness of communication styles across cultures. To raise students' consciousness of a particular feature of communication, teachers can expose their students to sample (chosen or created) conversations that demonstrate the behavior to be analyzed (e.g. how much participants spoke, and/or how many questions they posed). Upon observing the sample conversation, students should be guided through a discussion of what they observed and how the behaviors in question might be perceived across cultures (i.e., in the above-mentioned example, students might be able to reach the conclusion that low speaker incipiency and failure to ask questions to drive the conversation forward, will, at times, negatively affect IC). This process of deconstruction helps learners understand they might need to adjust some of their own behavior in order to better adapt to ELF norms.

Thus, the second phase of instruction provides students with a framework for initiating some changes in their conversational behavior. This would involve the teacher explicitly demonstrating to learners how they might be able to improve in a particular area. For instance, concerning the example given above, teachers would help students develop turn-taking and expansion techniques, as well as strategies that help them pose adequate return and follow-up questions in conversations. Lastly, in the third phase of instruction the teacher provides students with practice opportunities and feedback. This can be done by having students participate in role-plays or conversations in which they focus on applying the new conversational techniques they learned in the previous phase. The teacher and/or other students should observe the conversations (live, or if possible, via video playback) and offer constructive feedback.

In conclusion, this study contributes to our understanding of the listenership behavior of Japanese EFL learners. While there appears to be an association between proficiency level and performances relative to listenership behavior on some levels, participants also exhibited a great deal of individual variation in their performances. Follow-up studies could be designed to shed even more light on individual differences by increasing sample size and incorporating a more balanced ratio of female to male participants. Regarding the latter, the role of gender differences in listenership behavior could be examined by utilizing mixed-sex conversations. Moreover, future research in this vein would do well to investigate diverse groups of EFL learners and examine how listenership behavior is affected by other factors such as larger group dynamics, varying conversational registers, interlocutor familiarity, and the topic of the conversation.

References

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

Clancy, P. M., Thompson, S. A., Suzuki, R., & Hongyin, T. (1996). The conversational use of reactive tokens in English, Japanese, and Mandarin. *Journal of Pragmatics, 26*(1), 355-387.

Crawford, W. (2003). Back-channel transfer in cross-cultural second language conversation: Do native Japanese transfer more than native English speakers? *Journal of Kitakyushu National College of Technology, 36*(1), 155-163.

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

Cutrone, P. (2010). The backchannel norms of native English speakers: A target for Japanese L2 English learners. *University of Reading Language Studies Working Papers, 2*(1),

28-37.

Cutrone, P. (2011). *Towards the learning of listener responses in the Japanese EFL context.* PhD Dissertation, University of Reading, Reading.

Cutrone, P. (2014). A cross-cultural examination of the backchannel behavior of Japanese and Americans: Considerations for Japanese EFL learners. *Intercultural Pragmatics, 11*(1), 83-120.

Dewaele, J. M., & Furnham, A. (2000). Personality and speech production: A pilot study of second language learners. *Personality and Individual Differences, 28*(1), 355-365.

Duncan, S., & Donald, F. (1977). *Face to face interaction: Research, methods, and theory.* Hillsdale: Erlabaum.

Edelsky, C. (1981). Who's got the floor? *Language in Society, 10*(1), 383-421.

Eysenck, H. (1992). *A hundred years of personality research, from Heymans to modern times.* Houten: Bohn, Staeu Van Loghum.

Fujimoto, D. T. (2007). Listener responses in interaction: A case for abandoning the term, backchannel. *Journal of Osaka Jogakuin College, 37*(1), 35-54.

Harrington, N. G. (1995). The effects of college students' alcohol resistance strategies. *Health Communication, 7*(4), 371-391.

Hayashi, R. (1988). Simultaneous talk - from the perspective of floor management of English and Japanese speakers. *World Englishes, 7*(3), 269-288.

Hecht, M. L. (1978). The conceptualization and measurement of interpersonal communication satisfaction. *Human Communication Research, 4*(3), 253-264.

Heffernan, Neil & John Jones. 2005. *Top-notch students: Study skills for Japanese university students.* Tokyo: MacMillan Language House.

Ike, S. (2010). Backchannel: A feature of Japanese English. In A. M. Stoke (ed.), *The Japan Association for Language Teachers (JALT) 2009 conference proceedings* (pp. 205-215). Tokyo: JALT.

Kendon, A. (1977). Some functions of gaze-direction in two-person conversation. In A. Kendon (ed.), *Studies in the behavior of social interaction* pp. 13-51). Bloomington: Indiana University Press.

Lebra, T. S. (1976). *Japanese patterns of behavior.* Honolulu: The University of Hawaii Press.

Markel, N. (1975). Coverbal behavior associated with conversation turns. In A. Kendon, R. M. Harris & M. R. Key (eds.), *Organization of behavior in face-to-face interaction* (pp. 189-197). The Hague: Mouton.

Maynard, S. K. (1986). On back-channel behavior in Japanese and English casual conversation. *Linguistics, 24*(6), 1079-1108.

Maynard, S. K. (1990). Conversation management in contrast: Listener responses in Japanese and American English. *Journal of Pragmatics, 14*(1), 397-412.

Maynard, S. K. (1997). Analyzing interactional management in native/non-native English conversation: A case of listener response. *International Review of Applied Linguistics in Language Teaching (IRAL), 35*(1), 37-60.

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

Mizutani, N. (1982). The listener's responses in Japanese conversation. *Sociolinguistics Newsletter, 13*(1): 33-38.

O'Keeffe, A., & Adolphs, S. (2008). Response tokens in British and Irish discourse: Corpus, context and variational pragmatics. In A. Barron & K. Schneider (Eds.), *Variational pragmatics* (pp. 69-98). Amsterdam: John Benjamins.

Oshio, A., Abe, S., & Cutrone, P. (2012). Development, reliability, and validity of the Japanese version of Ten Item Personality Inventory (TIPI-J). *Japanese Journal of Personality, 21*(1), 40-52.

Sacks, H. E., Schegloff, E. A., & Jefferson, G. (1974). A simple systematics in the organization of turn-taking for conversation. *Language, 50*(1), 696-735.

Sato, Y. (2008). *Oral communication problems and strategies of Japanese university EFL learners.* PhD Dissertation, University of Reading, Reading.

Schegloff, E. A. 1982. Discourse as an interactional achievement: Some uses of "UH-HUH"

and other things that come between sentences. In D. Tannen (ed.), *Georgetown University roundtable on language and linguistics, analyzing discourse: Text and talk* (pp. 71-93). Washington: Georgetown University Press.

Spitzberg, B. H. (2000). A model of intercultural communication competence. In L. Samovar & R. Porter (Eds.), *Intercultural communication: A reader, 2ⁿ edition* (pp. 7-24). Belmont, CA: Wadsworth Publishing.

Stubbe, M. (1998). Are you listening? Cultural influences on the use of supportive verbal feedback in conversation. *Journal of Pragmatics, 29*(3), 257-289.

Thonus, T. (2007). Listener responses as a pragmatic source for learners of English. *CATESOL Journal, 19*(1), 132-145.

Tottie, G. (1991). Conversational style in British and American English: The case of backchannels. In K. Aijmer and B. Altenberg (eds.), *English corpus linguistics: Studies in honour of Jan Svartvik* (pp. 254-271). London: Longman.

Ward, N., & Tsukahara, W. (2000). Prosodic features which cue back-channel responses in English and Japanese. *Journal of Pragmatics, 32*(8), 1177-1207.

White, S. (1989). Backchannels across cultures: A study of Americans and Japanese. *Language in Society, 18*(1), 59-76.

About the Author

Pino Cutrone has been teaching in Japan for twenty years and is currently an Associate Professor at Nagasaki University. He received his PhD in Applied Linguistics from the University of Reading and has published widely in his field. His research interests include intercultural pragmatics, sociolinguistics, and EFL pedagogy.

Appendices

Appendix A
Transcription Conventions

- Listener responses are shown in italics below the primary speaker's talk at the point they occurred in the talk.

- To protect the identity of the participants, pseudonyms are used in the speaker labels on the left side of each transcribed line.

- To not confuse readers with the colons that are used for a different purpose described below, the speaker labels will be followed by a semi colon.

- To further preserve anonymity, pseudographs (i.e., notations in parentheses) will be used in instances where participants' private information such as name, address and/or telephone number has been uttered in the conversation.

- Numbers in parentheses indicate elapsed time in hundredths of seconds of pauses occurring in the conversations. Parentheses with a dot (.) indicates a micropause and/or hesitation under .5 seconds. Pauses are timed using transcription software in this study (Praat Version 5.0.18).

- The equal sign "=" indicates latching - i.e., no interval between the end of a prior piece of talk and the start of a next piece of talk.

- The beginnings of simultaneous speech utterances are marked by placing a left bracket at each of the points of overlap, and placing the overlapping talk directly beneath the talk it overlaps.

- Right-hand brackets indicate the point at which two simultaneous utterances end.

Metatranscription is shown as follows:

- Empty parentheses ( ) indicates that part of the transcription which is unintelligible.
- Words between parentheses indicate the transcribers' conjecture at the words or utterances

in the conversation that they are not completely certain of.

- Words between double parentheses may indicate comments and/or features of the audio materials other than actual verbalization.

- L stands for laughter.

- Other than apostrophes, which are used to show contraction between words, punctuation symbols in these transcriptions are not used as regular English punctuation markers indicating grammatical category. While other, non-regular, grammatical functions are shown by symbols such as slashes and double slashes, other punctuation symbols such as question marks and colons are used to indicate prosodic features in these transcriptions.

Nonverbal behavior is shown by the symbols indicated below.

- h stands for audible breathing. ^ stands for vertical head movement (head nod). > stands for horizontal head movement (head shake). S stands for smile. " indicates that eyebrows are raised. G indicates body or hand gestures.

- In cases where nonverbal behavior occurs concurrently with speech, symbols are placed directly above the speech it co-occurs with (instances where two types of nonverbal behavior occur simultaneously are shown by underlining them both). Nonverbal behavior that is continuous and occurs for a period longer than 2 seconds will be noted by signaling the beginning and the end of the behavior in parentheses where it occurs in the conversation. N.B. The parentheses containing the symbols below are solely used for separation purposes to make them easily identifiable in the specific examples below. Parentheses will not be used in this manner in the transcriptions as they have other specific functions, which have been outlined above.

- A slash (/) marks the grammatical completion point of an internal clausal boundary (i.e., a clause which is continuative).

- Two slashes side by side (//) mark the grammatical completion point of a final clause boundary (i.e., a clause which terminative). N.B. A final clause boundary is one that makes complete sense (i.e., fully meaningful) and could end the utterance there. In contrast, an internal clause is one in which the meaning is not complete, and there is a requirement for the utterance to go on in order for the meaning to be complete.

- A question mark (?) at the end of a word and/or utterance indicates a clear rising vocal pitch or intonation (i.e., one that is clearly heard, and is shown to rise by at least 600 Hz using Praat software).

- An inverted question mark (¿) at the end of a word and/or utterance indicates a clear falling pitch or intonation (i.e., one that is clearly heard and is shown to fall by at least 600 Hz using Praat software).

- A colon (:) as in the word "ye:s" indicates the stretching of the sound it follows (i.e., only marked in cases where the stretching was extended greater than .5 seconds).

- A hyphen at the end of an uncompleted word indicates the disfluency of a truncated word. For instance, if the word "bird" were truncated, it may be transcribed as "bir-".

- A part of a word and/or phrase containing CAPITAL letters indicates that it has been said with increased volume and/or more emphatically than the rest of the phrase (i.e., only marked when the highest point of the stressed part of speech was greater than 10 decibels the lowest part of the surrounding parts of speech).

- The underscore sign (_) indicates that the talk it precedes is low in volume.

- (~) indicates that the talk which follows is consistent with the person's regular voice and tone. This symbol is to be used after low volume talk to indicate the point in which the volume rises back to normal. When a pause occurs after the low volume talk and the talk that follows returns to normal, this symbol will not be shown.

Transcription of Conversation

>
1. Taro;   _mmm ~ no i've, (.) i've never (.75)

                    G   G
2. Gary;   oh now i remember his name// kevin smith// (.) my mistake//(.)

3. Taro;   [kev]in smith i i heard his name// (.) his (.)

       [( )]

              ^
4. Gary;   he's a very good director// and [writer]// (.)   (.)

                                      [uhum]   *oh really*

5. Gary;   and he appears in (.54) uu many of his movies/ as   [silent] bob// (.)

                                                     [(uhum)]

6. Gary;   who (.75) whose quirk. is he he doesn't speak// (.54)

                                 ((G begins))
7. Gary;   [except] maybe once// in the whole movie// (.)

                     ^
       _ *[(uhum)]*

                                   ((G   ends))
^
8. Gary;   [otherwise, he's, pan]tomiming// (.97)   = m (.)

       [(ooohh yeah)   ]   *oh really? =*

                              G
9. Taro;   how how long is (.51) is the movie// (.)

10. Gary;   (uu it) (.) depends// let's see// (.67) clerks (.) they had a few

                                      G
11. Gary;   appearances// (.67)ahh but (.) silent bob never spoke in that one// (.)

12. Gary; i, remember// right// =   (.)but he did speak uu in their next movie//

                   = mhm

13. Gary;   which [was (uu)] mall rats// (.)   (.61)

```
                         ^ ^
                     [mhm yeah]                           _mall ra-
```

```
                     ((G begins))
14. Gary;     uu (.99) (wha-) which is a (1.02) ahh let's see// mall rats was a: (.59)
```

---

15. Gary;   term/ used for (.) aaa (.80) u people/ who hang out in the mall//

```
                     ((G ends))
16. Gary;     but (.)[ never] actually (.) aa [ buy ] anything// (.)
                           ^                              ^
                     [(mhm)]                        [yeah]
```

```
17. Gary;     [it was] just sort of a place for them/ to (.) [hang out]// and (.88)

                   ^                                  ^
               [(mm)]                              [ (mmm) ]
```

```
18. Gary;     uuu (.) it was sort of a derogatory term// used for them//
                                                              ^
```

```
19. Gary;     [by]                 (.)some of the shop keepers. =
                       ^                                    ^
                 [ (mhm) ]                            = yeah
```

---

20. Gary;   (.67) uu in particular ben affleck. (.)

---

21. Gary;   i forget (.) what the name of his character was// but he was (.54)

```
22. Gary;      uuu the bad guy// for the most part// =                 = uu in that movie//

                                               = oh really_Lh =
```

```
                     ((G begins))
23. Gary;     uu (3.37) uuu bob uu (.) u silent bob (.)       = aaa (1.39) _ Lh (.59)
                                                                       ^
                                                                 uhum =
```

```
                                                              ((G ends))
24. Gary;     uu Lh it appears. that he used the force// aa =            (.)

                                                                       = force
```

```
                                           G
25. Gary;     did you ever see star [ wars ]//   (.)            (1.02) aa (.)

                                          G              "
                     [power]?      (mm)?                  AAaa =
```

```
                     ((G begins))
26. Gary;     star wars where they (.70) aaa (.)
```

---

27. Taro;　　yeah star wars　i (.)
　　　　　　　　　　　　　　　 ∧

　　　　　　　　　　　　　　　 ((G ends))
28. Gary;　　which which of the star wars movies have you seen// (.)

　　　　　　　　　　　　　　　　S
29. Taro;　　aaa (1.29) _ what is that// uu maybe (episodes) th- (.) (three) // i (.)

---

30. Taro;　　maybe i (.) three// (　　　　) (2.25)

　　　　　　　　　　　　　　　 ((G begins))
31. Gary;　　you know when a jedi, uu (.) reaches their arm out// and then something

---

　　　　　　　　　　　　　　　 ((G ends))
32. Gary;　(.)flies into it// (.) [and then] they catch it// (.75)

　　　　　　　　　　　　 ∧

　　　　　　　　　　　　 [　uhum　]

---

33. Gary;　　but =　　　　　　　　　　 (.51) but let's say they dropped your [life (　　　)]/
　　　　　　　　　　　　 S
　　　　　　= AAa (1.37) aa　　　　　　　　　　　　　　 *[(　Lh　)]*

　　　　　　　　　　　　　　　 ((G begins))
34. Gary;　　their [life saver]//　　　　= [and then] (.) they reach for it// =
　　　　　　[(　Lh　)]/ h (　) = [　yeah　]

---

35. Gary;　　　　　　 [and] then the life saver flies to them// (.51)

　　　　　= h Lh [Lh]

　　　　　　　　　　　 ((G ends))
36. Gary;　　[(　　　　　)] seeing something/ like that// (1.26)

　　　　　[　AAaa　]

---

37. Taro;　　maybe i (.) i, (.) ss (.67) saw that// (.) maybe (.67)

---

38. Gary;　　uuu (.) star wars is¿ (.) uuu (.) a big part of popular culture//
　　　　　　　　　　　　　　　　　　　　　　 ∧　　　 ∧

　　　　　　　　　　　　　　　 ((G begins))
39. Gary;　　in a[merica]//　　　　= aaa (2.06) and u kevin smith (the) (.54)
　　　　　 ∧
　　　　　[uhum]　(.) yeah =

　　　　　　　　　　　　　　　 ((G ends))
40. Gary;　　[　he] he's, no exception// he, uu (.)greatly enjoys the: star [(wa-)]
　　　　　　　　　　　　　　　　　　　　　　　　　 ∧∧
　　　　　[(　)]　　　　　　　　　　　　　　　　 *[uhum]*

---

41. Gary;　　star wars mythology// (.62) [uu and incor]porates it//
　　　　　　　　　　　　　　　　　 ∿

[oh yeah]

42. Gary;   a:t? least a few times// in aa (.) his movies// (.)   (.83) aa in (.72)
                                                                              uhum

((G begins))
43. Gary;   that uu in mall rats _u (.) _he would ~ he needed to get uu (.91)

((G continues))
44. Gary;   aaaa a video// (   ) (.) a videotape that was ...

Appendix B
WTC Questionnaire

DIRECTIONS: Below are twenty situations in which a person might choose to communicate or not to communicate in English. Presume that the person in each situation does not speak Japanese but can speak English. Also, presume you have *completely free choice.* Indicate the percentage of times you would choose *to communicate* in each type of situation. Indicate in the space at the left what percent of the time you would choose to communicate.

0 = never, 100 = always

   _____   1. *Talk with a service station attendant.
   _____   2. *Talk with a physician.
   _____   3. Present a talk to a group of strangers.
   _____   4. Talk with an acquaintance while standing in line.
   _____   5. *Talk with a salesperson in a store.
   _____   6. Talk in a large meeting of friends.
   _____   7. *Talk with a police officer.
   _____   8. Talk in a small group of strangers.
   _____   9. Talk with a friend while standing in line.
   _____   10. *Talk with a waiter/waitress in a restaurant.
   _____   11. Talk in a large meeting of acquaintances.
   _____   12. Talk with a stranger while standing in line.
   _____   13 *Talk with a secretary.
   _____   14. Present a talk to a group of friends.
   _____   15. Talk in a small group of acquaintances.
   _____   16. *Talk with a garbage collector.
   _____   17. Talk in a large meeting of strangers.
   _____   18. *Talk with a spouse (or girl/boy friend).
   _____   19. Talk in a small group of friends.
   _____   20. Present a talk to a group of acquaintances.

N.B. JEFLs were provided with Japanese explanations. Further, the asterisk (*) marking the filler items above, as well as the scoring table below, were not included on the questionnaires the JEFLs completed.

SCORING: The WTC permits computation of one total score and seven sub-scores. The sub-scores relate to willingness to communicate in each of four common communication contexts and with three types of audiences. To compute your scores, merely add your scores for each item and divide by the number indicated below.

Sub-score Desired	Scoring Formula
Group discussion	Add scores for items 8,15, and 19; then divide by 3.
Meetings	Add scores for items 6, 11, and 17; then divide by 3.
Interpersonal conversations	Add scores for items 4,9, and 12; then divide by 3.
Public speaking	Add scores for items 3, 14, and 20; then divide by 3.
Stranger	Add scores for items 3, 8, 12, and 17; then divide by 4.
Acquaintance	Add scores for items 4, 11, 15, and 20; then divide by 4.
Friend	Add scores for items 6, 9, 14, and 19; then divide by 4.

To compute the total WTC scores, add the sub-scores for stranger, acquaintance, and friend. Then divide by 3.

Appendix C
Conversational Satisfation Questionnaire (for NES Assessors)
Date: _____   Name: _____   **Key:**   1 = Yes

7 = No

Please score the sentences below based on how often you thought they generally occurred in the conversation. Based on the key shown above, circle the number that best corresponds to your opinion.

1. The Japanese person let his/her partner know that the partner was communicating effectively.
....1   2   3   4   5   6   7

2. The Japanese person showed his/her partner that they understood what their partner said.
....1   2   3   4   5   6   7

3. The Japanese person showed that they were listening attentively to what their partner said.
....1   2   3   4   5   6   7

4. The Japanese participant expressed a lot of interest in what their partner had to say.
....1   2   3   4   5   6   7

5. The conversation went smoothly...................................1   2   3   4   5   6   7

6. The Japanese encouraged his/her partner to continue talking.1   2   3   4   5   6   7

7. The feelings that the Japanese person expressed by means of listening feedback during the conversation seemed *authentic* (i.e., they conveyed what they were truly feeling   and not just agreeing and/or pretending to understand for the sake of harmony and/or to keep the conversation going smoothly)..1   2   3   4   5   6   7

8.   The Japanese person seemed impatient.........................1   2   3   4   5   6   7

9. The Japanese person seemed cold and unfriendly...............1   2   3   4   5   6   7

10.   The Japanese person was polite..1   2   3   4   5   6   7

11.　　The Japanese person appeared warm and friendly............1　　2　　3　　4　　5
6　　　7

12.　The Japanese person was impolite..................................1　　2　　3　　4　　5
6　　　7

　13. The Japanese person appeared interested and concerned....1　　2　　3　　4　　5
6　　　7

　14. The Japanese person interrupted their partner at times.......1　　2　　3　　4　　5
6　　　7

　15. The Japanese person seemed to want to avoid speaking......1　　2　　3　　4　　5
6　　　7

16. When the Japanese person did not understand something, they were able to clearly convey this to their　conversational partner with their listening feedback................1　　2　　3
4　　　5　　　6　　　7

17. The Japanese person's listening behavior seemed inadequate in some ways.1.　　2　　3　4
5　　6　　　7

If you answered "yes" (i.e., 1, 2 or 3) to question 17, please explain how and/or why you think their listening behavior seemed inadequate.
_____
_____
_____
_____
_____

18. Any other comments and/or observations regarding the Japanese participant's behavior in the conversation.
_____
_____
_____
_____
_____

Appendix D
Personality Questionnaire
Name（名前）:_____Date（記入日）: _____

Questionnaire　質問紙

Following the scale below, please write a number next to each statement below to indicate the degree to which you agree or disagree with that statement.

○１から 10 まてのことはがあなた自身にとのくらい当てはまるかについて，下の枠内の１から７ま ての数字のうちもっとも適切なものを括弧内に入れてくたざい。文章全体を総合的に見て，自分にとれたげ当てはまるかを評価してくたざい。

Disagree Strongly （全く違う	Disagree moderately （あまり	Disagree a little （少し違	Neither agree nor disagree （どちらでも	Agree a little （少しそう	Agree moderately （まあまあそ	Agree strongly （強くそ

と思う 1	そうだとは 思わない） 2	うと思う） 3	ない） 4	思う） 5	う思う） 6	う思う） 7

I see myself as... (私は自分自身のことを...)

1. _____ Extraverted, enthusiastic.(活発で，゛外向的だど思う)

2. _____ Critical, quarrelsome.(他人に不満をもち，もめこどを起こしやすいと思う)

3. _____ Dependable, self-disciplined.(しっかりしていて，自分に厳しいと思う)

4. _____ Anxious, easily upset.(心配性で，うろたえやすいと思う)

5. _____ Open to new experiences, complex.(新しいことか好ぎで，変わった考えをもつと思う)

6. _____ Reserved, quiet.(ひかえめて，おとなしいと思う)

7. _____ Sympathetic, warm.(人に気をつかう，やさしい人間だと思う)

8. _____ Disorganized, careless.(たらしなく，うっかりしていると思う)

9. _____ Calm, emotionally stable.(冷静て，気分か安定していると思う)

10. _____ Conventional, uncreative.(発想力に欠けた，平凡な人間だと思う)

Appendix E
Correlational Analyses

Key explaining dependent variables in order presented below: Total Words, Frequency (number of backchannels per interlocutor word), Number of Questions, MinBack (percentage of backchannels constituted by minimal backchannels), ExtBack (percentage of backchannels constituted by extended backchannels), BCs@FCBs (percentage of clause final boundaries attracting backchannels), SSBs (simultaneous speech backchannels), NONU (number of non-understanding situations), UNCONV (percentage of non-understanding situations constituted by unconventional backchannels), MinRep (percentage of non-understanding situations constituted by minimal backchannel as repair strategies), and FullRep (percentage of non-understanding situations constituted by Full-turn repair strategies).

		TOEFL	WTC	Extraversion
Total Words	Correlation Coefficient Sig. (2-tailed)	.495 .023*	.168 .465	.274 .229

BC/IL word	Correlation Coefficient Sig. (2-tailed)	-.054 .815	.101 .665	.209 .364
Questions	Correlation Coefficient Sig. (2-tailed)	.059 .798	-.050 .831	.334 .139
MinBCs	Correlation Coefficient Sig. (2-tailed)	.021 .930	-.433 .05*	.022 .924
Extended BCs	Correlation Coefficient Sig. (2-tailed)	.041 .861	-.140 .544	.475 .030*
BC@FCBs	Correlation Coefficient Sig. (2-tailed)	.203 .379	.128 .581	.258 .259
SSBCs	Correlation Coefficient Sig. (2-tailed)	.119 .606	-.133 .565	-.036 .876
NONU	Correlation Coefficient Sig. (2-tailed)	-.513 .017*	-.259 .256	-.428 .053
UNCONV	Correlation Coefficient Sig. (2-tailed)	-.490 .024*	-.271 .235	-.511 .018*-
MinRep	Correlation Coefficient Sig. (2-tailed)	N/A	N/A	N/A
FullRep	Correlation Coefficient Sig. (2-tailed)	N/A	N/A	N/A

English Language Teacher Agency in Classroom-based Empirical Studies: A Research Synthesis

Zhenjie Weng*
Jingyi Zhu
Grace J. Y. Kim
Ohio State University, USA

Abstract

This research synthesis analyzes a selection of classroom-based empirical studies on language teacher agency within ESL/EFL/bilingual contexts. Comprehensive analyses of the selected research center on several respects: theoretical frameworks, contexts, methodologies, major findings, methodological and ethical issues as well as implications. In doing this, this paper aims to present ESL/EFL/bilingual teachers' roles as legitimate and agentive actors in their contexts of work, to offer pedagogical implications for teachers and teacher educators, to better inform researchers of current literature and future research directions, and to support possible collaborations among different educational stakeholders.

Keywords: research synthesis, classroom-based empirical studies, English language teacher agency, ESL, EFL, bilingual

Introduction

A growing body of empirical studies centers on "teacher agency" in language education in recent years (e.g., Huyen Phan & Hamid, 2017; Kang, 2017; Nguyen & Bui, 2016). Teacher agency has been theorized regarding the activities teachers do in schools (Priestley, Biesta, & Robinson, 2013) as an "important dimension of teachers' professionalism" in response to curriculum or institutional changes (Biesta, Priestley, & Robinson, 2015, p. 625). Moreover, Priestley, Edwards, Priestley, and Miller (2012) state that teacher agency is "largely about repertoires for maneuver, or the possibilities for different forms of action available to teachers at particular points in time" (p. 211). Drawing upon previous scholarship, the current investigation aligns with the notion that teacher agency could be mediated within specific sociocultural contexts (Lasky, 2005), and within this framework, teachers have "the socially constituted capacity to act" on educational changes (Barker, 2008, p. 234).

Although the concept of agency has attracted attention in the education literature in recent years, teacher agency still remains understudied in the field of language education, particularly in classroom-based ESL/EFL/bilingual contexts. Among the limited research, a relatively large amount of studies have engaged in discussions of factors that affect teacher agency, such as teacher belief and perceptions (Shabir, 2017; Tang, Lee, & Chun, 2012) and social environment (Meierdirk, 2018). Little attention, however, has been paid to document teachers' enactment of change situated in language classrooms. In response to this, this study employs research synthesis as theoretical framework to offer major insights from the studies on teacher agency and identify literature gaps in the field for further research. Research synthesis, a "relatively sparse but rapidly growing literature" (Suri & Clarke, 2009, p. 397), is a "contemporary framework for reviewing" (Norris & Ortega, 2007, p. 806) that investigates and evaluates "past findings in a systematic fashion, always explicating the methodology followed in the review so as to enable replication by other reviewers" (Ortega, 2015, p. 225). To do so, research syntheses pay attention to relevant theories, critically analyze, and "attempt to identify central issues for future research" (Cooper & Hedges, 2009, p. 6). This research synthesis takes the epistemological stance of interpretivist that recognizes "the inevitable subjectivity in a synthesist's interpretive constructions" (Suri, 2013, p. 897) and attempts to identify the "plausible patterns" (Suri, 2013, p. 897) of teacher agency research across different empirical studies.

Guided by the framework, this synthesis aims to present ESL/EFL/bilingual teachers' roles as legitimate and agentive actors in their contexts of work, to offer pedagogical implications for teachers and teacher educators, to better inform researchers of current literature and future research directions, and to support possible collaborations among different educational stakeholders. In order to realize this, we attempt to answer the following research questions: (1) What is the

Email: weng.151@osu.edu; Address: 1945 N. High St.,138, Columbus OH 43210

nature—concerning theoretical frameworks, contexts, and methodologies—of language teacher agency research in ESL/EFL/bilingual contexts up to 2018? (2) What are the major findings, methodological and ethical issues, and implications discussed in the selected articles? (3) What are the future directions for research, teacher education, and professional development drawing upon the existing scholarship in teacher agency?

Literature Review on Teacher Agency in ESL/EFL/Bilingual contexts

In the context of language education, "learner agency" has been more frequently studied than "teacher agency;" however, the importance of which has been acknowledged (Biesta et al., 2015), particularly in the past two years (2017 and 2018). A cursory review of current literature reveals that language teacher agency, discussed in ESL/ EFL/ bilingual contexts, has been associated with other concepts, including beliefs (e.g., Ollerhead & Burn, 2016), autonomy (e.g., Hoang & Truong, 2016), identity (e.g., Dantas-Whitney, Clemente, & Higgins, 2012; Kayi-Aydar, 2015a), emotion (e.g., Benesch, 2018), language use in bilingual spaces (e.g., Henderson, 2017), (de)motivation (e.g., Song & Kim, 2016), and language policy (e.g., Hamid & Nguyen, 2016). Before proceeding to the discussion of differences among ESL/EFL/bilingualism on teacher agency, we first clarify that the distinction of language teaching and learning between ESL and EFL is not completely straightforward as Shin (2018) clearly states that "the often made distinction between [ESL] and

[EFL] is blurred for an increasing number of **transnational migrants** who cross and re-cross national boundaries" (p. 27, emphasis is in the original). Although this is true, research on teacher agency in ESL and EFL contexts shows different tendencies, scrupulously discussed in the next section. Related to ESL education is bilingual education, referring to learning contexts that involve two languages, oftentimes educating emergent bilinguals who are linguistically diverse (García, 2009; Palmer & Martínez, 2013; Wong, Athanases, & Banes, 2017). Commonly found program designs of bilingual education include dual language or immersion programs, in which two target languages are used in learning various content areas with the purpose of learners becoming bilingual, biliterate, and bicultural. Due to its dealing with linguistically and culturally diverse learners, bilingual contexts pose unique challenges and demands to teachers and teacher education. This synthesis focuses on the studies that are conducted in bilingual contexts and including English as one of the two languages.

In literature, studies on teacher agency in ESL contexts are much fewer than relevant studies in EFL contexts. Based upon extant literature, possible differences between EFL and ESL contexts reside in student population and constructs associated with teacher agency, including teacher roles or identities. We found studies on ESL teacher agency in contexts of Australia and the United States. For example, Ollerhead and Burns (2016) report two ESL teachers' response to policies in Australian adult ESL literacy classroom in which most students were refugees from Africa or Asia. The study explores the interaction of teacher roles, beliefs, and teaching approaches to uncover how those factors influence their exercise of agency. Besides, it is found that teachers' agency was affected by their own backgrounds and the institutional culture where the "one size fits all" policy compromised teaching efficiency (Ollerhead & Burns, 2016, p. 113). Furthermore, studies on teacher agency in ESL contexts also include teacher candidates. For example, Kayi-Aydar (2015a) investigates a teacher candidate's negotiation between her identity and her agency across time and space.

In addition to the aforementioned aspects, K-12 ESL teaching as language support for students who have limited English abilities is not part of mainstream education or academic discipline in most ESL contexts, such as the US, U.K., and Canada (Tyack & Tobin, 1994). In US contexts, for instance, pullout, push-in, and co-teaching models are mostly implemented in ESL. In a full pull-out model, a designated amount of time is given to ESL teachers or specialists to help English learners with explicit instruction each week; yet, push-in and co-teaching models require collaboration between ESL and grade-level teachers in their planning and instruction (Ovando & Combs, 2012). With this being said, ESL teachers are found marginalized, ignored, and invisible in schools. However, certain studies (e.g., Trickett et al., 2012) also demonstrate that despite the marginalized status, teachers counteracted the marginalization that they and their students experienced, became advocates of educational equality, interacted with content teachers, and built up relationship with administrators to influence class placements of their students.

In EFL contexts, teachers' exercise of their agency in the implementation of macro-level language education policies in local contexts has started to receive important consideration (e.g., Bouchard, 2016; Glasgow, 2016; Talalakina & Stukal, 2016). In many Asian countries—like China,

Japan, Malaysia, and South Korea—a global trend is that English was introduced earlier in curriculum and adopted as a medium of (higher) education (Hamid & Nguyen, 2016, pp. 26-27). That said, teachers have to prepare students for traditional examinations, easily resulting in test-driven approaches; yet at the same time, teachers need to provide students communicative resources to facilitate their participation in the globalized world, which entails Communicative Language Teaching (CLT). Hamid and Nguyen (2016) report that "it cannot be taken as a given that teachers will embrace the policy whole-heartedly and work towards policy goals. They may resist the policy in a covert manner if policy intentions do not reflect their interests, beliefs and realities" (p. 31). Under the tension, teachers' agentive action in policy implementation might be compromised by a series of factors, including teachers' communicative proficiency, pedagogical skills, their responsibility for students' performance on designated tests, institutional support, teacher professional development, teachers' interpretation of those policies, and other social pressures from parents and media.

In Ng and Boucher-Yip's (2016) edited book, several chapters address this phenomenon in EFL contexts. For example, in Hoang and Truong's (2016) chapter, in order to facilitate Vietnam's socio-economic development, the government aimed to profoundly improve English abilities of young graduates from all educational levels. Nevertheless, the pressure of restandardization which jeopardized many teachers' job security, the lack of power as well as professional and geographical constraints crippled the participant's agency to resist or mitigate the policies. In another chapter, Osman and Ahn (2016) explore a less studied context—a private university in Kazakhstan—to uncover English language teachers' response to the changes in light of new policies which aimed at the modernity and internationalization of its education. They report that the existence of local teaching teams, within which the teacher interviewees can interpret new changes to their colleagues and implement those changes, allowed the teachers' voice to be heard and exercise their agency; however, for the policies initiated outside of the teaching teams, there was less space for negotiation. The common theme across the aforementioned literature is that language policies in EFL contexts were related to the countries' socio-economic development. Those top-down policies were developed without participation of the teachers who were the agents, and factors like institutional contexts, teacher identity, power, and realities of local conditions have created a web of complex relationships for English language teachers to initiate their agency.

A recent trend of US bilingual education is a move towards additive bilingualism (Flores, 2001; García, 2009; Hopkins, 2013; Lambert, 1975; Palmer & Martínez, 2013; Ruiz, 1984; Wong et al., 2017). Additive bilingual programs are connected to language ideologies (Henderson, 2017; Palmer, 2011) that affirm "language as a resource" rather than "language as a problem" (Ruiz, 1984). Dual language programs are complex and dynamic, experiencing many tensions stemming from language hegemony in creating equitable learning opportunities to help students become bilingual, biliterate, and bicultural (Babino & Stewart, 2018; Henderson, 2017; Palmer, 2010, 2011). Bilingual teachers also face various challenges such as accountability demands and high-stakes testing in their efforts to connect with their students through culturally relevant pedagogy (Achinstein & Ogawa, 2012; Babino & Stewart, 2018; Wills & Sandholts, 2009; Wong et al., 2017) and linguistically responsive pedagogy (Lucas & Villegas, 2011).

In reviewing the spaces that teachers assert agency for bilingual children, Palmer and Martínez (2013) underscore the fact that teachers "acknowledge these constraints and honor the tremendous efforts sometimes required of teachers to provide authentic learning opportunities to bilingual children in school" (p. 270). Therefore, there is an increasing need for "critical orientation toward the challenge of educating bilingual students in the United States" (Palmer & Martínez, 2013, p. 274). Moreover, Bartolomé (2004) argues that teachers and educators who work in multilingual contexts need political and ideological clarity to "interrogate potentially harmful ideologies" (p. 98) for teaching is "not an apolitical undertaking" (p. 115).

Studies on teacher agency in bilingual contexts explore teachers as language policy makers (Henderson, 2017). Bilingual and dual language teachers achieve "varying degrees of agency" (Babino & Stewart, 2018, p. 274) based on their contextual experiences (Edwards, 2015). Dubetz and de Jong (2011) reviewed 30 research articles on bilingual teacher agency, and a common thread among their definitions of agency is a critical lens and a social justice orientation as a call to create equity for those who have been marginalized. In defining agency, they put an emphasis on "acting on behalf of others and encompass[es] individual and collective efforts to shape public policy in ways that ensure that individuals are treated equitably and have access to needed resources" (Dubetz & de Jong, 2011, p. 251). Thus, teacher agency in bilingual contexts presents complexity on

decision-making related to language policy implementation based on their language ideologies.

Methodology

Data Collection

The process of article selection was not linear but iterative and constantly refined, and the entire selection process was comprised of three major rounds. The first round started in March 2018, and then we met every month until April 2019 for data collection and analysis. In the first round of selection, we used the guiding keywords—"teacher agency," "language," and "classroom research"—and manually searched the literature in Scopus, Web of Science, Google Scholar, EBSCOhost, and ERIC. 53 articles were found relevant to the study, including journal articles and book chapters. In the second major round of selection, we undertook a focused selection procedure following the criteria:

1. "Teacher agency" and/or "agency" mentioned in "Abstract" or "Research Questions"
2. "Teacher agency" in the analysis
3. Classroom-based research
4. Empirical studies
5. Peer-reviewed journal articles
6. ESL/EFL/bilingual (bilingual when English is the target language) contexts

In the last round of selection, additional sources were identified from the lists of references in the selected articles during the initial stages of review process. We excluded several articles that were originally included because they do not emphasize "teacher agency" in classroom practices (e.g., Benesch, 2018; Dantas-Whitney et al., 2012; Hamid, Zhu, & Baldauf, 2014; Kayi-Aydar, 2015a). As a result, the article selection result was narrowed down to 32 articles. Among the 32 studies, 18 of them were published in 2018, followed by five studies published in 2017. Three studies were published respectively in 2015 and 2016. One source was published in 2009, 2010, and 2012 respectively. This indicates that research on teacher agency has only received attention recently. More detailed information about the focus of the studies and their research questions are presented in Appendix A.

Data Analysis

Data collection and analysis occurred simultaneously. When selecting and filtering articles, we also analyzed different components of the selected articles. Once we completed the collection of the 32 articles, we started intensive and systematic data analysis, using both top-down and bottom-up coding systems. According to the research questions we broached, we listed different categories, including theoretical frameworks, contexts, methodologies, methodological challenges, ethical issues, major findings, and implications (Appendix B). This coding system was generated through top-down process to focus on certain aspects of each article. During the process of sorting out different components of each article, we further added and revised (sub)categories that emerged after we read each article. For example, originally we only listed context regarding national settings. After reading each article, we became more specific about the context to not only include national settings but also K-12 as well as rural/urban settings. Upon completion, we arranged several meetings to collaboratively analyze the data and discuss findings. During these meetings, we talked through what we had found and commented for further revision.

Findings

In this section, we mainly focus on answering the first two research questions, namely, (1) What is the nature—concerning theoretical frameworks, contexts, and methodologies—of language teacher agency research in ESL/EFL/bilingual contexts up to 2018? and (2) What are the major findings, methodological and ethical issues, and implications discussed in the selected articles?

RQ1: What is the nature—concerning theoretical frameworks, contexts, and methodologies—of language teacher agency research in ESL/EFL/bilingual contexts up to 2018?

Theoretical Frameworks

A variety of theoretical frameworks and concepts have been applied to the study of teacher agency. Two essential themes have emerged among the 32 studies. First is that in most studies, the concept

of agency is employed to conceptualize teacher agency. For example, in Glas' (2016) study, literature on agency is reviewed first and then followed by the discussion on teacher agency. According to Glas (2016), agency, drawing upon sociocultural theory, is "mediated through psychological (and material) tools that were acquired culturally, through human interaction" and in relation to teacher agency, "the mediational means refer to their repertoire of motivational strategies, their teaching materials, or other tools that help them engage their students" (p. 444). Ilieva and Ravindran (2018) point out that agency is enabled only through the complex interplay among various affordances and constraints emergent in daily practices, and teachers' agentive action is performed in micro-level of activity and constrained by their social-professional environment (pp. 8-9).

Second, among the 32 studies, sociocultural theories and ecological perspectives are the two dominant theoretical stances in the analysis of teacher agency, in addition to "a number of perspectives including Bakhtinian, poststructural, ecological, and postcolonial thought" (Ilieva & Ravindran, 2018, p. 8) and positional theory. Concerning sociocultural theory, Feryok (2012) brings up that what distinguishes different branches of sociocultural theory is "how mediation is conceptualized" (p. 97). For example, Huyen Phan and Hamid (2017), relying on Cross' (2009) policy-as-tool within an activity system, claim that their freedom was constrained by factors like time, resource, and other social factors, even though the lack of implementation schemes granted the teachers freedom to develop their own teaching techniques (pp. 43-44). Newcomer and Collier (2015) perceive agency as a sociocultural concept defined by Barker (2008). Through the lens, agency is explored to understand how values, beliefs, knowledge, and experiences can affect a teacher's act within various contexts (Newcomer & Collier, 2015, p. 163). Feryok (2012), taking Leontiev's (1981) activity theory which addresses the gap between personal meaning and social reality (p. 96), considers how social forces and roles of personal experiences influence a language teacher to develop a sense of agency (p. 95). Activity theory is also applied in studies by Yang (2018) as well as Yang and Clark (2018) to analyze teachers' action within social and historical context of certain activity system and illustrate the interrelationship between belief, agency, and action.

Ecological perspective is evident in several articles (e.g., Glas, 2016; Haneda & Sherman, 2018; Hirver & Whitehead, 2018; Ilieva & Ravindran, 2018; Leal & Crookes, 2018; Mifsud & Vella, 2018; Vitanova, 2018). In current literature, although agency has been conceptualized as variable, capacity, or phenomenon/doing (Haneda & Sherman, 2018; Priestley, Biesta, & Robinson, 2015) in ecological perspective, agency as phenomenon or doing is more widely adopted in the exploration of internal and external factors on teachers' enactment of agency. For instance, Glas (2016) explores both internal and external constraints such as teachers' knowledge, family background, classroom dynamics, teachers' lesson plans and materials, professional support, and more in shaping, enabling and limiting teacher agency. Vitanova (2018) argues that "agency [in ecological perspective] is more of a relational phenomenon and is both positioned within an environment that is occupied by other individuals and is marked by temporality" (p. 28). In other words, an individual's degree of agency relies on situational contexts and different temporal frames. In the study, Vitanova (2018) explores how gender and race as identity social markers shape not only teachers' personal experiences but also their interactions with others and mediate their emerging professional agency.

Contexts

The contexts where these studies were conducted present a diverse picture. Among the 32 studies, 14 situated their investigations in EFL contexts with 11 focusing on participants from South and East Asia (e.g., Huyen Phan & Hamid, 2017 in Vietnam; Ishihara, Carroll, Mahler, & Russo, 2018 in Japan; Kang, 2017 in South Korea; Liyanage, Bartlett, Walker, & Guo, 2015 in China) and one from Western Asia (Feryok, 2012 in Soviet Armenia). Studies that are set in other continents are Glas (2016) in Chile and Tutunis and Hacifazlioglu (2018) in Istanbul, Turkey. Additionally, 10 other studies are ESL classroom-based research in English-speaking countries. It is noteworthy that seven of these studies were conducted in the US (e.g., Christiansen, Du, Fang, & Hirvela, 2018; Leal & Crookes, 2018; Newcomer & Collier, 2015) whereas the other ones were located in Australia (Ollerhead, 2010), New Zealand (White, 2018), and Canada (Ilieva & Ravindran, 2018) respectively. Also included in the current investigation are another eight studies centering around bilingual contexts/dual language programs mostly in the US (e.g., Palmer, Henderson, Wall, Zúñiga, & Berthelsen, 2016; Ray, 2009; Wong et al., 2017) with only one exception that is situated in Malta (Misfud & Vella, 2018).

In addition, participants in the identified research include non-native speakers (e.g., Kang, 2017; Nguyen & Bui, 2016) as well as American-born teachers (e.g., Feryok, 2012; Ray, 2009). With

regards to years of teaching in the aforementioned contexts, it ranges from limited teaching experience (e.g., three months in Newcomer & Collier, 2015; little or no teaching experience in Illieva & Ravindran, 2018) to more than 30 years (e.g., Feryok, 2012). Among the selected studies, a majority of the studies specify that the teachers had received certain training of second language teaching by the time of research or were receiving training in teacher preparation programs by the time of data collection (e.g., Varghese & Snyder, 2018).

The school settings also vary across the selected empirical studies. In the dual language/bilingual contexts, most research is administered in kindergarten, preschool, elementary, and secondary classrooms (e.g., Babino & Stewart, 2018; Ray, 2009; Varghese & Snyder, 2018). Among these studies, Newcomer and Collier (2015), however, recruited participants both from second to eighth grade and universities. In the ESL contexts, five studies were conducted in university settings, such as a community college in Leal and Crookes (2018), a vocational training college in Ollerhead (2010), and tertiary institutions in Miller and Gkonou (2018). In contrast, some studies in the dual language/bilingual cluster conducted their research in elementary schools (e.g., Colegrove & Zú ñ iga, 2018; Haneda & Sherman, 2018; Kayi-Aydar, 2015b). Note that some studies did not specify the school settings but highlighted the demographics of the language learners in their studies. For instance, in White (2018), even though it is unclear as to what school context in which this study is situated, it identifies that the student group is comprised of immigrant /refugee learners of English. In the EFL contexts, nearly half of the studies were conducted in elementary and secondary schools (e.g., Hiver & Whitehead, 2018; Ishihara et al., 2018; Kang, 2017; Palmer et al., 2016) whereas the other half are college-based (e.g., Feryok, 2012; Tao & Gao, 2017; Yang, 2018; Yang & Clarke, 2018; Zhang, 2018). In addition, Liyanage et al. (2015) examine both secondary school and college settings in the Inner Mongolian Autonomous Region in China. The geographic environments where these schools are located are mixed, such as urban district (Ray, 2009), remote mountainous area with many ethnic minority groups in Vietnam (Nguyen & Bui, 2016), outskirts of a major Australian city (Ollerhead, 2010), and a suburban context (Zhang, 2018). Note that a few studies highlight the economic and cultural dimensions of the settings, such as the marginalized communities in Colegrove and Zuniga (2018), the disadvantaged Istanbul neighborhoods in Tutunis and Hacifazlioglu (2018), and the high-performance school setting in Ray (2009). Finally, a few studies conduct research in multiple school settings (Liyanage et al., 2015; Newcomer & Collier, 2015; Nguyen & Bui, 2016; Palmer et al., 2016).

Methodologies
Research Design

In terms of research design, all the selected studies pursue qualitative inquiries. 25 studies further identify the types of qualitative study. Specifically, 16 of them are qualitative case studies, while five of them are qualitative ethnographic studies (e.g., Nguyen & Bui, 2016; Palmer et al., 2016; Varghese & Snyder, 2018) and three studies use narrative inquiry (Ishihara et al., 2018; Liyanage et al., 2015; Vitanova, 2018). Two studies (Kang, 2017; Tutunis & Hacifazlioglu, 2018) are qualitative study in nature but also utilize quantitative approach in data analysis. For example, Kang (2017) utilizes both interview transcripts and classroom observations to analyze his data, but he also employs quantitative approach "to determine whether the students' learning outcomes in classes where the LP [language play] was constructed were significantly better than those in classes where LP was not constructed" (p. 88). Mifsud and Vella's (2018) study is the only ethnography in research design which is different from qualitative ethnographic study in terms of the "degrees of orientation to theories from anthropology" (Heath & Street, 2008, p. 121).

Data Collection Method

Regarding data collection methods used in the selected 32 studies, 25 studies are multi-method in nature, which should be considered as a positive feature as researchers can triangulate their data in different ways to enhance the validity of their studies. Interview is the most common strategy, appeared in 30 studies (in most cases an open-ended, semi-structured format); several studies also mention the language used in interviews (e.g., Feryok, 2012 in English; Glas, 2016 in Spanish; Huyen Phan & Hamid, 2017 and Nguyen & Bui, 2016 in Vietnamese). Second most common strategy is classroom observation (n=17), followed by artifacts/documents (e.g., course syllabi, class handouts, students' assignments, curricula vitae, lesson plans, and posts) (n=7), and focus group discussion (n=6). Both field-notes, taken beyond classroom contexts, and survey are used in five studies. Informal conversations, written or digital narratives, journal entries and other (including

meeting or portfolios) are used in fewer than five research studies. Another point worth mentioning is that although various methods are used in the studies, not all of the studies clarify *how* and *why* interviews, classroom observations, and other methods are selected. Some excellent examples (e.g., Babino & Stewart, 2018; Ilieva & Ravindran, 2018) show how this can be achieved.

Most studies used one-on-one oral interviews with their participants to understand their decision-making process and factors influencing their decisions, which in turn is relevant to their exertion of agency. In Feryok's (2012) study, except for oral interviews, he also used semi-structured email interviews although how email interviews were conducted was not explained in the study. Six studies used interviews alone as their data collection method (Glas, 2016; Ishihara et al., 2018; Liyanage et al., 2015; Newcomer & Collier, 2015; Ray, 2009; Venegas-Weber, 2018). In Liyanage et al.'s (2015) study, eight teacher participants were interviewed individually for around 30 to 45 minutes. The interview questions are about those teachers' "(1) experiences of English testing; (2) perceptions of how testing was impacting on them professionally and personally as they dealt with demands of the examination-oriented education system; and (3) beliefs about what effective teachers do" (Liyanage et al., 2015, p. 255).

In Nguyen and Bui's (2016) study, although data were collected and analyzed from both interviews and classroom observations for its larger study, mainly interview data were used for this study. As the authors point out, a variety of topics were explored in interview, including:

> participants' views on the current English LPs [language policies], curriculum, and teacher training, along with students' performance; their own proposals for new or modified policies; their understanding and application of a linguistically and culturally responsive teaching approach; the role of the English language in the province; the roles of students' native languages and Vietnamese; and the influence of English on students' lives, education, and socio-economic opportunities. (Nguyen & Bui, 2016, p. 92)

Salient in these studies is that reflective interviews were used to probe and understand teachers' experiences and decisions they made.

Another finding is that in some studies situated in EFL contexts, interviews were conducted in participants' first languages, and later, interview transcriptions were translated into English. For example, in Kang's (2017) study, "the transcriptions were in Korean, translated into English for reporting" and "[a] professional Korean-English translator confirmed the translations done by the researcher" (p. 87). Similarly, in the study by Huyen Phan and Hamid (2017), the interviews with university English teachers were conducted in Vietnamese, indicating that original interview transcripts were in Vietnamese; nonetheless, the researchers did not clarify how they confirmed their English translations. In other studies, also in EFL contexts, like Liyanage et al. (2015) in China and Feryok (2012) in Armenia, interviews were in English for reasons including participants' advanced English proficiency and researcher-participant not sharing the same first language.

The next popular method used in this body of literature is classroom observations. Nevertheless, compared to relatively richer information about interview data, less information is provided about classroom observations. In many studies, classroom observations are mentioned as part of data collection; nonetheless, details on this data collection method vary across the selected articles. For example, in Ollerhead's (2010) study, classroom observation as a method of data collection was not discussed. Kang (2017) provides a brief introduction about classroom observations with respect to its frequency, "[o]bservations of the classroom were made eight times, once biweekly" (p. 87). More details were not found. An exception is Huyen Phan and Hamid (2017), which is straightforward about the purpose of classroom observations and provides detailed descriptions regarding types of data collected in observations, "[d]ata were gathered on the physical setting of the classes, the people being observed, and teaching and learning activities including the resources, pedagogic styles, curricula and their organisation...These data were descriptive fieldnotes" (p. 45). In analysis, the researchers incorporated data from both interview and classroom observations to support their argument.

Other methods such as survey, documents, focus group discussion, and more are used to collect supplementary data. For example, in Palmer et al.'s (2016) study, the researchers also observed Dual Language Bilingual Education (DLBE) and State of Texas Assessments of Academic Readiness (STAAR) related trainings. Through participating in the training along with the teachers, the researchers understood better on the resources teachers received and the potential pressures imposed on them (Palmer et al., 2016, p. 397). A very interesting data collection approach is introduced in Colegrove and Zúñiga's (2018) study in which the researchers used a multivocal, video-cued ethnographic method. Using this method, the researchers experienced several phases of

data collection. First, they had to record videos of their interest; second, they had to edit the video or several videos into a short film around 20 or 30 minutes with subtitles; third, the film was shown to multiple participants, including their teacher participant, students, and parents from the research site to collect interview data. Last, the researchers showed the video to participants from various external sites through focus group interviews. The benefit of this method lies in its involvement with different stakeholders whereas for researchers who are not familiar with technology, it might be a challenge.

Data Analysis

Regarding data analysis in those selected studies, a large portion of the studies emphasizes the iterative and comparative nature of data analysis in a qualitative fashion. To name a few, Feryok (2012) constantly compared the collected data related to the research topic. The author specifies that salient information was identified and (re)organized (p. 99). Nguyen and Bui (2016) emphasize a recursive process that sought patterns, themes, and categories that emerged from data. Similarly, Liyanage et al. (2015) united "recurring ideas and experiences" (p. 256) and repeatedly examined the interview components. Newcomer and Collier (2015) dealt with their data sets both individually and collaboratively. Likewise, Huyen Phan and Hamid (2017) state that the data analysis focused on the (re)interpretation and practice of learner autonomy in the English Language classrooms of the four English lectures (p. 46). Babino and Stewart (2018) utilized "constant comparative method" to examine each concept which led to emerging themes. Also, Zhang (2018) dealt with the transcribed texts repeatedly until preliminary codes were ready.

With regards to the procedures of data analysis, most studies conducted data collection and analysis simultaneously and followed an inductive and interpretive process for analyzing data while some mentioned the inclusion of deductive components (e.g., Venegas, 2018). Thematic analysis with coding and memos is a prominent technique used by the selected studies for data analysis (e.g., Ilieva & Ravindran, 2018; Ishihara et al., 2018; Li & De Costa, 2017; Liyanage et al., 2015; Mifsud & Vella, 2018; Newcomer & Collier, 2015; Palmer et al., 2016; Tao & Gao, 2017; Varghese & Snyder, 2018; Wong et al., 2017). In particular, Newcomer and Collier (2015) identified and refined patterns and themes as local theory that revealed the constraints faced by teachers while exercising agency (p. 166). Palmer et al. (2016) and Liyanage et al. (2015) specify that their studies included a deductive process. Palmer et. al. (2016) analyzed each school's data set thematically and explored "(dis)confirming evidence" (p. 398) during the identification of possible themes. Liyanage et al. (2015) uncovered recurring themes and meanwhile checked if the evidence fits the themes well.

Notably, conceptual or theoretical frameworks were used in several studies to assist the process of data analysis. For instance, Babino and Stewart (2018) employed theoretical propositions (Yin, 2013, cited in Babino & Stewart, 2018, p. 280) and grounded theory (Corbin & Strauss, 2008, cited in Babino & Stewart, 2018, p. 280) for guiding their analytic procedures. For another, using activity theory as theoretical framework, Yang and Clark (2018) and Yang (2018) scrutinized the interactions of different activity systems. Similarly, Ray (2009) employed template analysis with *a priori* themes drawing from Bandura's (1989) notion of human agency (p. 126), which allowed the researcher to compare data under a framework.

Discourse analysis is also another analytic tool that appears across several studies (e.g., Christiansen et al., 2018; Colegrove & Zuñiga, 2018; Glas, 2016; Yang & Clark, 2018). Taking Christiansen et al. (2018) as an example, the study utilized "narrative inquiry" to uncover participants' nuanced reflections on their agentive work. For another, in White (2018), data was analyzed relying upon the notion of narrative accounts and stance as "emergent product" through social interactions (p. 582). Similarly, Leal and Crookes (2018) employed a model of teacher agency for social justice for guiding the analytic procedures.

It is worth noting that a few studies used data analysis programs to assist their analysis procedures. For instance, Ray (2009) states that Super Hyper Qual was utilized to analyze a clean data set (p. 120) for a systematic analysis based on the teacher agency template drawn from its theoretical framework. Likewise, Kang (2017) indicates that the statistical procedure of the repeated measurement design was used to determine whether the students' learning outcomes in classes where the LP was constructed were significantly better than those in classes where LP was not constructed (p. 88). Miller and Gkonou (2018) involved qualitative data management software Atlas.ti to inspect the data. Additionally, Tutunis and Hacifazliogulu (2018) conducted analysis of each questionnaire partly from a quantitative analytic tool named Statistical Package for Social Sciences (p. 111).

RQ2: What are the major findings, methodological and ethical issues, and implications discussed in the selected articles?

Major Findings

Through iteratively and comparatively examining the sections of findings in these articles, five thematic categories emerged. Note that in several studies, findings entail aspects that go beyond the realm of teacher agency. For instance, in addition to focusing on teacher agency, findings regarding identity formation are also important in Hiver and Whitehead's (2018) research. In the current investigation, however, only findings that are closely relevant to teacher agency will be addressed. Also important is that the categories proposed by this synthesis may overlap with each other in the sense that some studies cover multiple themes. For instance, Zhang (2018) focuses on teachers' professional development but the findings are partly generated from teacher reflections. As the professional development portion appears more prominent, the current investigation decides to include this article in the category of professional development and agency instead of the teacher reflection category. The following discussion will provide more details concerning the thematic categories and highlight exemplar studies conducted in different language contexts.

Among the 32 studies, 11 explore policy constraints and enactment of teacher agency in classroom practices (e.g., Babino & Stewart, 2018; Haneda & Sherman, 2018; Huyen Phan & Hamid, 2017). In other words, the participant teachers' adjustments and resistance owing to the constraints of language policies are the major findings in those studies. For instance, in an ESL context, Ollerhead (2010) reveals that policy conditions constrain one participant teacher's "ability to act agentively as a teacher" (p. 616) while the other teacher has more freedom to make pedagogical decisions without receiving training and materials concerning the policies. Along with the same line of research but under a different language context (dual language program), the findings in Babino and Stewart (2018) reveal that the teacher participants who were pressured by the program climate that English is considered as the hegemonic language tried to collectively remodel their language programs in the best interest of their students. Similarly, in Huyen Phan and Hamid (2017) which is under an EFL setting, the findings uncover that teachers, motivated by their sense of responsibility, could empower their students by promoting learner autonomy in micro-level classroom practices under the circumstance of no explicit strategies provided by the macro-level policies.

10 studies focus on factors related to agency and how these factors are translated into classroom actions (e.g., Colegrove & Zú ñ iga, 2018; Feryok, 2012; Glas, 2016; Kang, 2017). These factors include previous personal experiences in the sociocultural environment, individual characteristics, teacher reflections and perceptions, and teacher and student attitudes and motivations. For instance, Feryok's (2012) study shows the personal experiences of the participant, Nune—as a student, a teacher, and a teacher trainer—deeply influenced her teacher agency and guided her individual actions in her local EFL context (p. 99). Several studies in this category show interest in the interplay of teacher reflections (or perceptions) and agentive actions. In a bilingual context, Ray (2009) examines the characteristics of teacher agency and reveals that teachers' sense of agency—such as mastery and vicarious experience—translates into instructional behaviors as manifestations (p. 128). Focusing on an ESL context, Colegrove and Zú ñ iga (2018) illustrate that teachers could create spaces for agentive classroom practices for marginalized students when they are not afraid of failures in implementing "dynamic teaching practices" (p. 188). In an EFL setting in Japan, Ishihara et al. (2018) discover that teachers' translingual practices (involving Japanese and Japanese cultures) could be used as a mediational tool to support their "agentive acts" (p. 89) in classrooms.

Discussions in another nine studies are concerned with teacher identity and agency (e.g., Hiver & Whitehead, 2018; Kayi-Aydar, 2015b; Leal & Crookes, 2018). Most studies have unveiled the interactive nature of professional identity development and teacher agency. For instance, Kayi-Aydar's (2015b) study, situated in an ESL context, finds that positioning and repositioning identities play crucial roles in shaping teachers' agentive acts in classrooms. In an EFL context, Li and De Costa (2017) discover that through constructing a teacher identity as someone who focuses on students' knowledge and career needs instead of simply preparing students for tests, teachers are able to exercise their beliefs and make decisions in classroom practices (p. 281). In a dual language context, Venegas-Weber (2018) suggests that developing "linguistic and cultural identities" (p. 165) potentially creates spaces for implementing agency in classrooms.

Two studies—Christiansen et al. (2018) and Zhang (2018)—discuss the interactions of professional development and teacher agency. Specifically, Christiansen et al. (2018) seek to

investigate the relationship between agency and expertise among graduate teaching assistants who teach ESL writing courses. The findings illustrate that professional learning community helps to make strong connections between expertise and teacher agency. Zhang (2018) explores an EFL distance education teacher's agency as a case study and reveals that the teacher copes with the challenges of professional development in the teaching context by developing academically in writing.

Methodological and Ethical Issues

Among the articles selected for analysis, almost half of the studies analyzed (*n*=15) do not explicitly mention any ethical or methodological limitation. The challenges and/or limitations mentioned in the studies are first categorized into ethical or methodological issues. Methodological issues include challenges or limitations encountered *during* or posed *in relation to* data collection, analysis, and reporting of the findings. Some studies (*n*=18) do not mention any methodological limitation or issue. The methodological issues mentioned in the studies are as followings: (1) small sample/data size or short time period, (2) "generalizability" (Yang & Clark, 2018), (3) limitation due to the type of data collected (i.e., data consisting of only self-report without any observation), (4) participant recruitment—recruiting those who showed to be resourceful, (5) limitations of the focus, (6) constraints in data collection, (7) missing themes that are outside of the theoretical framework, (8) decontextualization of data while coding, (9) time lag between the time of event and the interviews, and (10) challenge to report in a coherent and compelling manner "based only on selected excerpts from a fairly extensive dataset" (Hirver & Whitehead, 2018, p. 5).

One key limitation that affects studies on teacher agency is limited sources of data. The limitations on size and/or scope of data are mentioned more often than the others; those limitations relate to issues on generalizability, because the studies involve a relatively small number of participants or are conducted as a case study. Moreover, most of these studies employed interviews as a main method of collecting data. For instance, a number of articles (e.g., Babino & Stewart, 2018; Christiansen et al., 2018; Glas, 2016) collected data only from methods that involve self-reports such as interviews, focus groups, journal entries, or surveys without observations, and this poses a challenge that the responses from interviews may not fully portray their agencies as teachers and may not align with their actual performances. Another methodological limitation identified is from transcription. Feryok (2012) pinpoints that all transcription involves omission, and especially due to focusing on content analysis, not many conversational details such as intonation and pauses are included (p. 98). Most of the limitations stated in the studies are not exclusively applicable to research on teacher agency; yet these limitations do relate to research on the topic, especially considering that many of the studies on teacher agency used classroom observation and interviews for data collection.

Ethical issues are not explicitly mentioned in most (*n*=29) of these studies. The studies that mention ethical issues are Feryok (2012), Ilieva and Ravindran (2018), and Kayi-Aydar (2015b). For instance, Feryok (2012) was an instructor of the participants, which could have affected their responses. In order to minimize the effect, Feryok (2012) intended to conduct research as a "private individual" (p. 98), not as how she had been known as the spouse of a Western diplomat. Similarly, Kayi-Aydar (2015b) had the role of a professor of the participants and, thus, the participants could have attempted "to perform particular identities to please" (p. 102); to reduce the effect, the research was conducted *after* the semester was over. Another related limitation is with the recruitment of the participants. Ilieva and Ravindran (2018) explicitly state their "ethical dilemma" (p. 16), that the goal of social justice and equity—the stances that are enacted as a center—is "unstated in materials inviting applicants" (p. 16) to the program.

Several strategies were employed by the authors to address methodological and ethical issues in their studies. Kang (2017), Kayi-Aydar (2015b), and Ray (2009) further elaborate the ways they have exercised with caution. For example, Kang (2017) was extremely careful "to ensure that the interviewees would not be influenced to provide answers preferred by him" (p. 87). Similarly, Ray (2009) indicates several times that the researcher was attentive, especially during the data analysis phase, to "determine whether themes outside of antecedent/manifestation framework emerged from the data" (p. 126). Other strategies include triangulation (Feryok, 2012) by using three sets of data collected through different methods (p. 99) and collaboration with a professional translator for translation of transcripts (Kang, 2017) to address language-related issues.

Implications

Discussions and implications conveyed by the aforementioned studies are multifold, and some prominent ones are synthesized as follows. First, it is discussed that teacher agency is affected by varying factors, including time, resources, and other social factors. Navigating among all those factors, teachers, as change agents of pedagogical reforms, should be empowered while designing, implementing, and evaluating educational policies (Huyen Phan & Hamid, 2017; Liyanage et al., 2015; Newcomer & Collier, 2015; Nguyen & Bui, 2016; Ollerhead, 2010). This empowerment is reflected in several ways, such as offering specific policy implementation schemes, providing necessary professional training and teaching resources, and more importantly getting teachers' voices heard at institutional and departmental levels. These ways of empowering teachers in local contexts can enable them to better take control of their work and teach towards meaningful pedagogies (Huyen Phan & Hamid, 2017).

It is argued that support from higher levels like school principals or district leadership is one of the most important factors in enhancing teachers' agency (Priestley, 2011), which is also reflected in the studies. Specifically, Babino and Stewart (2018) imply that administrators may use their own agency to enact a collective leadership with teachers to create more equitable assessment decisions. Likewise, Colegrove and Zú ñ iga (2018) suggest that district level should provide space for teachers to expand their pedagogical repertoire through innovative programs like Project-Based Instruction (PBI). These suggestions on leaving space for teacher agency from higher levels are further confirmed by Glas (2016) that "It is necessary for decision-makers at higher levels to regain trust in the capacities of the individual teachers, in their abilities to relate to their students' contexts and to find the most appropriate materials, teaching–learning strategies and evaluation procedures" (p. 459).

Second, multiple implications from the 32 articles are for teacher education and professional training. It is suggested that teacher education should provide more relevant courses responsive to teachers' local contexts and help teachers to develop teaching techniques and methods feasible in their contexts (Kang, 2017; Nguyen & Bui, 2016). Several other studies (e.g., Babino & Stewart, 2018; Christiansen et al., 2018; Haneda & Sherman, 2018) suggest that the knowledge base of teacher education should be expanded to include the discussion of teacher agency interacting with other concepts like teacher expertise, teacher roles/identities, and teacher beliefs and to provide space for preservice teachers to develop agency while engaging in reflection (e.g., Colegrove & Zú ñ iga , 2018; Glas, 2016; Miller & Gkonou, 2018; Wong et al., 2017; Yang, 2018) and critical reflection (e.g., Ishihara et al., 2018; Kayi-Aydar, 2015b). This agency-oriented approach to teacher education should "guide student teachers to become more aware of their personal resources and learn to capitalize on them to seize available contextual opportunities" (Tao & Gao, 2017, p. 354).

Third, in terms of professional development, Nguyen and Bui (2016) reveal that participants in their study made their pedagogical changes and take a critical stance as language policy (LP) implementers because of "critical conversations" and their "engagement in the discussion on LP implementation with the researchers and other teachers" (p. 101). Therefore, they suggest more collaborative discussions among teachers within school contexts or a "third space" like "critical friend groups" and "mentoring and peer mentoring" (p. 101) for "critical consciousness and work towards realizing their potential as agents of [transformation]" (p. 101). This suggestion echoes what Palmer et al.(2015) argue in their study that collaborative spaces shape decision-making process (p. 410). Christiansen et al. (2018) propose Professional Learning Community (PLC) which include both expert and novice teachers to promote bidirectional interaction between agency and expertise for professional development. Yang (2018) particularly points to in-service EFL teachers' professional development in China. Yang (2018) proposes that "institutions in China should provide teachers, especially novice teachers, with access to wider professional notions, approaches and discourses about EFL teaching" (p. 50).

Next, future directions for research are also provided. This includes how professional development promotes dual language teachers' agency (Ray, 2009) and how agency is taken among different stakeholders (Palmer et al., 2016). Several studies (e.g., Kayi-Aydar, 2015b; Tao & Gao, 2017) on the connection between teacher identity and teacher agency comment that more research should be done on this topic in the context of educational change. Ishihara et al. (2018) discuss that wider range of teacher experience should be investigated, and Kayi-Aydar (2015b), in particular, suggest that mentor teachers' voice should be heard. A couple of studies, in addition, offer implications for research design. For example, Glas (2016) indicates that teacher and research collaboration on action research and large-scale quantitative study to complement qualitative studies

is needed; Tao and Gao (2017) recommend more longitudinal ethnographic study. Similarly, Babino and Stewart (2018) and Wong et al. (2017) ask for more exploration of comparative case study. Some other studies suggest new lens and approach to research (teacher) agency and theorize teacher agency and teacher identity, like complexity/dynamic systems theory (Hirver & Whitehead, 2018) and new materialism (Hirver & Whitehead, 2018). Last, several studies involving language policies have indicated the misalignment between school, district, community expectations, and language policies, which in turn has affected teacher agency. Therefore, more communication with and coordination from different stakeholders should be established and supported (Huyen Phan & Hamid, 2017; Palmer et al., 2016).

Discussion and Conclusion

Teacher agency, although has attracted attention in the education literature in recent years, still remains under-examined in general, even less in ESL/EFL/bilingual language education contexts. However, the selected literature underscores the importance of teacher agency. Nguyen and Bui (2016) point out that "[t]eacher agency is critical to the process of implementing educational changes" (p. 89), and Hamid and Nguyen (2016) claim that "agency is not exactly an exercise of free will; rather, teachers are in a way, forced to exert themselves if they wanted to help students to meet policy goals" (p. 35). Both statements indicate that teachers are essential mediators between macro and micro contexts; yet, the collection of articles in this synthesis conveys that more relevant studies should be conducted and more voices from teachers be heard. Hence, based upon the previous findings on the existing studies, we offer following future directions for research, teacher education, and professional development.

Future Directions for Research

To address the aforementioned issue, this research synthesis attempts to provide further directions for research. Based upon the analyses, several aspects on teacher agency in ESL/EFL/bilingual contexts can be further studied. First, the focus on current literature has been limited to policy and implementation, particularly in EFL contexts (e.g., Huyen Phan & Hamid, 2017; Nguyen & Bui, 2016) or educational change (e.g., Liyanage et al., 2015). In bilingual contexts, according to Babino and Stewart (2018), "most bilingual teacher agency research focuses on the single acts of individual teachers at the classroom levels" (p. 273). Consequently, more studies on different topics in different contexts can offer insights to other aspects of teacher agency. Second, the range of contexts examined in those studies could be expanded. It is found that most studies in bilingual contexts have examined teacher agency in elementary contexts (Kang, 2017; Newcomer & Collier, 2015; Ray, 2009; Palmer et al., 2016), half of the studies in ESL contexts are in university level, and around half of studies in EFL contexts have investigated in elementary and secondary contexts. Therefore, little research has been conducted in settings of kindergarten and high schools. In addition, national contexts beyond Vietnam, China, South Korea, United States, and Australia could be studied for further comprehension on teacher agency in this globalized age. Moreover, for bilingual education, California and Texas are the two main contexts in which bilingual education were studied; more states which provide bilingual education should be explored. This is essential because bilingual education in different states vary according to state policy, standards, and demographics..

Third, although interviews and classroom observation are useful methods in data collection in the study of teacher agency, other data collection methods—such as playback sessions, survey, mapping, collecting artifacts, and focus groups—can be used to triangulate and provide data from different angles. Fourth, studies (e.g., Babino & Stewart, 2018; Tao & Gao, 2017; Wong et al., 2017) suggest that different qualitative studies, like longitudinal ethnographic study and comparative case study, should be welcomed. Except for qualitative research, quantitative data "might have more persuasive power to counter educational policies that are currently undermining teacher creativity and their (sense of) agency" (Glas, 2016, p. 459); therefore, quantitative research should be encouraged as well. Fifth, although the native and nonnative teacher dichotomy in the discussion of teacher agency in the selected articles is not prevalent, this dichotomy is debated in research on teacher identity which is a construct closely related to teacher agency; therefore, this dichotomy should be discussed more in the study of teacher agency. Beyond this identity, other minority teachers (e.g., Leal & Crookes, 2018 on a queer English language teacher) should gain more attention. Last, although a variety of theoretical frameworks and analytical approaches have been drawn upon to study teacher agency, socio-cultural and ecological perspectives are dominant in research. Hence, more research on other perspectives or the mix of multiple perspectives might be

able to further contribute to the current literature.

Future Directions for Teacher Education and Professional Development

This research synthesis also offers implications for teacher education and professional development. First of all, several articles indicate that teacher education and training should take practitioners' local contexts into consideration, facilitate teachers to incorporate theories, concepts, and pedagogies from global to local contexts (Huyen Phan & Hamid, 2017, p. 52), and help teachers develop teaching techniques and methods feasible in their contexts (Kang, 2017). Second, Nguyen and Bui's (2016) study reveals the "mistrust" (p. 96) between practitioners and teacher educators due to "local and university trainers' lack of understanding of multiculturalism, multiple learning styles, and the socio-economic situations of minority students" (p. 96). To ameliorate this mistrust, teacher educators and teachers should have more communication and understanding of each other's working environment. Third, the selected studies (e.g., Babino & Stewart, 2018; Colegrove & Zúñiga, 2018; Glas, 2016) propose that modern educational system should afford teachers more leeway to enact their agency. To achieve the goal, all stakeholders (students, teachers, student families, and other higher levels) who care for humanizing pedagogy should participate in making the changes. For teachers, Leal and Crookes (2018) suggest that they should "develop an awareness of the contradictions between their 'sense of purpose' and the educational and social structures in which their work is located" (p. 38).

Fourth, several studies (e.g., Feryok, 2012; Newcomer & Collier, 2015; Ollerhead, 2010) are indicative of the fact that the enactment of teacher agency is dependent on teachers' experiences, ideas, and beliefs. Huyen Phan and Hamid (2017) state that the awareness of their teacher role and sense of responsibility are essential for students and academic well-being. Therefore, critical examination and reflection of their background and beliefs in teaching and learning might contribute to their awareness of teacher's role as an agent in classrooms. Fifth, adequate training in teacher education and professional development programs should be offered to equip and empower teachers with necessary guidance and support. Last, as language teachers become more professionalized in their work through exercising agency, the concept of teacher agency should be included in teacher education as part of teacher knowledge. Also, continued support for teachers from workplace should be given as agency is shaped by "social interactions and achieved in particular situations" (Mifsud & Vella, 2018, p. 273). In a word, studies on teacher agency should be continued in ESL/EFL/bilingual contexts in order to further understand teachers' agency and active participation in diverse educational settings.

References

Achinstein, B., & Ogawa, R. T. (2012). New teachers of color and culturally responsive teaching in an era of educational accountability: Caught in a double bind. *Journal of Educational Change, 13*, 1-39.

Babino, A. (2017). Same program, distinctive development: exploring the biliteracy trajectories of two dual language schools. *Bilingual Research Journal, 40*(2), 69-86.

Babino, A., & Stewart, M. A. (2018). Remodeling dual language programs: Teachers enact agency as critically conscious language policy makers, *Bilingual Research Journal, 41*(3), 272-297.

Bakhtin, M. M. (1984). *Rabelais and his world*. Translated by H. Iswolsky. Bloomington, IN: Indiana University Press.

Bandura, A. (1989). Human agency in social cognitive theory. *American Psychologist, 44*, 1175–1184.

Barker, C. (2008). *Cultural studies: Theory and practice* (3rd ed.). Thousand Oaks, CA: Sage.

Bartolomé, L. I. (2004). Critical pedagogy and teacher education: Radicalizing prospective teachers. *Teacher Education Quarterly, 31*, 97-122.

Benesch, S. (2018). Emotions as agency: Feeling rules, emotion labor, and English language teachers' decision-making. *System, 79*, 60-69.

Biesta, G., Priestely, M., & Robinson, S. (2015). The role of beliefs in teacher agency. *Teaching and Teacher Education, 21*, 624–640. doi:10.1080/02660830.2007.11661545

Bouchard, J. (2016). Monolingual education in policy discourse and classroom practice: A look into Japanese junior high school EFL classrooms. In P. C. L. Ng & E. F. Boucher-Yip (Eds.), *Teacher agency and policy response in English language teaching* (pp. 41-57). New York, NY: Routledge. Retrieved from https://ebookcentral.proquest.com/lib/ohiostate-ebooks/detail.action?docID=4626106

Christiansen, M. S., Du, Q., Fang, M., & Hirvela, A. (2018). Doctoral students' agency as second language writing teachers: The quest for expertise. *System, 79,* 19-27.

Colegrove, K., & Zúñiga, C. E. (2018). Finding and enacting agency: An elementary ESL teacher's perceptions of teaching and learning in the era of standardized testing. *International Multilingual Research Journal, 12*(3), 188-202.

Cooper, H., & Hedges, L. V. (2009). Research synthesis as a scientific process. In H. Cooper, L. V. Hedges, & J. C. Valentine (Eds.), *The handbook of research synthesis and meta-analysis* (pp. 3-16). New York, NY: Russell Sage Foundation.

Corbin, J., & Strauss, A. (Eds.). (2008). Basics of qualitative research: Techniques and procedures for developing grounded theory. Thousand Oaks, CA: Sage.

Cross, R.(2009). A sociocultural framework for language policy and planning. *Language Problems & Language Planning, 33*(1), 22–42.

Dantas-Whitney, M., Clemente, A., & Higgins, M. J. (2012). Agency, identity and imagination in an urban primary school in Southern Mexico. *Calidoscó pio, 10*(1), 114-124.

Dubetz, N. E., & de Jong, E. J. (2011). Teaching advocacy in bilingual programs. *Bilingual Research Journal, 34*(3), 248-262.

Edwards, A. (2015). Recognising and realising teachers' professional agency. *Teachers and Teaching, 21*(6), 779-784.

Feryok, A. (2012). Activity theory and language teacher agency. *The Modern Language Journal, 96*(1), 95-107.

Fitts, S. (2009). Exploring third space in a dual-language setting: Opportunities and challenges. *Journal of Latinos and Education, 8*(2), 87-104.

Flores, B. B. (2001). Bilingual education teachers' beliefs and their relation to self-reported practices. *Bilingual Research Journal, 25*(3), 275-299.

Fullan, M.(1993). *Changing forces: Probing the depth of educational reform.* London: Palmer Press.

García, O. (2009). *Bilingual education in the 21ᵗ century: A global perspective.* Malden, MA: Wiley/Blackwell.

Glas, K. (2016). Opening up 'spaces for manoeuvre': English teacher perspectives on learner motivation. *Research Papers in Education, 31*(4), 442-461. doi:10.1080/02671522.2015.1049287

Glasgow, J. P. (2016). Policy, agency and the (non)native teacher: "English classes in English" in Japan's high schools. In P. C. L. Ng & E. F. Boucher-Yip (Eds.), *Teacher agency and policy response in English language teaching* (pp. 58-73). New York, NY: Routledge. Retrieved from https://ebookcentral.proquest.com/lib/ohiostate-ebooks/detail.action?docID=4626106

Hamid, M. O., & Nguyen, H. T. M. (2016). Globalization, English language policy, and teacher agency: Focus on Asia. *International Education Journal: Comparative Perspectives, 15*(1), 26-44.

Hamid, M. O., Zhu, L., & Baldauf, R. B. (2014). Norms and varieties of English and TESOL teacher agency. *Australian Journal of Teacher Education, 30*(10). 77-95.

Haneda, M., & Sherman, B. (2018). ESL teachers' acting agentively through job crafting. *Journal of Language, Identity and Education, 17*(6). 402-415. doi:10.1080/15348458.2018.1498340

Harper, C., & de Jong, E. (2004). Misconceptions about teaching English-language learners. *Journal of Adolescent & Adult Literacy, 48,* 152-162.

Heath, S. B., & Street, B. V. (2008). *On ethnography: Approaches to Language and Literacy Research.* New York, NY: Teachers College Press.

Henderson, K. I. (2017). Teacher language ideologies mediating classroom-level language policy in the implementation of dual language bilingual education. *Linguistics and Education, 42,* 21-33.

Hiver, P., & Whitehead, G. H. K. (2018). Sites of struggle: Classroom practice and the complex dynamic entanglement of language teacher agency and identity. *System, 79,* 70-80.

Hoang, H., & Truong, L. B. (2016). Teacher agency and autonomy in rural Vietnam. In P. C. L. Ng & E. F. Boucher-Yip (Eds.), *Teacher agency and policy response in English language teaching* (pp. 188-202). New York, NY: Routledge. Retrieved from https://ebookcentral.proquest.com/lib/ohiostate-ebooks/detail.action?docID=4626106

Hopkins, M. (2013). Building on our teaching assets: The unique pedagogical contributions of bilingual educators. *Bilingual Research Journal, 36*(3), 350-370.

Huyen Phan, T.T., & Hamid, M. O. (2017). Learner autonomy in foreign language policies in Vietnamese universities: An exploration of teacher agency from a sociocultural perspective. *Current Issues in Language Planning, 18*(1), 39-56.

Ilieva, R., & Ravindran, A. (2018). Agency in the making: Experiences of international graduates of a TESOL program. *System, 79*, 7-18.

Ishihara, N., Carroll, S. K., Mahler, D., & Russo, A. (2018). Finding a niche in teaching English in Japan: Translingual practice and teacher agency. *System, 79*, 81- 90.

Kang, D. (2017). The multifaceted ecology of language play in an elementary school EFL classroom. *International Journal of Bilingual Education & Bilingualism, 20*(1), 84-101.

Kayi-Aydar, H. (2015a). Multiple identities, negotiations, and agency across time and space: A narrative inquiry of a foreign language teacher candidate. *Critical Inquiry in Language Studies,* *12*(2), 137–160. Retrieved from http://proxy.lib.ohiostate.edu/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=eric&AN=EJ1063925&site=ehost-live

Kayi-Aydar, H. (2015b). Teacher agency, positioning, and English language learners: Voices of pre-service classroom teachers. *Teaching and Teacher Education, 45*, 94-103.

Lambert, W. E. (1975). Culture and language as factors in learning and education. In A. Wolfgang (Ed.), *Education of immigrant students: Issues and answers* (pp. 55-83). Toronto, ON: Ontario Institute for Studies in Education.

Lasky, S. (2005). A sociocultural approach to understanding teacher identity, agency and professional vulnerability in a context of secondary school reform. *Teaching and Teacher Education, 21*(8), 899-916.

Leal, P., & Crookes, G. V. (2018). "Most of my students kept saying, 'I never met a gay person'": A queer English language teacher's agency for social justice. *System, 79*, 38-48.

Leontiev, A. N. (1981). The problem of activity in psychology. In J. V. Wertsch (Ed.), *The concept of activity in Soviet psychology.* Armonk, NY: M. E. Sharp.

Li, W., & De Costa, P. I. (2017). Professional survival in a neoliberal age: A case study of an EFL teacher in China. *Journal of Asia TEFL, 14*(2), 277-291.doi:10.18823/asiatefl.2017.14.2.5.277

Liyanage, I., Bartlett, B., Walker, T., & Guo, X. (2015). Assessment policies, curricular directives, and teacher agency: Quandaries of EFL teachers in Inner Mongolia. *Innovation in Language Learning and Teaching, 9*(3), 251-264.

Lucas, T., & Villegas, A. M. (2011). A framework for preparing linguistically responsive teachers. In T. Lucas (Ed.), *Teacher preparation for linguistically diverse classrooms: A resource for teacher education* (pp. 55-72). New York, NY: Taylor & Francis.

Meierdirk, C. (2018). The impact of the social environment on the student teacher's agency. *Teaching Education, 29*(1), 33-46.

Mifsud, C. L. & Vella, L. A. (2018). Teacher agency and language mediation in two Maltese preschool bilingual classrooms. *Language, Culture & Curriculum, 31*(3), 272-288.

Miles, M. B., Huberman, A. M., & Saldañ a, J. (2013). *Qualitative data analysis.* Thousand Oaks, CA: SAGE Publications.

Miller, E. R., & Gkonou, G. (2018). Language teacher agency, emotion labor and emotional rewards in tertiary-level English language programs. *System, 79*, 49-59.

Newcomer, S. N., & Collier, L. C. (2015). Agency in action: How teachers interpret and implement Arizona's 4-hour structured English immersion program. *International Multilingual Research Journal, 9* (3), 159-176. doi:10.1080/19313152.2015.1048179

Ng, P. C. L., & Boucher-Yip, E. F. (2016). *Teacher agency and policy response in English language teaching.* New York, NY: Routledge. Retrieved from https://ebookcentral.proquest.com/lib/ohiostate-ebooks/detail.action?docID=4626106

Nguyen, H. M., & Bui, T. (2016). Teachers' agency and the enactment of educational reform in Vietnam. *Current Issues In Language Planning, 17*(1), 88-105.

Norris, J., & Ortega, L. (2007). The future of research synthesis in applied linguistics: Beyond art or science. *TESOL Quarterly, 41*(4), 805-815.

Ollerhead, S. (2010). Teacher agency and policy response in the adult ESL literacy classroom. *TESOL Quarterly, 44*(3), 606-618.

Ollerhead, S., & Burns, A. (2016). Teacher agency and policy response in the Australia adult ESL literacy classroom. In P. C. L. Ng & E. F. Boucher-Yip (Eds.), *Teacher agency and policy response in English language teaching* (pp. 105-119). New York, NY: Routledge. Retrieved

from https://ebookcentral.proquest.com/lib/ohiostate-ebooks/detail.action?docID=4626106

Ortega, L. (2015). Research synthesis. In B. Paltridge & A. Phakiti (Eds.), *Research methods in applied linguistics: A practical resource.* New York, NY: Bloomsbury Publishing.

Osman, S., & Ahn, E. (2016). Navigating change: Kazakhstani English language teachers' response to multiscalar education reforms. In P. C. L. Ng & E. F. Boucher-Yip (Eds.), *Teacher agency and policy response in English language teaching* (pp. 133-147). New York, NY: Routledge. Retrieved from https://ebookcentral.proquest.com/lib/ohiostate-ebooks/detail.action?docID=4626106

Ovando, C. J., & Combs, M. C. (2012). *Bilingual and ESL classrooms: Teaching in multicultural contexts* (5th ed.). New York, NY: McGraw-Hill.

Palmer, D. (2010). Race, power, and equity in a multiethnic urban elementary school with a dual-language "strand" program. *Anthropology & Education Quarterly, 41*(1), 94-114.

Palmer, D. (2011). The discourse of transition: Teachers' language ideologies within transitional bilingual education programs. *International Multilingual Research Journal, 5*(2), 103-122.

Palmer, D., Henderson, K., Wall, D., Zúñiga, C. E., & Berthelsen, S. (2016). Team teaching among mixed messages: Implementing two-way dual language bilingual education at third grade in Texas. *Language Policy, 15*(4), 393-413.

Palmer, D., & Martínez, R. A. (2013). Teacher agency in bilingual spaces: A fresh look at preparing teachers to educate Latina/o bilingual children. *Review of Research in Education, 37*(1). 269-297.

Priestley, M. (2011). Schools, teachers, and curriculum change: A balancing act? *Journal of Educational Change, 12*(1), 1–23. doi:10.1007/s10833-010-9140-z

Priestley, M., Biesta, G., & Robinson, S. (2013). Teachers as agents of change: Teacher agency and emerging models of curriculum. In M. Priestley & G. Biesta (Eds.), *Reinventing the curriculum: New trends in curriculum policy and practice* (pp. 187- 206). New York, NY: Bloomsbury.

Priestley, M., Biesta, G., & Robinson, S. (2015).Teacher agency: An ecological approach. London, UK: Bloomsbury.

Priestley, M., Edwards, R., Priestley, A., & Miller, K. (2012). Teacher agency in curriculum making: Agents of change and spaces for manoeuvre. *Curriculum Inquiry, 42*(2), 191-214.

Ray, J. M. (2009). A template analysis of teacher agency at an academically successful dual language school. *Journal of Advanced Academic, 21*(1), 110-141.

Ruiz, R. (1984). Orientation in language planning. *National Association for Bilingual Education Journal, 8*(2), 15-34.

Shabir, M. (2017). Student-teachers' beliefs on the use of L1 in EFL classroom: A global perspective. *English Language Teaching, 10*(4), 45-52.

Shin, S. J. (2018). *Bilingualism in schools and society: Language, identity, and policy.* London, UK: Routledge Taylor and Francis.

Song, B., & Kim, T. Y. (2016). Teacher (de)motivation from an activity theory perspective: Cases of two experienced EFL teachers in South Korea. *System, 57*, 134–145. Retrieved from http://proxy.lib.ohio-state.edu/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=mzh&AN=2017651508&site=ehost-live

Stritikus, T. T. (2003). The interrelationship of beliefs, context and learning: The case of a teacher reacting to language policy. *Journal of Language, Identity and Education, 2*, 29-52. doi:10.1207/S15327701JLIE0201_2.

Suri, H. (2013). Epistemological pluralism in research synthesis methods. *International Journal of Qualitative Studies in Education (QSE), 26*(7), 889–911. https://doi.org/10.1080/09518398.2012.691565

Suri, H., & Clarke, D. (2009). Advancements in research synthesis methods: From a methodologically inclusive perspective. *Review of Educational Research, 79*(1). 395-430.

Talalakina, E. V., & Stukal, D. (2016) Incorporating academic skills into EFL curriculum: Teacher agency in response to global mobility challenges. In P. C. L. Ng & E. F. Boucher-Yip (Eds.), *Teacher agency and policy response in English language teaching* (pp. 120-132). New York, NY: Routledge. Retrieved from https://ebookcentral.proquest.com/lib/ohiostate-ebooks/detail.action?docID=4626106

Tang, E., Lee, J., & Chun, C. (2012). Development of teaching beliefs and the focus of change in the process of pre-service ESL teacher education. *Australian Journal of Teacher Education, 37*(5), 90-107.

Tao, J., & Gao, X. (2017). Teacher agency and identity commitment in curricular reform. *Teaching and Teacher Education, 63,* 346-355.

Toohey, K. (2007). Autonomy/agency through socio-cultural lenses. In A. Barfield & S. Brown (Eds.), *Reconstructing autonomy in language education.* Basingstoke, England: Palgrave Macmillan.

Trickett, E. J., Rukhotskiy, E., Jeong, A., Oberoi, A., Weinstein, T., & Delgado, Y.(2012). "The kids are terrific: It's the job that's tough": The ELL teacher role in an urban context. *Teaching and Teacher Education, 28,* 283–292. doi:10.1016/j.tate.2011.10.005

Tutunis, B. & Hacifazlioglu, O. (2018). The impact of reflective practices of English language teachers on the development of a sense of agency. *Journal of Education and Training Studies, 6* (10), 107-116.

Tyack, D. & Tobin, W. (1994). The "grammar" of schooling: Why has it been so hard to change? *American Educational Research Journal, 31,* 453–479. doi:10.2307/1163222

Varghese, M. M., & Snyder, R. (2018). Critically examining the agency and professional identity development of novice dual language teachers through figured worlds. *International Multilingual Research Journal, 12*(3), 145–159. Retrieved from http://proxy.lib.ohio-state.edu/login?url=http://search.ebscohost.com/login.aspx?direct=true &db=eric&AN=EJ1182676&site=ehost-live

Venegas-Weber, P. (2018). Teaching and knowing in "Nepantla": "I wanted them to realize that, that 'is' being bilingual." *International Multilingual Research Journal, 12*(3), 160-172.

Vitanova, G. (2018). "Just treat me as a teacher!" Mapping language teacher agency through gender, race, and professional discourses. *System, 79,* 28-37.

Wertsch, J. V. (1998). *Mind as Action.* Oxford, London: Oxford University Press.

White, C. J. (2018). Agency and emotion in narrative accounts of emergent conflict in an L2 classroom. *Applied Linguistics, 39*(4), 579–598.

Wills, J. S., & Sandholtz, J. H. (2009). Constrained professionalism: Dilemmas of teaching in the face of test-based accountability. *Teachers College Record, 111*(4), 1065-1114.

Wong, J. W., Athanases, S. Z., & Banes, L. C. (2017). Developing as an agentive bilingual teacher: Self-reflexive and student learning inquiry as teacher education resources. *International Journal of Bilingual Education and Bilingualism,* 1–17.doi:10.1080/13670050.2017.1345850

Yang, H. (2018). An analysis of the relationship between Chinese EFL teachers' agency and beliefs from an activity theory perspective. *Asian EFL Journal, 20*(3), 32-54. Retrieved from www.scopus.com

Yang, H., & Clark, M. (2018). Spaces of agency within contextual constraints: A case study of teacher's response to EFL reform in a Chinese university. *Asia Pacific Journal of Education, 38* (2), 187-201.

Yin, R. K. (2013). *Case study research: Design and methods.* Thousand Oaks, CA: Sage Publications.

Zhang, X. (2018). Mitigating suburban English writing teachers' constrained professional development through distance education: One case study. *The International Review of Research in Open and Distributed Learning, 19* (5), 238-254.

Appendix A
Focus of Studies and Research Questions

Sources	Focus of Study	Research Questions or Objectives
Babino & Stewart (2018)	Dual Language (DL) teachers enact agency among different tensions [1]as policymakers in the microsystems of their classrooms to truly accomplish DL programs' three-fold goals: bilingualism, biliteracy, and bicultural	RQ1: How do DL teachers perceive their agency as language policy makers in the classroom? RQ2: How do these same DL teachers exhibit a critical

	competence of students and advocate equity for minoritized students	consciousness by acknowledging the tensions and acting on that knowledge?
Christiansen, Du, Fang, & Hirvela (2018)	Three graduate teaching assistants' quest to achieve a desirable level of expertise in teaching second language writing became a contributing factor in their engagement with teacher agency	Within a professional learning community setting, what is the relationship between agency and expertise?
Colegrove & Zú ñ iga (2018)	A first-grade ESL teacher explores her agency (and that of her students) in implementing and experimenting with project-based instruction (PBI) in her economically, culturally, and linguistically diverse classroom.	How does a first-grade ESL teacher implementing PBI examine her teaching approach in a high-stakes testing environment? In what ways does a first-grade ESL teacher discover agency and that of her students while implementing PBI?
Feryok (2012)	An Armenian EFL teacher's early experiences and actions outside of the professional development classroom influenced her engagement with teacher agency constrained also by local, national, and international contexts	How does a language teacher develop a sense of agency?
Glas (2016)	Reports from 13 Chilean English language teachers' difficulties in motivating their students indicate the relevance between teachers' agency and learner motivation	RQ1:How motivated are students to learn English, according to their teachers' perception? RQ2: What are the reasons they mention to explain either the presence or the absence of learner motivation? RQ3:What contextual constraints and challenges with a potentially negative influence on learner motivation are mentioned by the teachers? RQ4: What internal constraints can be inferred from the teachers' accounts that impede their use of motivational strategies? RQ5: How do contextual and internal constraints interact? RQ6: What spaces for manoeuvre are perceived by

		the teachers that help create and maintain learner motivation in spite of contextual constraints?
Haneda & Sherman (2018)	Elementary ESL teachers' agentive action and redesigning of their work through job crafting to optimally support English learners	RQ1: How did ESL teachers act agentively through job crafting to bring their practice into greater alignment with their conceptua- lizations of what constitutes ESL teachers' work? RQ2: What factors allowed for or hindered their job crafting?
Hirver & Whitehead (2018)	Language teachers' co-construction of a sense of agency and professional identity through their classroom practice	What roles do the phenomenological manifestations of teacher agency in instructional practices play in the process of teacher identity formation?
Huyen Phan & Hamid (2017)	University English teachers' exercise of agency motivated by their sense of responsibility to their students and their academic well-being in micro-macro foreign language policy processes.	To understand how teachers (re)interpret and appropriate the concept of LA, how they empower learners in the EFL classroom through involving them in the decision-making process, creating opportunities for learners' self-reflection and optimising learner's target language use in the EFL class
Ilieva & Ravindran (2018)	International graduate students' co-construction of teacher identity and teacher agency when engaging with native speaker ideology and tensions in reconciling teaching goals and	To understand the material effects of one teacher education program on enactments of agency by program graduates

	professional contexts	
Ishihara, Carroll, Mahler, & Russo (2018)	Two former assistant language teachers' agency construction through drawing on linguistic and cultural resources from both English and Japanese	RQ1: How was teacher agency constructed, constrained, and (re)negotiated discursively in relation to their positioning by others in the local context? RQ2: How was the teachers' translingual practice related to the negotiation of their agency?
Kang (2017)	The complex bi-directional influence between teacher/learner agency and Language Play (LP) production which is related to teacher authority and inter-student power	RQ1: How do an elementary school NNEST and her EFL students construct LP in their classroom? RQ2: What are the underlying factors for such construction of LP?
Kayi-Aydar (2015b)	The influence of three pre-service classroom teachers' identity (re)negotiations on teachers' agency, interactions and classroom practice	How do teachers of ELLs position themselves and (re)negotiate identities in relation to their social context in their accounts of experiences and how do such positionings interact with their agency?
Leal & Crookes (2018)	A queer English language teacher's exercise of agency with marginalized identities for social justice and the analyses were through four aspects: "sense of purpose," "competence," "autonomy," "reflexivity"	RQ1: What is Jackson's "sense of purpose" (as defined by Pantic) when exercising agency for social justice in the classroom? RQ2: What are the conditions, as understood by Jackson (Pantic's "competence"), supportive of her exercising agency for social justice in the classroom? RQ3: How does Jackson exercise agency (Pantic's "autonomy") for social change in the classroom?

Li & De Costa (2017)	An EFL English teacher's negotiation of her professional identity in relation to the exercise and investment of her professional agency within the affordances and constraints of the given work context	RQ1: How did Ms. Q negotiate her teacher identity in relation to contextual affordances and constraints at her school? RQ2: How did Ms. Q's exercise of teacher agency affect the ways in which she negotiated her teacher identity?
Liyanage, Bartlett, Walker, & Guo (2015)	Inner Mongolian English language teachers' exercise of agency amidst the instructional demands of an exam-oriented community, and a misalignment created by an exam remaining centered on discrete skills rather than students' proficiency in language use within New English Syllabus expectations	To determine how teachers' professional practices are mediated by an examination-success-oriented mind-set of the public, curricular and policy directives, and teacher agency.
Mifsud & Vella (2018)	Two Maltese preschool teachers' agency and mediation of languages in their bilingual classrooms, which are influenced by the teachers' background and language beliefs, the sociolinguistic context (national and local), as well as the school language policies	To reflect on how two teachers of bilingual preschool classrooms in Malta were agentive in their language mediation strategies and to uncover the complex interplay of personal beliefs about language, classroom practices, and the individual needs of learners
Miller & Gkonou (2018)	Language teachers' agentive exercise of emotion and the role of emotion labor in producing emotional rewards in the teaching practice of English language teachers employed by tertiary-level institutions in the U.S. and U.K.	RQ1: What are the most common emotions experienced by tertiary-level English language teachers while teaching? RQ 2: How is teacher agency enabled and constrained in teachers' emotion labor? RQ 3: How does their exercise of agency, through emotion labor, lead to emotional rewards? RQ 4: How can teachers' reported emotions and emotion labor be understood from the perspective of ethical self-formation and teaching-as-caring?
Newcomer & Collier (2015)	Elementary teachers and teacher educators' exercise of agency in their interpretation and implementation of Arizona's model	To show how teachers exercise agency to counter some of the restrictions associated with Arizona's

	of Structured English Immersion (SEI)	specific SEI program, the 4-hour ELD model
Nguyen & Bui (2016)	Vietnamese English language teachers' agency in response to the national English Language Education Policy reforms at the local level	RQ1: What are teachers' attitudes towards the government-initiated English policies in Vietnam? RQ2: To what extent do the teachers possess the capacity for change in accordance with Fullan's (1993) theory on change agentry?
Ollerhead (2010)	Investigation of two adult ESL teachers' interpretation and response to Australia's Language, Literacy, and Numeracy Program (LLNP) policy through their diverging backgrounds and pedagogical and personal attitudes and beliefs	To examine specific policy-driven constraints and enablements experienced by each teacher in the course of her teaching work
Palmer, Henderson, Wall, Zúñiga, & Berthelsen (2016)	Two schools' teams of third grade teachers worked together to negotiate the intersection of DLBE implementation and high stakes accountability pressures	How do two teams of 3rd grade bilingual teachers negotiate the intersection of two-way dual language program implementation and high stakes standardized testing?
Ray (2009)	The characteristics of teacher agency at an academically successful Dual Language (DL) elementary school, including the identification of the factors that inform teachers' sense of agency (antecedents) and the instructional behaviors that result from that sense of agency (manifestations)	What are the characteristics of teacher agency at an academically successful dual language school?
Tao & Gao (2017)	The interaction of teacher agency and identity commitment of Chinese university L2 teachers that facilitates professional development during curricular reform	RQ1: How did teachers enact agency in facilitating professional development during curricular reform? RQ2: How did their identity commitment mediate teachers' enactment of agency to facilitate their professional development during curricular reform?

Tutunis & Hacifazlioglu (2018)	30 language teachers' reflective practices on the development of a sense of agency in disadvantaged neighborhoods in İstanbul, Turkey	RQ1: What are the initial challenges encountered by the English teachers in the first two years of their teaching to young learners in disadvantaged neighborhoods? RQ2: How do English teachers overcome the initial challenges? RQ3: What are the reflection experiences of English teachers before and after training?
Varghese & Snyder (2018)	Four pre-service teachers' development of professional identities and sense of agency as dual language teachers in the interactions with the teachers' personal linguistic,racial,and cultural backgrounds and external affordances, including their own language ideologies and those present in their contexts	To examine how four teachers in a mainstream elementary teacher education program (TEP) develop their sense of agency and figured worlds of dual language teaching
Venegas-Weber (2018)	The complexity of teachers' professional identity development and their possibilities for agency within nepantla, focusing on their negotiating of their linguistic and cultural identities as English-or Spanish-only teachers in a dual language program with a strict language separation model	RQ1: How do Chicanx/Latinx bilingual teachers' learning and development within nepantla shape their professional identity as bilingual and bi/multicultural teachers? RQ2: How does this professional identity interact with their agency?
Vitanova (2018)	ESL teachers'micro-aspects of emergent agency mediated by causative social factors, like gender, race, and culture	**Centrall question:** How do social factors such as gender and race mediate teachers' emergent agencies? **Other related questions** are: As teachers' agency emerges, how do these identity markers influence their relationships with others, for example, colleagues, students, or superiors? How does the past affect their future choicesdboth personal and professionaldas revealed through storytelling?
White (2018)	The interrelationships between emotion and agency from a dialogical perspective in multiple accounts of an incident of emergent	To examine the interrelationships between agency and emotion in teacher narrative accounts

	conflict in an L2 class for immigrants and refugees	
Wong, Athanases, & Banes (2017)	Through self-reflexive and student-learning inquiries to examine a bilingual teacher's teaching practices and agentive decision-making in a dual-language program	RQ: How, and to what degree, did one bilingual educator leverage self-reflexive and student-learning inquiry as resources for decision-making and teaching?
Yang (2018)	Teachers' agency in resolving the contradiction between their beliefs and their practices in regard to EFL reading instruction in a Chinese university from an activity theory perspective	To analyse the relationship between teachers' beliefs and teacher agency within a joint-activity system, including the activity systems of the teachers,their students and the departmen
Yang & Clark (2018).	EFL teachers' pedagogical agency in implementing College English curriculum reform from an activity theory perspective within and between the macro policy initiation level, the university implementation level, and the teacher's classroom level	To investigate teacher pedagogical agency in implementing College English curriculum reform in the wider Chinese context
Zhang (2018)	A Chinese suburban English writing teacher' s exercise of his agency while encountering multiple constraints in systemic functional linguistics-based distance education	RQ1: How does the teacher develop himself through SFL-based distance education? RQ2: How does the teacher conduct follow-up writing instruction on his own?

Appendix B
An Example of Coding and Data Analysis

Article	Theoretical frameworks	What is the Methodology?				Ethical issues	Methodological challenges
		Research design	Methods	Data collection	Data analysis		

| Ray (2009) | theory of emergent interactive agency, described in Bandura's (1989) social cognitive theory. Human agency means people change their situations or themselves through intentional actions determined by the interplay of behavior, internal personal factors, and external environment (Ray, 2009, p. 116). | qualitative instrumental case study approach | interview | "Data were collected through a series of semistructured interviews. Each participant was interviewed once, with interviews lasting from 1 to 3 hours" (p. 126). | "The clean data set was put into SuperHyperQual (Padilla, 2004) and analyzed using template analysis" (p. 120) | Not mentioned in the article | "A potential limitation of this style of analysis is the chance of missing themes that do not fit within the framework" (p. 126). "Although the sample size for this study is small, these findings suggest agentic factors that may be associated with high student achievement" (p. 135) "Another potential limitation has to do with the coding process itself. By removing fragments of text for use in the coding process, it is possible that the data might lose some of their meaning. This threat of decontextualization is countered by both the researcher's attention to detail and the need to amalgamate the data" (p. 136). |

About the Authors

Zhenjie Weng is a doctoral student at the Ohio State University in the Department of Teaching and Learning. Her research interests include second language writing, teacher identity, language teacher education, and World English(es).
Telephone: (724)-4679518; Email: weng.151@osu.edu; Address: 1945 N. High St.,138, Columbus OH 43210

Jingyi Zhu is a doctoral student at the Ohio State University in the Department of Teaching and Learning. She is currently enrolled in the Foreign, second, and Multilingual Language Education Program. Her research interests include second language reading, instructed second language acquisition, Children's/YA literature, and multicultural education.

Grace J. Y. Kim is a doctoral student at the Ohio State University in the Foreign, Second, and Multilingual Language Education program in the Department of Teaching and Learning. Her research interests include bilingual education, teacher education and curriculum design in language immersion programs, and heritage language education.

Multilingual, Middle-Eastern Students' Varied Responses to Directive and Non-directive Strategies in Peer Tutoring

Maria Eleftheriou'
American University of Sharjah, UAE

Abstract

This article presents the findings of a qualitative study examining the tutorial experiences of English as an additional language (EAL) tutees and their peer writing tutors in a Middle-Eastern university where the language of instruction is English. Data from stimulated recall activities, written observations, and interviews were analyzed to answer the following research questions:

1. Which type of tutoring approach (e.g. non-directive vs. directive) do EAL tutees find most effective?
2. Which type of tutoring approach (e.g. non-directive vs. directive) do tutors of EAL tutees find most effective?

The findings revealed that both tutors and tutees preferred the directive approach for lower order concerns (LOCs) and a non-directive approach for higher order concerns (HOCs). This study shows that diverse tutoring models that accommodate the background and experiences of Middle Eastern students, and their particular strengths and weaknesses, should be considered. Based on these findings, the author recommends tutorial training that emphasizes flexibility and recognizes the unique nature of each tutorial situation. These findings may signal a direction for the development of writing center pedagogy that focuses on the linguistically and culturally diverse students in the Middle East.

Key Words: Writing Center; EAL students; non-directive strategies; directive strategies; peer tutoring

Introduction

Writing center scholarship has developed in Western contexts and promotes a non-directive, collaborative approach, so writing centers in North America typically eschew directive approaches to peer tutoring. While a directive approach attempts to transfer the tutor's presumably superior language knowledge to the tutee through explicit explanations, a non-directive approach is intended to promote a sense of responsibility in students, teaching them to recognize and correct their own errors by eliciting knowledge rather than transmitting it. The orthodoxy of this non-directive peer tutoring perspective has been challenged in recent scholarship that analyzes tutor-tutee interaction to assess the effectiveness of different tutoring strategies. Tutors assisting tutees with English as an additional language (EAL) in writing centers have been found to use more directive approaches with their EAL tutees than with English as a first language (L1) tutees, because their EAL tutees require more language support and expect tutors to engage in directive instruction. The influence of the orthodox model extends to the Middle East, where a purely non-directive approach may be even less suited to address the particular challenges experienced by students working to develop their English writing skills. Students in the Middle East who seek help in English-language writing centers will typically be EAL, may not have had prior experience with the style of teaching involved in non-directive, collaborative tutorials, and they may not share the cultural assumptions implicit in the approach.

Little previous research has been done on English language writing centers in the Middle East, and most studies of EAL writers have been conducted in North American contexts. The present study thus seeks to examine tutors' and tutees' experiences of the effectiveness of directive and non-directive strategies of instruction with EAL students at MEU, a pseudonym for a university located in the Middle East. The findings of this study suggest that a mixed and context-sensitive approach to the use of directive and non-directive strategies would be most effective in this environment. These insights could eventually provide guidance for language instruction more

Email: meleftheriou@aus.edu, College of Arts and Sciences, Sharjah, UAE, 97165153353

generally in other multilingual or multicultural environments.

EAL Students in the Writing Center

In the mid-1990s, research on writing center pedagogy began to take into account the growing number of EAL students using English-language writing center services (Williams & Severino, 2004). Such research began to raise questions about the peer relationship and the effectiveness of a strictly non-directive approach for tutorials (Harris & Silva, 1993; Powers, 1995; Thonus, 2004; Williams & Severino, 2004). Some studies have indicated that EAL students might benefit from a more directive tutoring approach (Blau & Hall, 2002; Shamoon & Burns, 1995; Thonus, 2001, 2004; Williams & Severino, 2004), particularly with respect to lexical, syntactical errors (Mousse, 2013; Myers, 2003; Nakamaru, 2010; Nakatake, 2013; Rafoth, 2015; Weigle & Nelson, 2004). Studies focusing on EAL students have indicated that these students exhibit problems with morphology, lexical mastery, and syntax (Myers, 2003; Williams & Severino, 2004). They may lack the vocabulary to clearly articulate their ideas and may also require significant feedback on their word choice (Cogie, 2006; Minnet, 2009). Williams's (2005) research has indicated that EAL tutees and their English L1 tutors engage in more extensive diagnosis of the tutee's writing challenges (e.g. grammar, lexical issue) than is typically reported with English L1 tutees. Furthermore, compared with their English L1 peers, EAL tutees tend to elicit greater directive instruction from their tutors. The non-directive strategy of focusing on higher order concerns (HOCs) rather than lower-order concerns (LOCs) may be difficult to implement when students are still struggling with basic elements of a new language. Williams and Severino (2004) have noted that this focus on HOCs is intended to increase student agency and ownership of the text. Grammar and syntax improvement is assumed to occur naturally; improvements that do not come naturally are thought to be more easily learned once the student establishes the necessary connection to the text.

Though some scholars maintain that tutors should resist giving feedback on sentence-level concerns (Staben & Nordhaus, 2009), attitudes about error feedback and correction for EAL tutees have changed recently. The close link between rhetorical proficiency and linguistic proficiency in the development of EAL writing ability has been recognized (Blau & Hall, 2002; Myers, 2003; Nakamaru, 2010; Williams, 2004). Rafoth (2015) has claimed that adhering to a strictly non-directive approach "has opened writing centers to criticism for succumbing to the monolingual bias that treats errors the same way for native speakers, who can fall back on their intuitions about what sounds correct, and non-native speakers of English, who cannot" (p. 109). Blau and Hall (2002) have argued "that sentence-level errors can create global concerns, such as "errors in a student's thesis statement that make his or her central point confusing" (p. 36). Myers (2003) has recommended using a more traditional approach to tutoring EAL students, such as rephrasing students' sentences, inserting corrections into students' texts, and even offering practice exercises that target specific areas of weakness. She has argued that tutors must relinquish the attitude that giving EAL students the assistance they need amounts to appropriation and is therefore "unethical" (p. 66).

Despite concerns about tutor appropriation of student texts, most practitioners believe tutors can be good language resources for students without taking ownership of the text. Severino (2009) has argued that appropriation does not take place if tutors offer reasons for the changes they suggest to their tutees. Reid (1994) has also suggested that "intervening" by offering phrases and options to students to "provide [them] with adequate schemata (linguistic, content, contextual, and rhetorical)" (p. 286) is not appropriation; rather it "demystifies the writing process" (p.286). Pyle (2005) has suggested that tutors be given "more leeway to... give the L2 students... more robust feedback in the arena of expression [e.g. grammar and word choice], as well as training... in how best to do so" (as cited in Kastman Breuch & Clemens, 2009, p. 134). Myers (2003) has argued that combining rhetorical and linguistic feedback is essential and in keeping with the commonly accepted role of the tutor as cultural informant. There is a consensus among writing center researchers that some combination of non-directive and directive practices may provide the best solution to the variety of learning challenges faced by EAL students (Blau & Hall, 2002; Mousse, 2013; Williams, 2004; Williams & Severino, 2004).

While EAL students may lack the intuitions for correct English writing that L1 students can rely on, EAL students have another resource that many English L1 students lack: native mastery in a

language other than English. The native language that EAL students speak is often treated merely as an impediment to good English writing, as a source of interference errors. While it would be a mistake to deny or minimize the effect of interference errors, it is reductive to treat the native language of EAL students solely as a source of errors when it is also a potentially rich source of language knowledge upon which EAL students may draw in their English writing. The English language is highly adaptable, and its vocabulary and grammar are in constant flux. Its role as the lingua franca of the modern world results in an incredible diversity among it speakers, and the EAL speakers among them are a significant source for innovation. Accordingly, although it may be necessary to adopt a more directive approach with EAL tutees, non-directive strategies remain important for EAL tutorials, if students are to maintain agency over their texts, and if their non-English native languages are to be allowed to inform their writing in English. Horner, Lu, Royster and Trimbur (2011) have argued persuasively for a translingual approach to writing, one that addresses "how language norms are actually heterogenous, fluid and negotiable" (p. 305). They have proposed "a new paradigm that sees difference in language not as a problem to overcome or as a barrier to manage, but as a resource for producing meaning in writing, speaking, reading and listening" (p. 303). According to Canagarajah (2013), the translingual approach opens language pedagogy to the possibility of multiple languages producing "synergy, treating languages as always in contact and mutually influencing each other, with emergent meanings and grammars" (p. 41).

In the context of a writing center, within the translingual approach, "a tutor's job is no longer just about pointing out textual "divergences" from a singular notion of American academic English and then instructing a multilingual writer on how to "fix" that "mistake" (Olsen, 2013, para. 15). Although scholars advocating a translingual approach also call "for more, not less conscious and critical attention to how writers deploy diction, syntax, and style, as well as form, register and media" (Horner et al., 2011, p. 304), this attention will occur within a framework that does not view writing instruction as a search for errors to correct. Horner et al. are writing within a North American context, but this insight may be especially relevant for MEU, where the language of instruction is English, and tutors and tutees are EAL speakers. Adopting insights from the translingual approach, tutors could be trained to anticipate the possibility that the other languages spoken by EAL speakers could be an asset, a source of novel ways of expressing meanings in English.

Possible Challenges Facing Arabic L1 Tutees in the Writing Center

Researchers have suggested that some Arabic L1 students do not respond positively to non-directive teaching methods because the pre-university education of Arabic L1 students emphasizes passive learning and rote memorization (Gardiner-Hyland, 2014; Martin, 2006). Although the UAE government and UAE educational institutions have attempted to encourage active learning methods, teacher-centered, behaviorist methods with a focus on rote learning and memorization are still prevalent in many of the school systems in the country (Hall, 2011; Nunn & Langille, 2016; Sperrazza & Raddawi, 2016). Martin (2006) has pointed out that teachers in the UAE are seldom challenged, and students are not expected to be active participants in their learning. Mynard (as cited in Martin, 2006) has noted that students from Arab societies may not challenge rules or take initiative or risks in learning situations because they fear shame. Secondary school teachers may be trained in constructivist methods, but confronted by student resistance to these methods in their classrooms, they fall back to the traditional behaviorist methods that were used during their schooling (Gardiner-Hyland, 2014). The highly hierarchical and directive methods of teaching English experienced by most Arabic L1 speakers in their pre-university education is coupled with a related traditionalist approach to language learning that focuses on grammar instruction as the main route to language fluency. While an increased focus on grammar for Arabic L1 students is advocated by some scholars (Al-Buainani, 2006; Al-Jamhour, 2001; Al-Khasawneh, 2010), other researchers argue that a focus on grammar instruction and outdated approaches to English language instruction are to blame for Arabic L1 students' low English proficiency levels. Al Hamzi (2006) has observed that a commitment to sentence level construction and grammar consideration is one of the central drawbacks of EAL writing instruction for Arabic L1 students. A narrow focus on grammar is often "in the service of the institution and at the expense of multilingual writers" (Olson, 2011, para. 9).

In the UAE, the majority of universities are Western and mostly American, and at the core of their liberal arts curriculum is critical thinking (Hall, 2011; Sperraza & Raddawi, 2016). Incoming students are required to complete courses in academic writing intended to develop critical thinking

skills. However, coming from primary and secondary school systems that follow traditionalist teaching methods, these students are often ill-prepared to meet the expectations of their professors because they "are not used to being in charge of their own learning and text creation, and they struggle to cope with the demands of critical thinking and independent learning" (Hall, 2011, p. 430). For Arabic L1 students who have been conditioned to believe that progress in language learning comes from learning rules of grammar and correcting errors in their application, the non-directive and collaborative process employed in the writing center may seem useless or even obstructive.

Another challenge for the non-directive methods used in the writing center may result from Arabic diglossia (Hall, 2011; Rivard, 2006). The written form of Arabic that students learn in school is the classical Arabic of the Qur'an, which is quite different from the variety of spoken dialects of Arabic that function as the vernacular. Written mastery of Arabic is quite distinct from the ability to speak clearly or even eloquently in demotic Arabic. Consequently, Arab L1 students may lack writing mastery in their native language, which will in turn make it difficult "for them to achieve competency in second language (L2) writing" (Hall, 2011, p. 428). If the issue of Arabic diglossia is considered together with the insight of Rafoth (2015) that EAL students may have problems with non-directive methods because of their inability to draw on intuitions about what sounds correct, we can see that the problem Rafoth has noted could be compounded in the case of Arabic L1 students, who may be unaccustomed, even in their native language, to relying on speech-based intuitions when writing.

Research Questions

The purpose of this study was to understand tutors' and tutees' experiences with the use of directive and non-directive strategies in peer tutoring sessions with EAL tutees and tutors in a Middle Eastern university where the language of instruction is English. Given the specific nature of writing center work, as well as the wide range of potential factors that may influence tutors' and tutees' experiences, an interpretive approach was deemed appropriate to address the research questions. This approach attempts to identify how individuals construct meaning through their experiences, and it allows for examination of interaction authentically (Thonus, 2002) and within its natural environment (Kim, 2003; Rowlands, 2005). The following research questions were addressed:

1. Which type of tutoring approach (e.g. non-directive vs. directive) do EAL tutees find most effective?
2. Which type of tutoring approach (e.g. non-directive vs. directive) do tutors of EAL tutees find most effective?

Methodology

Research Context and Participants

MEU, located in the United Arab Emirates, enrolls thousands of international students annually. The university curriculum is modeled on the curriculum of American universities, and it has received American accreditation. The population for this study is drawn from a multilingual, multicultural, and polydialectical context. While most students enrolled at MEU are Arabic speakers, they form a significantly different mix from Arabic speakers in other Middle Eastern countries: they attend a university where the language of instruction is English, many of them speak more than two languages and use English as their language of commerce. Students at MEU, including tutors and tutees, have high levels of what Byram (1997) describes as "intercultural communicative competence and intercultural awareness" (p. 3). Tutors are students themselves and share this common ground with their tutees. A tutorial session may include tutors and tutees with various nationalities and ethnicities; for instance, an Iranian tutor could be working with an Indian tutee, an Afghani tutor with a Sudanese tutee, or a Syrian tutor and a Lebanese tutee.

Despite the diversity of the students' language backgrounds, the writing they produce for their courses at MEU is expected to conform to, and is graded on the basis of proficiency in, Standard American English (SAE). All students at the university are required to take 15 credits of writing courses as part of their major programs, and many other courses have assessment components that require writing, as part of a university wide effort to include writing across the curriculum. Students with lower levels of initial English proficiency face serious hurdles, not merely

in comprehending course content in English classrooms with English textbooks, but in conveying what they have learned in papers, tests and other types of assessment that are expected to conform to SAE. The effect of this disparity in English proficiency upon student academic performance is a serious concern for faculty members and students. The faculty of the Writing Studies department, responsible for the teaching of writing to students in all disciplines, is aware of the cultural and linguistic challenges that EAL students face, particularly when they first arrive at MEU. The instructors use a variety of approaches to negotiate the institutional requirement for SAE. For example, instructors may offer individualized attention to students who are not able to meet the expectations for SAE in their classrooms. Additionally, in courses that are not focused specifically on academic writing, instructors may tolerate or even encourage regional or translated idiomatic expressions, and syntactic and organizational structures that may sound odd to the Western ear, but that allow students to express their points. The ability of Writing Studies instructors to adopt a broader translingual approach that would endorse the legitimacy of hybrid forms of English is limited, however, by the fact that MEU has no major in Writing Studies, and therefore, the primary function of Writing Studies courses is to prepare students for the writing requirements in the courses of their chosen discipline, where SAE is expected.

Following the American writing center model, both in the use of the term "peer tutoring" and in the organization of the Writing Center, the MEU Writing Center offers one-on-one tutoring sessions by appointment or on a drop-in basis to all students throughout the university. Most of the clients at the Writing Center are EAL students who are seeking help with their writing assignments. Instructors teaching writing-intensive courses often encourage or even require their students to visit the Writing Center for supplemental help. Students can also self-refer. Peer tutors are recruited from undergraduate composition courses at the recommendation of their instructors, who attempt to identify the most competent and confident English writers. Frequently, tutors are multilingual, and occasionally, they are even unable to identify their native language. They are employed for 6-15 hours a week and paid 30 dirhams an hour. They have fulfilled the requirements for tutoring at the Center: a credit-bearing course on peer tutoring and writing. The peer tutoring course is designed to help students become familiar with and to think critically about writing and peer tutoring issues and to develop a practical approach to peer tutoring in writing. Tutors are required to participate in in-service training activities throughout the academic year to discuss issues and concerns that arise during tutorial sessions.

Participants

Tutee and tutor participants were solicited at the beginning of the academic semester. Participants were selected based on the time they arrived for their tutorials. Tutees who arrived earliest for their appointments were recruited first. Fifteen tutees and 15 tutors participated in the study. The 15 tutees were seeking assistance with writing assignments for their writing composition courses. Both these tutees and the tutors who were paired with them were asked to participate in the study.

Tutees

For the purposes of this study, the stimulated recall interviews and the interviews of 15 tutees and 15 tutors were separately examined. Nine of the tutees were female; 10 tutees were freshman, two were juniors, two were sophomores, and one was a senior. Fourteen of the tutees were Arabs, and one student was Pakistani. Most tutees spoke only Arabic at home (13), eight were taught in both Arabic and English in high school, and six were taught only in Arabic. Twelve tutees spoke Arabic as a first language. Tutees had a variety of academic majors (i.e., electrical engineering, mass communication, mechanical engineering, civil engineering, international studies, design management, chemical engineering, computer science, marketing and management, and architecture).

Table 1
Means and Standard Deviations of Tutees' English Fluency Ratings

	N	Mean	*SD*	Min	Max
Written Score	15	2.27	0.88	1	4
Oral Score	15	2.47	1.13	1	4

Table 1 shows the tutees' fluency in oral and written English. Tutees were asked to rate their written and oral language skills on a scale ranging from 1 = Strong to 5 = Weak. The average tutee fluency for both written and oral English was rated as moderate. The average and standard deviation for English fluency represents a typical sample of Writing Center clientele in a given semester.

Tutors

Most tutors were female (13) and seven were 21 years of age. As with the tutees, the tutors had a range of academic majors (i.e. mass communication, mechanical engineering, chemical engineering, marketing and management, international studies, architecture, finance and management, visual communication, as well as English literature). Tutors had a variety of ethnic backgrounds, and they spoke a variety of languages (e.g., Arabic, English, French, Bengali, Farsi, Konkani, Tamil, Pashtu, Malayalam).

Table 2
Means and Standard Deviations of Tutors' English Fluency Ratings

	N	Mean	*SD*	Min	Max
Written Score	15	1.07	0.26	1	2
Oral Score	15	1.13	0.35	1	2

Tutors were also asked to rate their English fluency on a scale ranging from 1 = Strong to 5 = Weak. Although the languages spoken at home were diverse, the average tutor English fluency in both written and oral was rated as strong. The average and standard deviation for tutors' English fluency represents a typical sample of Writing Center staff in a given semester (see Table 2).

Tutees are referred to as E1 through E15, while tutors are referred to as U1 through U15, in each case according to the order in which the tutorials took place.

Data Collection

The study took place over a six-month period. In an effort to enhance the validity of my findings, I employed three methods for gathering data: observation, stimulated recall, and interviews.

Data collection methods	Timeline	Purpose
Observation	Concurrent with tutorial session	Noted elements of tutorial session Identified tutorial strategies
Stimulated recall	Within 24 hours of tutorial session	Prompted recall of thoughts during tutorial sessions
Interviews	Within 72 hours of stimulated recall	Asked questions to gather perceptions of tutorial effectiveness

Observation

Tutorials were video recorded. Video recordings of tutorials were then watched in order to (1) take extensive notes on all elements of the tutorial (e.g., communication and dynamics between tutors and tutees, effectiveness of tutorial strategies, reactions of tutees, overall effectiveness of tutorials, and any other relevant factors) and (2) identify strategies later discussed by tutors and tutees during their stimulated recall activity. To address the possibility of researcher bias, a colleague who taught the tutor training course watched the videos with me. We then compared notes and discussed our observations, finding no significant discrepancies.

Stimulated Recall

As recommended by previous researchers (e.g., Gass & Mackey, 2000), stimulated recall using video recordings was conducted immediately after tutorial sessions for tutees and within 24 hours for tutors. During stimulated recall sessions, participants were asked to pause the video when they wanted to comment about a particular segment. The researcher also paused the video to ask tutors/tutees what they were thinking or feeling during certain interactions of interest (e.g. if a tutee seemed confused, annoyed, frustrated, satisfied, or pleased), or if the video had not been stopped for some time.

Semi-Structured Interviews

I conducted semi-structured interviews with all participants within 72 hours of the stimulated recall. Interviews were based on a prepared set of questions asking tutees and tutors which tutoring strategies they found effective or ineffective; however, I allowed emerging circumstances to guide the interview process, asking for clarification or additional information and encouraging participants to elaborate on their answers. The interviews were recorded and transcribed.

Data Analysis

I adopted a theory driven approach in order to examine the discourse about the non-directive and directive strategies. Once collected, I examined the data in order to identify references to non-directive and directive strategies. When the type of strategy being referred to in the stimulated recall was unclear, my research assistant, a tutor with two years' experience in the Writing Center, watched sections of the tutorials with me to help identify the directive/non-directive moves used in tutor talk. Strategies such as negotiating the tutorial agenda with tutees, discussing writing with tutees, and encouraging tutees to think critically and independently were considered non-directive. Strategies such as telling tutees how to rephrase a sentence, taking notes for tutees, and identifying and correcting tutees' errors were considered directive. The research assistant and I had minor disagreements about the labeling of a number of the tutoring moves, but these were typically when a combination of directive and non-directive strategies were employed. In these cases, comments were labeled "non-directive and directive." When there was insufficient information from the comment or the videotaped tutorial to label the strategy being referred to as "directive or non-directive," the strategy was left unlabeled.

Once the non-directive and directive labels had been added to the transcriptions, an independent rater and I began the coding. At the time, I was an advanced doctoral student in TESOL and Applied Linguistics and had over 10 years of experience working in and directing writing centers. The rater had an MA in Education and experience with coding and analysis, but she did not have experience with writing centers. Guidelines offered by Hycner (1985) on analysing data phenomenologically were followed. First, the rater and I looked for patterns in the use of non-directive and directive strategies from both tutors and tutees. Having completed the process of identifying patterns, the next step was to combine and catalogue related patterns into sub-themes. In addition to identifying patterns in the data, I calculated the frequency of themes and subthemes, non-directive and directive comments, and positive and negative comments. I then compared the notes I had taken while watching the tutorials with the themes/subthemes and frequencies that were generated through the study.

Limitations

My status as Director of the Writing Center meant that I was responsible for hiring and supervising tutors and overseeing daily operations of the Center. As Lerner (2002) has noted, such insider status may bias results but does not necessarily invalidate assessments by writing center directors. Following his recommendations, I attempted to bring a sense of neutrality to the process of participant observation and to balance the roles of insider and outsider. I took measures to reduce anxiety and ensure authentic responses (making participation voluntary, assuring tutors that their responses would not impact their positions). Furthermore, the tutors are accustomed to my observing their tutorials and discussing them with me as part of their work. They indicated that my observation did not make them nervous during sessions or throughout the study, that they understood the study's objectives, and that our discussions were based on trust and mutual respect. I also used triangulation to acquire data to test the veracity of

my interpretations and consultations with colleagues unaffiliated with the Writing Center to double-check my conclusions.

Results and Discussion

The stimulated recall (SR) and the interviews (I) of 15 tutees and 15 tutors were analyzed separately. The themes that emerged from the data were used to answer the research questions related to directive and non-directive tutoring strategies. Themes related to the more general tutoring experience will be addressed in a future article.

Tutees' Perceptions of Non-Directive and Directive Approaches

Fourteen tutees mentioned non-directive strategies as being effective 43 out of 55 times, and directive strategies as being effective 35 out of 55 times. The interview data yielded slightly different results, but the number of differences was too low to be considered significant. Patterns emerged that show that tutees found each approach effective for specific concerns and in particular contexts.

Tutees favored a directive approach when discussing LOCs: they reported that they appreciated clarity and directness when their tutors addressed issues such as verb tenses, punctuation, and word choice. E2 paused the recording as her tutor explained how to identify and correct a run-on sentence to say:

> She made it very simple, she was direct. She wasn't trying to give me another example. No, she just sticks to the material, and she said 'this is what's wrong, and this is how you fix it'...and that was good.
>
> (SR, E2)

Some tutees mentioned not having the knowledge necessary to identify and correct their grammatical errors and therefore appreciated a more directive approach. In the video of Tutorial 15, the tutor pointed to the word "discriminate" and instructed the tutee to use "discriminate against." The tutee paused the tape at this instance to say:

> I was not English educated, and sometimes, I feel I have some problems in my writing like I don't know the correct way of using a certain word, so... I really appreciated how she told it and explained it to me.
>
> (SR, E15)

These reports are consistent with findings with EAL tutees in North American contexts (Blau & Hall, 2002; Mousse, 2013; Myers, 2003; Rafoth, 2015; Thonus, 1999, 2001; Williams, 2004) who point out that EAL tutees may indeed require the directive approach of an informed tutor. Mastery of LOCs is critical to both reading and writing texts, and while English L1 learners often have some operative understanding of these functions, EAL learners must learn these elements before proceeding to fuller text comprehension (Williams & Severino, 2004). Williams (2004) has observed that deciding whether to use a directive or non-directive approach with EAL students is "not a yes/no question. There is much that no amount of questioning, indirect or otherwise, could ever elicit from these writers because there is so much that they simply do not understand about their L2 and academic writing" (p. 195). Williams explained that tutor attempts to use non-directive strategies with their EAL tutees sometimes resulted in "almost absurdly circuitous interactions, in which the writer engaged in a sort of guessing game" (p. 195). In Tutorial 15, U15 said she attempted to elicit the correct answer from her tutee, but that it led to "a guessing game," so she decided to switch to a directive approach. Her tutee said that she appreciated the more directive approach because her high school education had left her unfamiliar with idiomatic expressions or nuanced meanings.

Although most tutees preferred a directive approach when dealing with LOCs, some tutees, even those with low English proficiency, also complained about this approach, especially if it entailed line-editing or corrections without instruction. In Tutorial 4, the tutor asked the tutee to rewrite his thesis statement, providing the tutee with the words to rephrase his sentence. E4 expressed his dissatisfaction with this method by pausing the recording during this section of the tutorial to say:

> I believe that the main goal or aim of the Writing Center is to help students be better in writing, try to find their mistakes...without any tutor...But here...she's just... telling me to write down the sentence, and this is not useful.
>
> (SR, E4)

Tutees frequently requested additional explanations so that they could understand where they had

erred and how they could improve. In North American contexts, it has been observed that EAL tutees may require more rules, explanations, and illustrations than English L1 students to make sense of the language (Mousse, 2013; Rafoth, 2015; Williams & Severino, 2004). Perhaps our tutees, who are predominantly EAL learners, sometimes require a similar style of instruction. Drawing on Shamoon and Burns's research (1995), Myers (2003) has claimed that the directive approach is most valuable when it includes rhetorical strategies such as modeling and imitation. In the recordings, tutors can be seen employing these types of traditional directive strategies: explaining rules relating to the particular errors while correcting them, illustrating how errors can be avoided and fixed in other writing contexts, and modeling different ways of writing sentences. Tutees were particularly responsive to situations when the tutor used examples to illustrate their errors. When discussing comma usage, for instance, U3 not only mentioned the grammatical rule for comma placement, she wrote a sentence on a notepad and explained where one would place the punctuation. She then asked the tutee to address the comma issue in his writing, and he was able to identify the position of the misplaced comma. The tutee identified this strategy as effective, thereby illustrating Myer's (2003) claim that "showing is better than telling" (p. 66).

It was observed in many instances that tutees seemed to want their tutors to be authoritative, confident, and knowledgeable, but they still wanted to retain a level of control over the session. In Tutorial 12, the tutor used a directive strategy by pointing to the word "expressions" and recommended that her tutee change it to "words and actions." It is clear from the video that E12 seemed hesitant to make the change: she was silent, paused for a few seconds before finally scratching out her own word to make the suggested change. E12 said during her stimulated recall that she was not convinced and did not understand the tutor's reason for this change: "I didn't like that she told me to change it. I thought it just made the sentence longer." It is interesting to note that tutees with higher English proficiency, such as E12 above, rejected tutors' suggestions and questioned the tutors' advice more often than tutees with lower English proficiency. E5, the other tutee who rated her English proficiency as strong and who was instructed in English in high school, felt confident rejecting her tutor's suggestions as well. Throughout the tutorial, the tutor and tutee had a relatively egalitarian relationship: it was a highly interactive tutorial with both tutor and tutee taking turns leading the tutorial session. The tutor suggested that the tutee combine two of her ideas into one paragraph rather than discuss the ideas in two separate paragraphs. E5 did not challenge her tutor's suggestion during the tutorial, but during the stimulated recall, she paused the tape to tell the researcher she had ignored her tutor's suggestion because "essay writing is all about opinions."

Both the stimulated recall and interviews revealed that there were more instances of the lower level EAL students accepting the tutor's advice in instances where they did not completely agree with the recommendations. In Tutorial 6, the tutor recommends that the tutee change the word "religion" in his text to "Islam" to make it more specific. During the stimulated recall, Tutee E6 stated that he did not agree with his tutor's suggestion: "I not want like to mention the name of the religion. Want to just to write religion, but she told me to write Islam. I write Islam but not completely agree. She know correct." The tutor's suggestion, and the tutee's unfortunate belief that the suggestion was "correct," appears to have distorted, rather than revealed, the tutee's intention. The tutor was not in a productive dialogue with the tutee and failed to recognize the importance of word choice (LOC) and its relation to the tutee's argument (HOC). Had the tutor explored the word choice with the tutee, rather than defaulting to the assumption that the tutee had made a mistake, had she considered "what the [writer] [was] doing with language and why" (Horner et al., 2011, p. 305), the student would not have had his meaning distorted.

Instances when tutees expressed discomfort with the degree to which their tutors altered their texts, or complained they did not understand the rationale for these changes, illustrate the risk to student agency in the use of the directive approach by tutors. Instances of compromised agency can be seen primarily in situations where tutors suggest modifications without explanation. The translingual approach may provide insights that would help tutors to avoid employing the directive approach in ways that compromise the writer's agency. This approach stresses the importance of "discussing with multilingual writers the various reasons behind a question or suggestion about language use" (Olsen, 2013, para. 15). With a translingual approach, writing tutors do not focus on removing errors from tutee writing; rather, they help "multilingual writers draw from their different discourses and make active decisions about utilizing various features from them" (Olsen, 2015, para. 12). Canagarajah (2006) has noted that "not every instance of nonstandard usage by a student is an unwitting error; sometimes it is an active choice motivated by important cultural and ideological considerations" (p. 609), and that to assume otherwise denies multilingual students agency. This

observation would have been valuable guidance to the tutor in Tutorial 6 above.

While tutees tended to find the directive approach more effective in addressing LOCs, both sets of data revealed that they found a non-directive approach most effective when addressing HOCs such as structure, organization, argument, and coherence. Strategies such as negotiating an agenda, asking questions about the assignment, asking tutees for clarification on their intended meaning, outlining with tutees, and asking them to write independently were deemed effective by most of the tutees, even those with low English proficiency. Outlining was mentioned by several tutees as being effective. In Tutorial 2, the tutor told the tutee that she had serious problems with coherence in her essay. The tutor can be seen in the video writing numbers and headings and drawing boxes on a sheet of paper and then asking the tutee to write her idea in each box. Her tutee, E2, paused the tape during this activity to say: "I may look grumpy, but I am really happy. I liked her method, I said so, and it was making sense and boxes really help." The non-directive strategy of asking tutees to explain or clarify their ideas orally in order to help them rephrase their sentences was similarly effective. Tutors can be seen asking their tutees questions such as "Can you explain what you mean here?" Most tutees responded favorably to this non-directive strategy. E3 paused the tape as his tutor asked him "What do you mean in this sentence?" to say that he appreciated this strategy because his tutor could not know his intended meaning without asking. Interestingly, tutees with lower English proficiency also reported satisfaction with non-directive strategies. In one instance, U1 asks her tutee to create a thesis statement, but he waits with pen in hand for the tutor to provide him with the words to write the thesis. The tutor does not comply and instead gives the tutee a notepad to write the sentence on his own. During the stimulated recall, the tutee said:

> I was really comfortable and happy at the same time. She asked me to write something, and then I asked her to do it, and she said, "No you can do it." I really liked that. (SR, U1)

These results are surprising in one respect: most of our writing center clientele have experienced a traditional, authoritative style of schooling (Gardiner-Hyland, 2014; Martin, 2006; Richardson, 2004). Discovery and critical thinking, important elements in the non-directive approach, are not generally encouraged. Although our tutees have been accustomed to directive approaches, they seem to respond well to non-directive approaches and to the possibility of gaining a higher level of agency and responsibility.

While the non-directive strategies mentioned above were considered effective by most tutees, there were some exceptions with tutees at all proficiency levels. In Tutorial 8, the tutor asked her tutee, a student with moderate English proficiency, to clarify the meaning of her topic sentence and handed her a pad of paper to rephrase the sentence on her own. The tutee paused the recording during this portion of the tutorial to say:

> Okay, this part, it was stressful. When I sit at home, I'm relaxed, the words come up. But then, I felt that I looked stupid in front of her. (SR, E8)

Similarly, several other tutees (E2, E11) reported finding their tutor's questions "stressful," "confusing," or "unclear," or indicated that they felt "lost" when the tutor used a non-directive approach, such as trying to engage in a discussion about the logic of their arguments. Interestingly, for first-time users of the Writing Center like E8, E2 and E11 above, opinions of non-directive strategies seemed to change as their tutorials progressed. While their comments were more negative at the beginning of the stimulated recall sessions, they became more positive toward the end of these sessions and during their interviews. They became more adept at handling the tutor's non-directive methods and even realized the value of these techniques. For example, during the stimulated recall, E11 first reported that he did not like being asked about the problems in his writing but then later said:

> Over here, when he told me ... "what the wrong here?"... He gave me a chance to try to... to know what my mistake and not to tell me what my mistake, so it's help me.

While studies conducted in North American contexts with EAL tutees indicate that they expect directive approaches and perceive their tutors as representatives of the academic institutions (Blau & Hall, 2002; Thonus, 2001, 2004), the findings of this study suggest that our tutees prefer non-directive approaches under certain circumstances, perhaps because, as Ronesi (2009) has observed, MEU students are "multicultural and multilingual, and often multidialectal" (p. 77). The tutees may respond well to non-directive approaches because they are more adaptable, having had years to develop high levels of intercultural communicative competence. Although these tutors and tutees come from families, educational institutions and societies that are largely hierarchical and patriachical, they respond well to the egalitarian nature of peer tutoring relationships. Not only do they have a tacit understanding of each other, they appreciate the opportunity to engage in

discussions about their writing where they are able to generate insights and ideas rather than just to receive them. Their ability to "decentre" (Byram, 1997, p. 3) and understand how messages will be perceived in another cultural context has been tested and developed in ways that may not be characteristic of many EAL students in the US. The students at MEU, living as they do in a multilingual community, may enjoy the benefit that Canagarajah (2007) has described from "the constant interaction between language groups" and how different languages can "overlap, interpenetrate, and mesh in fascinating ways" (p. 930).

All tutees responded positively when their tutors used a combination of directive and non-directive strategies. Several tutees mentioned that a switch from a non-directive to directive approach worked well because they did not have the necessary knowledge to work independently. In Tutorial 8, the tutor noted that the tutee had not included topic sentences in her paragraphs. She started by asking her tutee to summarize and write down the main idea for each paragraph, but her tutee responded negatively to this request. At this point, U8 changed her approach, explaining the function of topic sentences and suggesting various methods of phrasing one. E8 paused the video at this point to say:

> I liked this way more than the other one when she asked me to come up with the points. I liked it when she suggested the points that I should write, and I expressed it in my own words. This way I learned topic sentences.

Similarly, E6 stated in both his stimulated recall and interview that he appreciated how his tutor asked him to identify and correct his own errors first but then became more directive when it was clear he did not have sufficient knowledge to self-correct. In his interview, E6 stated that he found his tutor's strategy effective: "She told me read sentence first, then after I finished, she asked me if that's right. Find mistake yourself first, then if I can't, she will say it. This is good." Although tutees appreciated having a role in the discussion of their writing, they acknowledged the value of having tutors assume more authoritative roles as language informants.

Tutors' Perceptions of Non-Directive and Directive Approaches

The tutors felt positively about both non-directive and directive strategies; however, there was a stronger preference for non-directive strategies. The preference for the non-directive approach was apparent in both sets of data. Most of the tutors who answered the interview question "Which strategies, if any, did you find ineffective?" felt that the directive strategies they used were the least effective. For example, U12 said the "directive approach just put [her tutee] off" and that she "should have been more facilitative because this is someone that clearly would have been very co-operative with [her]" (I).

Tutors were more critical of a directive approach in the interview data than in the stimulated recall session with 17 out of 22 comments on the directive approach being negative; nevertheless, both sets of data show that different tutors displayed consistency in their preference for a combination of directive and non-directive approaches. Like the tutees, tutors favored a directive approach when discussing LOCs, particularly in cases where rules were not obvious and a native-like fluency was required. U4 reported that she used the directive approach when correcting her tutee's prepositions: "It becomes very difficult to explain to the tutees when to use a certain preposition...so, you tend to just say 'it's just like this,' and you have to remember it" (SR). U15 can be seen using the directive strategy frequently when addressing LOCs throughout her tutorial. She spoke about her choice to use directive strategies in response to her tutee's incorrect use of a preposition:

> I just felt that she didn't know it, that she couldn't know it, you know? I tried to get her to figure it out, but she was not going to, and when I found out that she was not...., I just gave it to her. (SR)

It is interesting to note that many tutor/tutee pairs commented on the same instances when discussing the effectiveness of the directive approach. For example, in one pair, both tutor and tutee (U15 and E15) commented favorably on the tutor's use of the directive approach when explaining the error in "discriminate people." The tutee said that she was not familiar with the expression "discriminate against" and appreciated her tutor's explanation. The tutor said she felt that a directive move such as identifying the error and correcting it for her tutee was appropriate and effective in this situation. Consistent with these findings, Blau & Hall (2002) have proposed "a rethinking of conventional tutoring strategies" (p. 29) for EAL tutees and have suggested that tutors should feel comfortable using a directive approach when addressing LOCs. Similarly, Rafoth (2015) has maintained that students who are struggling with low proficiency may need a tutor who can provide

answers to language questions.

It is worth noting that some tutors mentioned avoiding the directive approach even though they felt such an approach may have been necessary. For example, one tutor (U9) reported that she was unsure how much content-based information she could provide for her tutee because she did not want to overstep boundaries. Another tutor (U1) also worried that she was being "too directive with her tutee and found herself not helping him with vocabulary because she "didn't want to introduce too many new words into his essay" (SR). Consistent with these tutors' concerns, Severino (2009) has argued that, while the directive approach can provide tutees with relevant information, it can also compromise tutees' agency – especially in situations where tutors suggest modifications without explanation. It is especially problematic in this context because tutees with a low to mid-level English proficiency followed their tutors' advice unwaveringly, even when they were skeptical about the recommendations. The tutors' concern about text appropriation is reasonable and has been noted in the literature (Bringhurst, 2006; McHarg, 2013). Ethical dilemmas about whether or not to intervene directly or how much to intervene are complex. However, as Bringhurst (2006) has observed, a non-directive approach with struggling writers can present another ethical concern: maintaining a strictly non-directive approach may be depriving tutees of the knowledge-based guidance they require in order to gain new skills.

Like the tutees, most tutors favored a non-directive approach, especially when addressing HOCs. They reported that non-directive strategies such as asking tutees for clarification, outlining, mapping, and asking their tutees to write independently were more effective for addressing issues such as unity and coherence of essay assignments. U9 said that asking for clarification was an effective strategy with her tutee. In the video, U4 can be seen turning the tutee's paper away and asking "what are you trying to say here?" During her stimulated recall, she explained why this non-directive strategy was effective:

> There are some writers who... have a lot of ideas and they made the connections in their heads, and they know exactly where they're going, but they don't write that down on paper. So as a reader, I was really, really confused... But when she explained it to me, I understood: ... So I just asked her to sort of like make the connections herself and put them down on paper. And...at this point I felt she did understand.(SR, U2)

Furthermore, tutors found that their tutees were most animated and receptive when they used non-directive strategies and could engage their tutees in discussion and the writing process. This is consistent with the findings from the tutees' SR and interviews. As mentioned earlier, this finding does not correspond entirely with studies that have been conducted in North American contexts, which report that EAL students elicit more directive strategies from their tutors (Blau & Hall, 2002; Thonus, 2001, 2004), and it is surprising when considering that most students who participated in the study have been through more traditional directive-style schooling. MEU students may be involved in language studies in ways that differ from EAL learners in American universities. Being multilingual, they may be able to relate English grammatical structures to their formal education in other languages. Tutors are in dialogue with tutees, drawing upon structures from the tutees' backgrounds to highlight aspects of English, relating idiomatic expressions and turns of phrase to SAE, and extending their understanding of writing as a process through which writers collaborate with readers in order to allow meaning to emerge.

Although most tutors reported finding non-directive strategies effective with HOCs, there were some notable exceptions. Several tutors claimed that, despite a tutee's English proficiency, their lack of knowledge sometimes prevented the use of a non-directive approach. For example, in the video, U2 can be seen trying to use non-directive strategies at the beginning of the tutorial to help her tutee, whose English proficiency is high, organize her ideas; however, she changed her approach when she noticed her tutee was not being receptive. As noted above, it is interesting that, in several such cases, though tutors felt the non-directive strategies were not successful at the beginning of the tutorials, they felt more positively about them by the end. Tutors attributed this change to their tutees' unfamiliarity with the writing center approach. U2 described her tutee's reaction:

> Initially, she was sort of sensitive. I was new to her, and she was new to the writing center concept, and she really didn't know how to react or...what to expect. She didn't even know how to react to criticism, so initially she was defensive, but then she warmed up and then she became more comfortable to asking questions and being more interactive too.(I, U2)

Both sets of data revealed that all the tutors who used a combination of directive and non-directive strategies found this combination effective. Fourteen tutors mentioned non-directive strategies 72 times: 45 comments were positive, 25 were negative, and two were neutral. Fourteen

tutors mentioned directive strategies 50 times: 22 comments were positive, and 15 were negative. Tutors who believed a combination of approaches worked best said that they often started the session with the goal of using non-directive strategies but changed their approach if they noticed their tutees were not receptive to it. Tutors reported that their strategies were most successful when they took into consideration their tutees' personalities and English proficiency. U6 started off using a non-directive approach but felt that her tutee's proficiency was too low to respond to this approach:

> I found that there were a lot of times when he wouldn't really understand my question, and he would nod, but then I kind of understood that he's really lost. So then I would be a little directive and try to tell him, "OK, do this," and this worked much better.(I, U6)

U3 said she started off the tutorial using a directive approach because her tutee's spoken English was difficult to understand; however, she switched to a non-directive approach when she noticed that "his written English was better than his spoken English" (SR). She went on to say that this approach suited her tutee much better. Tutorial 8 is another instance where a non-directive approach was ineffective with HOCs. The tutee, a first-time user of the Writing Center, initially responded less positively to the non-directive approach, leading the tutor to change her style and become more directive. As the tutee became more relaxed, however, the tutor reintroduced some of the non-directive strategies toward the end of the tutorial, and the tutee's perception of the non-directive approach became more positive.

As Jane Cogie (2001) has said, "fostering student authority is not a matter of following a single approach and avoiding another" (p. 47). Tutors should use their discretion, which will improve with experience, to determine the appropriate approach for each tutee, each assignment and during each phase of the tutorial. Blau and Hall (2002) have cautioned against treating all EAL students in the same way and advise tutors to consider students' individual differences by assessing each situation as it arises. Likewise, Carino (2003) has advised tutors to try to anticipate when to focus on global issues such as content, when to pay closer attention to LOCs such as grammar, or when to spend time on both. He has argued that tutors need to prepare to deal with both "interpersonal and intertextual" features of tutorials and advises tutors to avoid "all-too tempting sort of rules of thumb" that can lead to "prescriptive dictums that can unintentionally cement a strained social relationship between tutor and tutee" (p. 113). It is not surprising then that tutorials deemed most successful by tutors and tutees were those in which tutors were most self-reflective and self-critical during their stimulated recall and interviews. They demonstrated an ability to take into account tutees' level of ability, their personalities, and their responses to different choices and decisions. These tutors were attuned to the changing needs and levels of confidence of their tutees and navigated between the directive and non-directive approaches as they interacted with their tutees. They were involved in a genuine dialogue with the students, engaging in a shared attempt to find appropriate English expression for meanings that emerged from a confluence of English and another language, or even other languages.

Conclusion and Implications

Although traditional writing center doctrine has insisted that tutors adopt collaborative and non-directive approaches in tutorials, recent literature has explored the potential for more directive strategies when working with EAL students (Bringhurst, 2006; Thonus, 2001, 2002, 2004; Williams, 2004, 2005;). Studies conducted in North American contexts (Blau & Hall, 2002; Rafoth, 2015; Thonus, 2001, 2004; Williams, 2004, 2005) have indicated that EAL tutees prefer authoritative tutors who use directive approaches, and an emerging consensus suggests that a combination of non-directive and directive practices may be more appropriate in addressing the needs of EAL students (Blau & Hall, 2002; Carino, 2003; Henning, 2001; Mousse, 2013; Rafoth, 2015). This study's findings are in line with the emerging consensus, in that they demonstrate the value of both approaches in different situations. In contrast to the studies of EAL tutees in North American contexts, however, this Middle Eastern study indicates that while tutors and tutees find directive approaches useful for LOCs, both tutors and tutees prefer non-directive approaches when addressing HOCs. Thus, while directive approaches should be recognized as valuable tools, tutors should not abandon collaborative and non-directive strategies, even when dealing with low proficiency EAL students. The participants in this study may be more responsive to the non-directive approach than EAL students in a North American context because both tutors and tutees are typically EAL students, making it more likely that they interact as genuine peers. Additionally, their knowledge of multiple languages allows them to see languages other than English

as facilitating rather than obstructing effective communication in English, suggesting that insights derived from a translingual orientation can inform the teaching of writing at MEU, and writing centers and writing classrooms in the region. It is incumbent upon writing centers to continue to challenge the orthodoxy of dominant theory and to keep asking themselves, "how can we serve writing students better?" The results of this study indicate the dynamic nature of writing center research and the need to be wary of fixed theories and categories that are not responsive to the changing needs of tutors and tutees. This study reveals the importance of recognizing the differences not only between EAL tutees here and EAL tutees in North American contexts, but also among different populations of EAL tutees within our Writing Center to better meet the individual tutee's particular needs.

The implications of this study may extend beyond writing centers in the region to English language classrooms. Using an appropriate mix of directive and non-directive strategies, in a manner that can flexibly adjust to the particular needs of individual cohorts, could result in a better fit between instruction methods and student needs, potentially improving outcomes in student English fluency. Additionally, treating students as knowledgeable individuals with valuable skills in other languages, rather than merely as failed English speakers, could bolster student confidence in a context where fear of making errors can make students excessively risk-averse and therefore reluctant to express themselves in English.

A more general implication of this study is that a student's learning history does not determine the student's learning future. Students who had been educated in a highly directive, authoritarian system have proven to be quite capable of appreciating, preferring, and even demanding non-directive teaching strategies once they have been exposed to them. Teaching English to EAL students, whatever their backgrounds, should proceed as a dialogue with the student, one in which the student is seen as an active participant, contributing to and setting the direction for the ongoing pursuit of clear English writing.

Recommendations

The findings of this study suggest that writing centers could benefit from tailoring tutor training and writing center policy to suit their clientele.

The following recommendations are based on the findings of this study.

- Writing centers should employ flexible tutoring models that accommodate the experiences of tutors and tutees and their particular strengths and weaknesses.

- Tutors should be open to the use of the directive approach, especially when addressing LOCs, and with students with low English proficiency. However, tutors should use the approach with caution, being careful to avoid appropriation. Activities such as role playing, stimulated recall or close vertical transcriptions could help tutors to determine when to use the directive approach, and when to stick to non-directive strategies.

- Tutors should be trained in the use of directive strategies such as modeling, asking leading or closed questions, offering suggestions, and identifying and correcting errors. These strategies should be described in detail, and situations where such strategies would be appropriate should be explained. Tutor training should include participation in mock tutoring sessions that simulate difficult tutorial situations, followed by sessions of stimulated recall with the writing center director, in order to enable tutors to develop instincts to guide their use of directive approaches.

- Tutors should be cautioned to be careful when advising lower level EAL students. Lower-level students often accept recommendations they do not agree with or understand because they see their tutors as authorities.

- First time users of the writing center may not be familiar with the non-directive approach and may be more resistant to it than repeat users. The study showed that tutees sometimes warmed to the non-directive approach even later within the first tutorial. If tutors are made aware of this situation, it may help them gain and retain the confidence of new users, without foreclosing opportunities to deploy the non-directive strategies that both tutors and tutees report as the most helpful, especially when dealing with HOCs.

References

Al- Buainani, H. A. (2006, March). *Students' writing in EFL: Towards a teaching methodology.* Paper presented at the Second International AUC OXF Conference on Language and Linguistics. Cairo, Egypt. Retrieved from http://www1.aucegypt.edu/conferences/aucoxf/

Al-Hazmi, S. (2006). Writing reflection: Perceptions of Arab EFL learners. *South Asian Language Review, XVI* (2), 36-52. Retrieved from http://www.asian-efl-journal.com/site_map.php

Al-Jamhour, A. (2001). A cross-cultural analysis of written discourse of Arabic-speaking learners of English. *Journal of King Saud University (Language and Translation), 13*(1), 25-44. Retrieved from http://digital.library.ksu.edu.sa/V13M163R431.pdf

Al-Khasawneh, F. (2010). Writing for academic purposes: Problems faced by postgraduate students of the college of business, UUM. *ESP World, 9*(2), 28. Retrieved from http://www.esp-world.info

Blau, S., & Hall, J. (2002). Guilt-free tutoring: Rethinking how we tutor non-native-speaking students. *Writing Center Journal, 23*(1), 23-44. Retrieved from http://www.english.udel.edu/wcj/

Bringhurst, D. (2006). Identifying our ethical responsibility: A criterion-based approach. In C. Murphy & B. Stay (Eds.), *The writing center director's resource book* (pp. 165-178). Englewood, New Jersey: Lawrence Erlbaum Associates.

Brooks, J. (1991). Minimalist tutoring: Making the student do all the work. *The Writing Lab Newsletter, 15*(6), 1-4. Retrieved from http://www.writinglabnewsletter.org/new/

Byram, M. (1997). T*eaching and assessing intercultural communicative competence.* Clevedon, NJ: Multilingual Matters.

Canagarajah, S. (2006). *The place of World Englishes in composition: pluralization continued. College Composition and Communication, 57*(4), 586-619. Retrieved from https://www.jstor.org/stable/20456910?seq=1#page_scan_tab_contents

Canagarajah, S. (2007). Lingua franca English, multilingual communities, and language acquisition. *Modern Language Journal, 91*, 923–939. doi:10.1111/j.0026-7902.2007.00678.x

Canagarajah, A. S. (2013). Negotiating translingual literacy: An enactment. *Research in the Teaching of English, 48*(1), 40-67. Retrieved from https://www.jstor.org/stable/24398646

Carino, P. (2003). Power and authority in peer tutoring. In M.A. Pemberton & J. Kinkead (Eds.), *The center will hold: Critical perspectives on writing center scholarship* (pp.96-113). Logan, UT: Utah Tate UP

Cogie, J. (2006). ESL student participation in writing center sessions. *The Writing Center Journal 26*(2), pp. 48–66. Retrieved from http://www.english.udel.edu/wcj/

Gardiner-Hyland, F. (2014). Exploring the impact of teacher education pedagogy on EFL reading teacher identities: A United Arab Emirates Case. In Bailey, K. & Damerow, R (Eds.), *Teaching and learning English in the Arabic-speaking world (pp.83-100).* London: Routledge.

Gass, S., & Mackey, A. (2000). *Stimulated recall methodology in second language research.* Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Grimm, N. (1996). The regulatory role of the writing center: Coming to terms with a loss of innocence. *The Writing Center Journal, 17*, 5-29. Retrieved from http://www.english.udel.edu/wcj/

Hall,   K. (2011). *Teaching composition and rhetoric* to *Arab EFL learners.* In C. Gitskaki (Ed.), *Teaching and learning in the Arab world* (pp. 422–440). New York: Peter Lang Publishing.

Henning, T. (2001, March). Theoretical models of tutor talk: How practical are they? Paper presented at the annual meeting of the Conference on College Composition and Communication, Denver, CO. Retrieved from ERIC database. (ED451569)

Horner, B., Lu, M., Royster, J., & Trimbur, J. (2011). Language difference in writing: Toward a Translingual approach. *College English, 3*(73), 303-321.

Hycner, R. H. (1985). Some guidelines for the phenomenological analysis of interview data. *Human Studies, 8*, 279–303. doi:10.1007/BF00142995

Kastman Breuch, L., & Clemens, L. (2009). Tutoring ESL students in online hybrid (synchronous and asynchronous) writing centers. In S. Bruce & B. Rafoth (Eds.), *ESL writers: A guide for writing center tutors* (2nd ed.; pp. 132-148). Portsmouth, NH: Boynton/Cook Publishers.

Kim, S. (2003). Research paradigms in organizational learning and performance: Competing modes

of inquiry. *Information Technology, Learning, and Performance Journal, 21*(1), 9-18. Retrieved from http://www.osra.org/itlpj/kimspring2003.pdf

Martin, J. (2006). Online Information literacy in an Arabian context. *Learning and Teaching in Higher Education: Gulf Perspectives, 3*(2). Retrieved from http//www.zu.ac.ae/lthe03_03_03_martin.htm

McHarg, M. (2013). A sociocultural exploration of English faculty perceptions of the Writing Center in the Qatari context. *Arab World English Journal. 4*(4). Retrieved from http://www.awej.org/index.php?option=com_content&view=article&id=386:molly-mcharg&catid=44&Itemid=133

Minett, A. (2009). Earth aches by midnight: Helping ESL writers clarify their intended meaning. In S. Bruce & B. Rafoth (Eds.), *ESL writers: A guide for writing center tutors* (2nd ed.; pp. 66–77). Portsmouth, NH: Boynton/Cook Publishers.

Moussu, L. (2013). Let's Talk! ESL Students' Needs and Writing Centre Philosophy. *TESL Canada Journal, 30*(2), 55-68.

Myers, S. A. (2003). Reassessing the proofreading trap: ESL tutoring and writing instruction. *The Writing Center Journal, 24*(1), 51–67. Retrieved from http://www.english.udel.edu/wcj/

Nakamaru, S. (2010). Lexical Issues in Writing Center Tutorials with International and US-Educated Multilingual Writers. *Journal of Second Language Writing, 19*(2), 95–113. http://dx.doi.org/10.1016/j.jslw.2010.01.001

Nakatake, M. (2013). Challenges and possibilities in tutorials in a writing center in Japan. *The Language Teacher 37*(6). Retrieved from http://jalt-publications.org/tlt/articles/3458-challenges-and-possibilities-tutorials-writing-cente r-japan

Nunn, R., & Langille, J. (2016). Operationalizing a global and holistic characterization of competence in a local context in current trends in language testing in the Pacific Rim and the Middle East: Policies, analysis, and diagnosis. In A. Aryadoust. & J. Fox (Eds.), *Cambridge Scholars* (pp. 300-314). UK: Cambridge Scholars Press.

Olson, B. (2013). Rethinking our work with multilingual writers: The ethics and responsibility of language teaching in the writing center. *Praxis: A Writing Center Journal. 10*(2). Retrieved from http://www.praxisuwc.com/journal-page-102

Rafoth, B. (2015). *Multilingual writers and writing centers.* Boulder: Utah State University Press.

Reid, J. (1994). Responding to ESL students' texts: The myths of appropriation. *TESOL Quarterly, 28,* 273-292. doi:10.2307/3587434

Richardson, P. (2004). Possible influences of Arabic-Islamic culture on the reflective practices proposed for an education degree at the Higher Colleges of Technology in the United Arab Emirates. *International Journal of Educational Development, 24*(4), 429-436. doi:10.1016/j.ijedudev.2004.02.003

Rivard, J. N. (2006). *An investigation into diglossia, literacy and tertiary-level EFL classes in the Arabian Gulf states.* Unpublished Master's thesis, McGill University, Montreal, Canada.

Ronesi, L. (2009). Multilingual tutors supporting multilingual peers: A peer-tutor training course in the Arabian Gulf. *Writing Center Journal, 29*(2). Retrieved from http://www.english.udel.edu/wcj/

Rowlands, B. (2005). Grounded in practice: Using interpretive research to build theory. *The Electronic Journal of Business Research Methodology, 3*(1), 81-92. Retrieved from http://www.ejbrm.com

Severino, C. (2009). Avoiding appropriation. In S. Bruce & B. Rafoth (Eds.), *ESL writers: A guide for writing center tutors* (2nd ed.; pp. 51–65). Portsmouth, NH: Boynton/Cook Publishers.

Shamoon, L., & Burns, D. (1995). A critique of pure tutoring. *Writing Center Journal, 15,* 134-151. Retrieved from http://www.english.udel.edu/wcj/

Sperrazza, L., & Raddawi, R. (2016). Academic writing in the UAE: Transforming critical thought in the EFL classroom. In A. Ahmed & H. Abouabdelkader (Eds.), *Teaching EFL writing in the 21st century Arab world* (pp. 157-188). Hampshire: Palgrave Macmillan. doi: 10.1057/978-1-137-46726-3_7 157-188

Staben, J., & Nordhaus, K. (2009). Looking at the whole text. In S. Bruce & B. Rafoth (Eds.), *ESL writers: A guide for writing center tutors* (2nd ed.; pp. 78–90). Portsmouth, NH: Boynton/Cook Publishers.

Weigle, S. C., & Nelson, G. L. (2004). Novice tutors and their ESL tutees: Three case studies of tutor roles and perceptions of tutorial success. *Journal of Second Language Writing, 13*(3), 203–225. doi:10.1016/j.jslw.2004.04.011

Thonus, T. (2001). Triangulation in the writing center: Tutor, tutee, and instructor perceptions of the tutor's role. *Writing Center Journal, 22,* 59-81. Retrieved from http://www.english.udel.edu/wcj/

Thonus, T. (2002). Tutor and student assessments of academic writing tutorials: What is success? *Assessing Writing, 8,* 110-134. doi:10.1016/S1075-2935(03)00002-3

Thonus, T. (2004). What are the differences? Tutor interactions with first- and second-language writers. *Journal of Second Language Writing, 13,* 227-242. doi:10.1016/j.jslw.2004.04.012

Williams, J. (2004). Tutoring and revision: Second language writers in the writing center. *Journal of Second Language Writing, 13,* 173-201. doi:10.1016/j.jslw.2004.04.009

Williams, J. (2005). Writing center interaction: Institutional discourse and the role of peer tutors. In B. H. K. Bardovi-Harlig (Ed.), *Institutional talk and interlanguage pragmatics research* (pp. 37-65). Mahwah, NJ: Erlbaum.

Williams, J., & Severino, C. (2004). The writing center and second language writers. *Journal of Second Language Writing, 13,* 165-172. doi:10.1016/j.jslw.2004.04.010

About the Author

Maria Eleftheriou is an Assistant Professor at the American University of Sharjah. In addition to directing the AUS Writing Center, Maria researches the relevance of established writing center models in contemporary Middle-Eastern contexts. Her research interests include tutor training models, writing center assessment, on-line writing instruction and writing assessment.

Differential Allocation of Attention to Meaning and Form in Reading Comprehension for Monolingual and Bilingual Learners of English

Mohammad Nabi Karimi
Ebrahim Zangani⁵
Nahid Fallah
Kharazmi University, Tehran, Iran

Abstract

In the light of evidence that attention can facilitate and enhance learning, this study attempts to investigate the allocation of attention to meaning and form simultaneously in reading comprehension in the foreign language among monolingual and bilingual learners of English. To this end, three groups of learners namely, bilingual (knowing Kurdish, Persian), monolingual (knowing Persian) learners of English and mixed (equal number of bilingual and monolingual learners) as the control group were selected based on the results of a language proficiency test ((MTELP). The learners in the two experimental groups were asked to read a written text for meaning and circle a designated lexical form. The learners in the control group were only required to read the text for meaning and answer the follow-up questions. To examine whether the type of attentional condition and bilinguality, as the two independent variables of the study, have any significant effect on comprehension scores, a two-way ANOVA was run. Think aloud technique was also employed to elicit the learners' targeted thought processes from the groups which had to circle the targeted lexical form. The results indicated that the experimental group who paid simultaneous attention to form and meaning and processed the targeted item for both form and meaning gained better comprehension scores regardless of the number of languages they knew. Moreover, it was shown that deeper levels of processing are associated with better comprehension ability. Therefore, we may conclude focus on the lexical form might improve comprehension as indicated by comprehension scores and may be an effective way to make texts more comprehensible. **Key Words**: Attention, monolinguals, bilinguals, meaning, form, reading comprehension

Introduction

Majority of the studies that have been directed at uncovering language learning processes operate under the assumption that language learners are monolinguals learning a second language (Kramsch, 2012), yet with the advent of globalization and the expansion of mobility and communication technology, a considerable number of language learners tend to be bilingual or even multilingual. Therefore, examining and exploring how learning a third language is different from learning a second and what processes are involved in merit further investigations. This endeavor should be undertaken to shed some light on how third language learners learn and may have implications as to how they can be helped in their journey to become trilingual.

It is a rather established fact that prior knowledge influences our understanding and interpretation of a text in a new language (Ellis, 2006). In this regard, Ellis proposes the concept of "learned attention" to capture the idea of L1 affecting L2 processing and adopts McWhinny's (1987) competition model to illustrate this point. He explains how second language learners selectively attend to some parts of a sentence and ignore the redundant non-salient features. In an argument similar to VanPatten's (2004) primacy of meaning principle, he posits that language learners attend to content words rather than forms of low cue validity. He states that content words overshadow non-salient linguistic features. What could be of interest for further research is that maybe learners with different L1 backgrounds attend to different parts of a sentence. Moreover, it is still not clear how being a third language learner and knowing more than one language would influence processing at different levels of comprehension.

Any theory of sentence comprehension should account for "(i) the representations, which are encodings in memory, (ii) what information is extracted from incoming words, and (iii) how that information is used to combine the incoming" (Malko, Ehrenhofer, & Phillips, 2016, p.1). The majority of current research indicates the mixed storage and high interconnectivity of language

systems in the mind of a multilingual (Szubko-Sitarek, 2015). Learners, therefore, are highly likely to make use of all the linguistic resources available to them to make sense of a particular sentence. However, it is still unclear whether they use L1 or L2 or both to process the meaning of a sentence in a third language. In other words, what linguistic resources do they exploit to comprehend the meaning of a sentence in another language? Do they reconstruct meaning in L1, L2, or both?

So far most of the studies on input processing have been done in second language acquisition and it is not clear whether knowing a third language would be an asset in attending to meaning and form simultaneously. Furthermore, as some studies have demonstrated low proficiency language learners often resort to their L1 and their knowledge of the world to interpret and understand the meaning of sentences and texts, thus, it is likely that multilingual learners use both L1 and L2 and maybe even outperform second language learners due to knowing two languages. No previous studies, to our knowledge, have been undertaken to compare the performance of low proficiency third language learners in allocating attention to meaning.

Attention has long been associated with processing and subsequent learning (Schmitd, 1990). The role of attention is central in theoretical and applied linguistics. In theory, most, if not all, theatrical approaches to language learning posit a role for attention whether they regard it as the detection or noticing (Schmidt, 2001; Tomlin & Villa, 1994).

Recently, Robinson (2017) while emphasizing the interrelation between attention and awareness distinguished different levels of these two concepts. Two levels of attention, namely perceptual attention (attending to different issues automatically and unconsciously) and focal attention (attending to issues consciously) come into the fore in language learning issues. When learning occurs without attention, it simply means that there is no focal attention to input, just choosing some parts of data for more processing in memory. To justify such kind of learning, Robinson maintains that at the perceptual processing stage, rudimentary detection of input before selection takes place which assists the learner to learn. As such he concludes that this learning usually occurs without awareness as awareness is a prerequisite to focal attention. Similarly, awareness has also different levels ranging from noticing surface structures to understanding rules and regularities.

Review of the Related Literature

It is claimed that attention facilitates and enhances learning (Baars & Gage, 2010; Logan, 2005) since it induces learners' attention to intended linguistic forms and leads to noticing. Robinson (1995) defined noticing as "detection plus rehearsal in short-term memory, prior to encoding in long-term memory" (p. 296). He claimed that only detected input is focally attended to and noticed. While detected input goes into short-term memory, focally attended input goes into working memory. Schmidt (2001) emphasized the importance of attention in all types of learning both conscious and unconscious and believed that little learning can occur without attention. Tomlin and Villa (1994), Truscott (1998) and Carroll (2006) believed that attention and input and not necessarily awareness contribute to learning.

Attention in the field of SLA has received substantial research interest and many applied linguists have paid much heed to it. Schmidt (2001) attributed a pivotal role to attention in every aspect of the SLA process. He believed that attention helps us to understand L2 development, variation, fluency, individual differences and the role of instruction. Learners may attend to form or meaning or both when processing input (Greenslade, Bouden, & Sanz, 1999; VanPatten, 1990; Leow, Hsieh & Moreno, 2008). However, Norris and Ortega (2000) believe that attention to form and meaning simultaneously is more efficacious than either alone. VanPatten (2004) in his "primacy of meaning principle" postulated that "learners process input for meaning before they process it for form" (p.14). VanPattan (1990) investigated learners' attention to meaning and form simultaneously in aural input. He used control and experimental groups to assess the learners' attentional resources. Low proficiency learners in the experimental group were required to listen to the passage and identify the occurrences of the three target L2 forms, namely inflació n, la and – n. The results indicated that low proficiency learners cannot process both form and meaning in aural L2 mode. Attention to grammatical items negatively influenced learners' comprehension compared to the control group but attending to lexical items did not. Greenslade et al. (1999) replicated VanPattn's study in the written mode. In this study, the learners were required to circle the same target forms in the passage. He obtained similar results as VanPattn's. Learners who attended to a lexical form understood the passage as well as control group but comprehension was somehow blocked when they attended to grammatical form. Wong (2001) surveyed simultaneous attention to form and

meaning in both aural and written modes. This was another replication of VanPattn's study. Wong used the English translation of the Spanish text that VanPattn used. He found that in the aural mode, results corroborated VanPattn's study whereas in the written mode the same level of comprehension was achieved by all groups. Wong propounded that learners' attentional resources may be varied in the written and oral modes. Leow et al. (2008) followed the same line of inquiry but with some modifications in the methods used like the use of think-alouds to elicit learners' targeted thought processes, the use of multiple choice questions to check learners' comprehension of the text, and the use of new grammatical forms. In the written mode, no differences were found between the experimental and control groups. They believed that differing cognitive constraints for processing different modalities account for varied sets of results. They also found that target forms may be processed at different levels by different learners. Morgan-Short, Heil, Botero-Moriarty, and Ebert (2012) studied simultaneous attention to form and meaning and the reactivity of think-aloud protocols in the written mode. Their results showed that learners' attention to lexical or grammatical forms did not influence the comprehension of the reading passage. They also found that thinking aloud had a reactive effect when reading a passage for meaning and concurrently paying attention to form. Finally, it was found that learners with deep processing of forms experienced greater understanding.

Tomlin and Villa (1994) discussed three subparts of attention, namely alertness, orientation, and detection. Alertness is concerned with interest and motivation. Orientation is associated with input-flooding and attention to form. They refer to detection as "cognitive registration of stimuli" which is crucial for learning and processing. LaBerge (1995) stated that attention has the potential to facilitate or increase processing. Accordingly, selected materials require more attention (Neill, Valdes & Terry, 1995) and ignored information receive little attention (Van der Heijden, 1981). It has also been found that language components require different attentional resources. In other words, the way that learners allocate attention and awareness to learning vocabulary and morphology is different from learning syntax (VanPatten, 1994; Schwartz, 1993). Meanwhile, attention should be subject to a particular learning domain, that is, it should be concentrated and it must not just be global (Schmidt, 2001). Put it differently, to learn phonology, attention must be focused on the L2 sounds and to learn vocabulary, the learner should attend both to word forms and contextual clues.

Similarly, Robinson, Mackey, Gass, and Schmidt (2012) maintained that when processing grammatical gender, learners pay more attention to morphophonological cues (noun endings) rather than syntactic cues. Lew-Williams (2009) found that learners pay less attention to syntactic cues for gender like agreement on adjectives and determiners. Mackey, Gass, and McDonough (2000) elicited learners' perceptions of oral feedback using retrospective interviews in order to survey learners' attention to different parts of the language. They focused on learners' production of the target language and examined their attention on specific aspects of language especially deviated forms. The role of attention was also investigated by Gass, Svetics, and Lemelin (2003) on the learning of three aspects of language, namely lexicon, syntax, and morphosyntax. By manipulating attention, they found that syntax received the largest amount of attention while without focused attention, lexicon drew the learners' attention. They also highlighted the role of proficiency in lowering the effect of directed attention.

In the same fashion, Armengol and Cots (2009), firstly investigated the nature and objects of attention in two university students and secondly, they surveyed the relationship between the attention processes and the final written products. These students were multilingual and underwent think-aloud protocols while engaging in writing an essay in two languages (Spanish and English) other than their first language (Catalan). The researchers found that the participants of the study made use of their multilingual resources in creating a text in a specific language. With regard to attention episodes as indicated in think-aloud protocols, it was found that the subjects focused on procedure-related and language-related issues separately. The objects of attention in procedure-related awareness episodes were content, text structure and cohesion, rhetoric, and writer's block while in language-related awareness episodes were grammar, spelling, sentence cohesion and structure and word choice. With regard to the relation between the attention processes and the final written products, the researchers observed that explicit and implicit nature of awareness episodes may be of importance and needs attention, as one participant exhibited good implicit knowledge whereas the other showed good explicit knowledge about writing.

Godfroid and Uggen (2013) investigated German beginning second language learners' attention to irregular verb morphology during sentence processing by means of eye-tracking techniques. He found that learners paid more attention to stem-changing verbs than those verbs

which were regular, demonstrating the delayed effect of irregularity of verbs on reading times. Godfroid, Boers, and Housen (2013) wanted to know whether more attention contributes to more learning. Specifically, they aimed to address whether L2 learners devote more attention to unknown words when reading for pleasure. In other words, they assessed the role of attention in incidental vocabulary learning in the second language using the eye-tracking technique. The results revealed that the subjects of the study allocated more time for processing the unknown words than familiar words. The results of their study were in line with Ellis' (2002) and Rayner's (2009) findings in which low-frequency words required more processing time than high-frequency words.

Finally, Dolgunsöz (2015) measured learners' attention while reading L2 text and learning gains by means of eye-tracking technique. The results indicated that learners spent less time on familiar words than unfamiliar words. A positive correlation was also found between attention and learning gains. He also discussed merits and drawbacks of eye tracking methodology compared to other techniques of measuring attention like note-taking, underlining and verbal protocols. Eye tracking is the robust method of gathering attentional data without suffering from reactivity and memory decay. However, there are some technical challenges facing researchers regarding the use of this technique which makes it demanding for them.

Although there have been few investigations regarding how learners allocate attention to meaning, the paucity of studies concerning whether multilingual learners differ in allocating attention to meaning and form from bilingual learners was a significant incentive in conducting this study. Accordingly, this study intends to investigate whether multilingual learners utilize linguistic resources of first, second or third language when attending to meaning. In other words, whether knowing a third language (being multilingual) will be an asset in attending to specific aspects of meaning compared to bilingual or monolingual learners. This research specifically aims to focus on the following research questions:

1. Does simultaneous attention to form and meaning of lexical items in the written text have a significant effect on comprehension?
2. Does the allocation of attention to meaning affect comprehension differently for bilinguals and monolinguals?
3. Is there a statistically significant effect on reading comprehension due to the interaction effect of bilinguality and attentional condition?
4. Do different levels of processing lexical items make a difference in comprehension?

Methodology

Participants

The study utilized a convenient sampling method and the participants were recruited from among the low intermediate level students based on their English proficiency score at university entrance exam majoring in Mechanical engineering at Kashan University, Iran. The subjects were male and female, 26 monolingual and 28 bilingual. The learners' first language was Farsi or Kurdish. The Kurdish language is usually spoken in Northwestern Iran (Kurdistan province). To achieve the purpose of the study, three intact classes of third-semester English course with students' and instructors' consent were selected. To ensure the exact proficiency of the students in English, a language proficiency test was administered among the participating students. To save time, a reduced form of Michigan Test of English Language Proficiency (MTELP) was applied. Those with one standard deviation above or below the mean based on the result of MTELP were selected. Then, the subjects were divided into two groups based on whether they were monolingual or bilingual. An equal number of subjects (10) from each group was selected to be assigned to the third group, that is, control group. The first group constituted bilingual learners of English (25 participants). These learners speak Kurdish as their first language, Farsi as their second language at school and community and English as their third language. The second group consisted of monolingual learners whose first language was Persian and they learn English as a foreign language (25). The final number of participants for the monolinguals was 18 and for the bilingual group was

16. The third group which consisted of the equal number of monolingual and bilingual learners (10 bilinguals and 10 monolinguals) acted as a baseline to compare the performance of students in the first and second groups (20).

Instrumentation

The materials for this study were adopted from Wong (2001). Wong herself adopted the text from

VanPatten (1990) which was originally written in Spanish for listening and reading comprehension. This study utilized multiple choice questions to check the learners' comprehension of the text as they were used in Leow et al. (2008) and Morgan-short et al. (2012). Meanwhile, one content word was chosen for the students to attend to in the reading comprehension task. That lexical item was the word "*inflation*" as it was used in VanPattan (1990) and Wong (2001). The verbal protocol was also used to elicit learners' targeted thought processes at the end of the test. The type of technique applied in this study was think- aloud protocol.

Procedure

As this study attempted to probe simultaneous attention to meaning and form among monolingual and bilingual learners of English, three groups of monolingual, bilingual and mixed as control group were selected. Learners first received instructions regarding how to perform their tasks. In the control group, the learners were asked to read the text for meaning and answer the follow-up questions. Students in monolingual and bilingual groups were told to read the passage for meaning too and circle the word "inflation" whenever they encountered in the text and then answer the comprehension questions. Reading comprehension questions were designed in Persian since it was believed that the goal is to check the overall comprehension of the text and not the questions. The subjects were asked to complete their tasks in ten minutes. Then, the participants of the study (except the control group) were asked to verbalize what was going through their mind when performing the tasks. As the control group was not required to pay attention to form while reading for meaning, they were not asked to verbalize their thought. For further analyses, the learners' reports were audio-taped by researchers or recorded by the participants themselves using their cell phones. Students' recordings then were emailed to the researchers and transcribed and coded along with researchers' audio-taped files. The obtained results of each group were analyzed so as to shed light on the attentional resources they employed in comprehending the meaning of the sentences. To be included in analyses, the subjects were required to detect the minimum of at least 60% of the target item, as it was also highlighted in previous studies. Subjects received one point for every correct answer to multiple choice questions and zero points otherwise. To operationalize attention, the lexical items circled or mentioned in think-alouds were coded as instances of attention. Think-aloud protocols were coded by the researchers. If learners asserted that they went back to read the text in order to answer the questions, the researchers eliminated them from the subjects' pool.

Results

To address the first and the second research questions investigating whether the type of attentional condition and bilinguality, as the two independent variables of the study, have any significant effect on comprehension scores, a two-way ANOVA was run with two between-subject factors (number of known languages and attentional condition). Before running the ANOVA, descriptive statistics were calculated for the groups. Table 1 presents the mean scores and standard deviations for all groups.

Table 1
No. of known languages and attentional condition

Group	Mean	SD	no.
Monolingual	3.16	1.24	18
Bilingual	3.06	1.34	16
Mon.(control)	2.3	1.25	10
Bi. (control)	2.4	1.17	10
Total	2.83	1.28	54

As shown in table 1, there does not seem to exist a large difference between the monolinguals and bilinguals in their comprehension of the foreign text. Table 2 presents the descriptive statistics for the effects of the two independent variables, which are bilinguality and attentional condition. It shows that monolingual and bilingual speakers who focused on the content words had better mean scores than those who did not.

Table 2
Descriptive Statistics for the effects of the two independent variables

Lang. status	Attention	Mean	SD	no.
Monolingual	Circle	3.16	1.24	18
	Not circle	2.3	1.25	10
	Total	2.85	1.29	28
Bilingual	Circle	3.06	1.34	16
	Not circle	2.4	1.17	10
	Total	2.8	1.29	26
Total	Circle	3.11	1.27	34
	Not circle	2.35	1.18	20
	Total	2.83	1.28	54

The results from ANOVA, as presented in Table 3, showed that there was no statistically significant interaction between the effects of bilinguality and attentional condition, $F(1, 50) = .082$, $p = .77$. In other words, simple main effects analysis showed that there was no significant difference between monolinguals and bilinguals in terms of comprehension, $F(1, 50) = 0.00$, $p = .99$. There was, however, a significant difference in comprehension scores between the group which circled the lexical item and the group which did not, $F(1, 50) = 4.59$, $p = .03$. Overall, the results showed that bilinguality or monolinguality do not seem to have a particular role in learners' ability to attend to the meaning and to comprehend a text. In other words, it was only the focus on the lexical target that determined the learners' comprehension scores.

Table 3
Tests of between-subjects effects to examine the effects of attention and bilinguality

Source	Type III Sum of Squares	Df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	7.563a	3	2.521	1.577	.207	.086
Intercept	375.553	1	375.553	234.904	.000	.825
attention	7.352	1	7.352	4.599	.037	.084
No. of known langs.	5.459E-5	1	5.459E-5	.000	.995	.000
attention*No. of known langs.	.131	1	.131	.082	.776	.002
Error	79.938	50	1.599			
Total	521.000	54				
Corrected Total	87.500	53				

a. R Squared = .086 (Adjusted R Squared = .032)

It was assumed that learners' overall comprehension was likely to deteriorate when they had to put in additional cognitive resources and efforts to process meaning and form simultaneously and that the average comprehensions score for those who had to circle the lexical item and process it more deeply would be lower than those who only processed the text for meaning. However, the findings of the study, as shown in Table 3, did not support these assumptions. In fact, the results proved quite the opposite. As shown in Table 3, those who paid simultaneous attention to form and meaning and processed the targeted item for both form and meaning gained better comprehension scores regardless of the number of languages they knew. Overall, the findings from the present study do not support the claim that deeper processing of lexical items hinders processing for meaning and overall comprehension of the text.

Results of Coding: Operationalization of Attention to Meaning and Form

To answer the second question of the study, think-aloud protocols were gathered for the groups which had to circle the target form. They were coded to make sure that the participants were following the instructions and had paid attention to the targeted form while their attention was simultaneously focused on meaning. Attention to meaning was, therefore, operationalized as the participants' sustained effort to comprehend the meaning of the text while simultaneously mentioning and/or circling the form. The participants who did not report the targeted form by a

minimum of 60% were excluded from the study. The think-aloud protocols revealed that for the groups, whether bilingual or monolingual, there were some participants who did not process the input for meaning. Others looked back at the text to answer the comprehension questions (backtracking). As mentioned earlier, they were eliminated. Adding these participants to the data pool would have jeopardized the validity of the study as processing for meaning was the basic requirement of the study and could have a detrimental effect on the results of the study. Half of the think-aloud protocols were coded by two raters. The interrater reliability was calculated and found to be 85%.

Depth or Levels of Processing

There were also differences in the learners' level of attention and the processing of the target lexical form. The participants' level of processing ranged from mere attention to form, pronouncing the word, raising the intonation while reading the word, to translating or interpreting. To explore these levels, concurrent data were analyzed by two coders to categorize the levels of processing (inter-coder reliability was 90%). The think-aloud data revealed three levels: The first level was associated with the simple circling of the target form. The second level was providing a report of processing for example by reading with a noticeable intonation, and the third level was translating the target form (see Table 4 for analysis).

Table 4
Descriptive statistics of different levels of processing

processing	*Mean*	*SD.*	*no.*
First level	2.5	1.30	8
Second level	2.44	.88	9
Third level	3.76	1.14	17
Total	3.11	1.27	34

Table 5
Tests of between-subjects effects for different levels of processing

Source	Type III Sum of Squares	*Df*	Mean Square	*F*	*Sig.*
Corrected Model	14.248ᵃ	2	7.124	5.622	.008
Intercept	257.173	1	257.173	202.957	.000
process	14.248	2	7.124	5.622	.008
Error	39.281	31	1.267		
Total	384.000	34			
Corrected Total	53.529	33			

a. R Squared = .266 (Adjusted R Squared = .219)

As it was already mentioned, there were different levels of processing; some of the participants reported the target form in addition to circling it. They made some comments about it or changed their intonation or reading it louder or paused after it. All of these cases were considered as indications of a deeper level of processing compared to when participants were only circling the word and therefore this could suggest allocating more attentional resources to it (Craik, 2002). As attention to form was the only predictor of learners' ability to comprehend the text, a comparison was made between the learners with different levels of attention to the text to examine the effects of different levels of processing on the learners' comprehension of the presented text.

As shown in Table 6, the difference between the participants processed the text at a deeper level (the third level) had significantly higher comprehension scores than the other two groups. The difference between the first level and the second level participants did not even approach significance. However, it is not surprising that most of the participants embarked on interpreting and translating the word, given its salience in the text. However, at the start of reading, they just noticed the word and then gradually they realized that the whole text depended on that word. They tried to make sense of the word by guessing or interpreting the word from the context. As shown in the table, most of the learners showed signs of deeper levels of processing. This may not be surprising, as the circled item is a lexical item and is, therefore, more salient and also crucial to comprehending

the text. What is clear, nevertheless, is that deeper levels of processing, as it was mentioned, are associated with better comprehension ability.

Table 6 *Multiple Comparisons of different levels of processing*

(I) process	(J) process	Mean Difference (I-J)	Std. Error	*Sig.*	95% Confidence Interval Lower Bound	Upper Bound
First	Second	.0556	.54698	.994	-1.2907	1.4018
	Third	-1.2647	.48263	.035	-2.4525	-.0769
Second	First	-.0556	.54698	.994	-1.4018	1.2907
	Third	-1.3203	.46404	.021	-2.4623	-.1782
Third	First	1.2647	.48263	.035	.0769	2.4525
	Second	1.3203	.46404	.021	.1782	2.4623

Based on observed means.
 The error term is Mean Square(Error) = 1.267.
*. The mean difference is significant at the .05 level.

Leow et al. (2008) pointed out that the degree of engagement with form depends on the saliency of the target forms. Participants tend to notice and process content words more effectively than function words (Ellis, 2006). Therefore, attention to form, in this case, does not appear to reduce attention to meaning and, hence, does not seem to affect comprehension detrimentally. All in all, this result indicates that there seems to exist a clear relationship between the level of processing and the comprehension scores. The results of the study, therefore, corroborate the finding of previous studies that simultaneous attention to form and meaning does occur but at different levels.

Discussion

The current study has investigated the effects of bilinguality and type of attentional condition on simultaneous attention to form and meaning. With regard to the effects of knowing an additional language on attention to form and meaning, no significant difference was observed between monolingual and bilinguals. To the best of researchers' knowledge, no previous studies have addressed this issue, therefore drawing comparisons to other studies is not possible. Further research, especially of neurological type, is recommended to uncover the complexities of bilingual and multilingual subjects. With regard to the effect of type of attentional condition, it was found that it did not affect comprehension negatively. This finding is in line with the results from Morgan-short et al. (2012), Leow et al. (2008), and Wong (2001) but different from the findings from the studies on written comprehension such as Greenslade et al. (1999) and from the studies on aural comprehension by VanPatten (1990) and Wong (2001). This difference was not observed in Greenslade et al.'s study. Our study also confirms VanPatten (1990)'s finding in that focusing on lexical items may lead to better overall comprehension of the text whereas a more non-salient redundant feature may hinder comprehension. A possible explanation for this result might be methodological issues. In Greenslade et al.'s (1999) and VanPatten's (1990) studies, the participants were asked to mark all the targeted forms during exposure and nearly 67% of the –*n* forms were in one paragraph. Thus, it was not clear whether participants were processing the form and meaning simultaneously. In our study, we made sure that there is an equal distribution of the target form across the passage. Also, some of the participants in VanPatten's study admitted that they were not following the instructions for paying simultaneous attention to form and meaning which make the data collected from those participants questionable. In the present study, however, special care was taken to make sure that the participants were following the instructions as precisely as possible.

As suggested in previous studies (e.g. Leow et al., 2008; Wong, 2001), modality can be a possible explanation for successful simultaneous attention to form and meaning. Four studies administered via a written text, including Morgan-Short et al., Leow et al., Wong and the current study report that comprehension is not negatively affected by the attentional condition. The two studies administered in aural mode, i.e., VanPatten (1990) and Wong (2001), on the other hand, have reported negative effects of attention to form on comprehension.

Another possible explanation for the differences between the studies conducted in written mode could be the learners' proficiency level. In the present study, the participants' proficiency level

was controlled by administering a proficiency test at the beginning of the study. The learners in Wong's study recalled 12 idea units out of 52 whereas the learners in Greenslade et al. (1999) recalled 22.5 idea units out of 53. This clearly indicates that the participants in Greenslade et al. 's (1999) study were more proficient. For the studies to be comparable, similar levels of proficiency should be ensured to rule out the effect of mediating variables.

Since the current study and Leow et al. (2008) used different measures of comprehension from those of Greenslade et al. and Wong, it would be difficult to compare them. If these studies employed the same measure of comprehension, then their comparison would have been more plausible and would enable a more valid comparison. Although the proficiency level of the participants could possibly interfere with the results in such research studies, it does not challenge the conclusion that modality is an important variable. It seems that in the written mode, attention to form while reading for meaning does not have a negative effect on comprehension. However, this does not appear to be true about the aural mode.

Another issue worth considering is why studies conducted through different modes of presentation produced differing results with regard to simultaneous attention to form and meaning. Some suggest that cognitive constraints may account for the existing differences as attentional capacity may be controlled differently by cognitive factors in written and oral modes (Wong, 2001). In other words, due to the simultaneous nature of the oral mode, learners' cognitive capacities might have a more determining role in their ability to comprehend. In the written mode, however, this might not be quite a factor as the text is not transient in nature. Accordingly, a possible explanation could be that aural mode is constrained by the limitations of the processor, whereas in written mode, processing is not constrained in the same way. However, it should be noted that in the studies in aural mode timing was controlled, but in the written mode the participants were asked to circle the items at their own pace. Recently, research in cognitive psychology implies that attentional constraints and processes are basically the same in different modalities (Chun, Golomb & Turk-Browne, 2011). However, in the written mode, attentional constraints become more evident when the timing is controlled through the rapid visual presentation of input.

To further explore the issue of attention during the process of comprehension, this study also examined the role of depth of processing. To determine this and finding the relationship between the level of processing and the overall comprehension the verbal protocols were analyzed. Leow et al. (2008) had argued that the framework of level of processing proposed by Craik and Lockhart (1972) may explain the results of the studies which reported the lack of effects of attentional condition on comprehension.

In Leow et al.'s (2008) study since most of the targeted items were forms, few participants had attended to the targeted forms at a deep level of processing. They stated that this shallow processing of form used limited attentional resources and most of the attention was given to processing meaning. They suggested that "the non-significant difference in comprehension between experimental conditions might be attributed to the relatively low level of processing reported in all experimental groups in regard to the targeted form" (p. 686). The only conclusion they could arrive at giving their limited number of participants was that attention to form did not have negative effects on comprehension. Interestingly, the findings of the current study corroborate this statement as it was shown that deeper levels of comprehension lead to increased comprehension scores.

In the current study, similar to the work done by MorganShort et al. (2012), there was sufficient data to run statistical analysis that showed that a deeper level of processing is associated with a better comprehension score. This conclusion, however, does not seem to be in congruence with VanPatten's primacy of meaning principle. Nevertheless, these results do seem to be consistent with the predictions made by Craik's (2002) levels of processing framework, which claims that retention of items in memory depends on the level of processing of the item when encoding. The framework postulates that recalling items which have been processed deeply is more likely than those who have undergone shallow processing. The findings of this study confirm the predictions of this framework. When the lexical form is processed deeply, attention to it does not interfere with processing for meaning which leads to better comprehension.

Moreover, the type of linguistic target might have played an important role in the findings too. As the results of the tables display, most of the students processed the targeted form at a deep level. *Inflation*, being a lexical item, carried more semantic weight and therefore was noticed more easily and processed more deeply compared to grammatical items which carry less meaning and sometimes seem to be semantically redundant (Morgan-Short et al., 2012). The same holds true in Leow's and Morgan-Short's studies as participants in these studies processed *Sol* which is a lexical

item more deeply than the morpheme *–n* which carries less meaning. The low level of processing of grammatical forms in these two studies did not appear to impose any cognitive load when processing the text for meaning and therefore no effect is seen on comprehension contrary to the results of the Greenslade et al. (1999) and VanPatten (1990).

Conclusion

The current study attempted to expand the findings of previous studies with an advantage in which the proficiency level of learners was controlled for. The primary goal was to see if there were any differences between monolingual (Persian as the first language) and bilingual (Kurdish as a first and Persian as second language) learners of English. The results confirmed the findings of Leow et al. (2008) to a large extent. No significant difference was observed between bilingual and monolingual learners regarding the level of processing of the targeted lexical form as far as the subject of attention to meaning and form is concerned. However, in the current study, the results of the analysis of data yielded interesting findings with regard to the level of processing. It indicates that learners should be encouraged to focus more on the content words and process them more deeply since deeper processing is associated with better reading comprehension according to the findings of the present study. Furthermore, the results showed that, as with Leow et al. (2008) and Wong (2001), paying attention to lexical forms when trying to understand meaning did not have any negative effect on comprehension, as VanPatten's primacy of meaning principle would predict. In fact, the results showed that focus on the lexical form does have a positive effect on comprehension. This means that using input enhancement techniques such as textual input enhancement could positively affect comprehension especially if the content words are targeted. The analysis of verbal protocols further revealed that the more deeply the lexical item is processed the greater the comprehension score would be. Further research probably should consider the effect of modality, the timing of input, level of processing and L2 proficiency so that we may arrive at a clearer understanding of the issue.

Like any other study, however, this research project suffers from a number of shortcomings. The most prominent ones are as follow: First of all, the number of participants was somehow low. This will jeopardize the generalizability of findings beyond the current research. The mortality of some participants due to backtracking was also disappointing. Second, this study focused on the allocation of attention to lexical meaning and ignored the grammatical form. Future studies can investigate the allocation of attention to grammatical form and meaning among multilingual learners. Third, eye-tracking technique, a new research method, could also be employed to investigate the issue of attention as it was applied in some researches (see Godfroid & Uggen, 2013; Godfroid et al., 2013). The use of eye-tracking is also a better technique to control for the potential effect of reactivity. Fourth, this study specifically focused on written input and ignored aural input. Future studies could be conducted with aural input or both written and aural input being presented within the same time period to control for the effects of timing. Finally, although our results were discussed with reference to levels of processing, it has a number of shortcomings. The absence of an objective measure of the depth of processing was an issue (Craik, 2002). Clear distinctions need to be made between different levels of processing to enable researchers to make more reliable and consistent conclusions and comparisons.

References

Armengol, L., & Cots, J. M. (2009). Attention processes observed in think-aloud protocols: Two multilingual informants writing in two languages. *Language Awareness, 18* (3), 259-276.

Baars, B. J., & Gage, N. M. (2010). *Cognition, brain, and consciousness: Introduction to cognitive neuroscience.* London: Elsevier Academic Press.

Bowles, M. (2010). *The think-aloud controversy in second language research.* London. Routledge.

Carrol, S. E. (2006). Shallow processing: a consequence of bilingualism or second language learning? *Applied Psycholinguistics, 27* (1), 53–56.

Chun, M. M., Golomb, J. D., & Turk-Browne, N. B. (2011). A taxonomy of external and internal attention. *Annual Review of Psychology, 62*, 73–101.

Dolgunsö z, E. (2015). Measuring attention in second language reading using eye-tracking: The case of the noticing hypothesis. *Journal of Eye-movement Research, 8* (5), 1-18.

Ellis, N. (2006). Selective attention and transfer phenomena in L2 acquisition: Contingency, cue competition, salience, interference, overshadowing, blocking, and perceptual learning.

Applied Linguistics, 27 (2)*,* 164–194.

Ellis, N. (2002). Frequency effects in language processing: A review with implications for theories of implicit and explicit language acquisition. *Studies in Second Language Acquisition, 24* (2)*,* 143-188.

Craik, F. I. M. (2002). Levels of processing: Past, present... and future? *Memory, 10* (5), 305–318.

Craik, F. I. M., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior, 11(6),* 671–684.

Gass, S. M., & Mackey, A. (2000). *Stimulated recall methodology in second language research.* Mahwah, NJ: Erlbaum.

Gass, S., Svetics, I., & Lemelin, S. (2003). Differential effects of attention. *Language Learning, 53*(3), 497–545.

Godfroid, A., Boers, F., & Housen, A. (2013). An eye for words: Gauging the role of attention in incidental L2 vocabulary acquisition by means of eye-tracking. *Studies in Second Language Acquisition, 35* (3), 483-517.

Godfroid, A., & Uggen, M.S. (2013). Attention to irregular verbs by beginning learners of German. *Studies in Second Language Acquisition, 35*(2), 291–322.

Kramsch, C. (2012). Authenticity and legitimacy in multilingual SLA. *Journal of Critical Multilingualism Studies 1*(1), 107–128.

Kormes, J. (1998). Verbal reports in L2 speech production research. *TESOL Quarterly 32* (2), 353-358.

Leow, R. P., Hsieh, H., & Moreno, N. (2008). Attention to form and meaning revisited. *Language Learning, 58* (3), 665–695.

Lew-Williams, C. (2009). Real-time processing of gender-marked articles by native and non-native Spanish-speaking children and adults. Ph.D. Dissertation. Stanford University.

Logan , G. (2005). The time it takes to switch attention. *Psychonomic Bulletin & Review, 12* (4), 647–653.

Greenslade, T. A., Bouden, L., & Sanz, C. (1999). Attending to form and content in processing L2 reading texts. *Spanish Applied Linguistics, 3*, 65–90.

LaBerge, D. (1995). *Attentional processing: The brain's art of mindfulness.* Cambridge, MA: Harvard University Press.

Mackey, A., Gass, S., & McDonough, K. (2000). How do learners perceive interactional feedback? *Studies in Second Language Acquisition, 22*, 471–97.

MacWhinney, B. (1987). The competition model. In B. MacWhinney (Ed.), *Mechanisms of language acquisition* (pp. 249-308). Hillsdale, NJ: Lawrence Erlbaum.

Malko, A. Ehrenhofer, L., & Phillips, C. (2016). "Theories and frameworks in second language processing". *Bilingualism: Language and Cognition, 20.*

Morgan-Short, K., Heil, J., Botero-Moriarty, A., & Ebert, S. (2012). Allocation of attention to second language form and meaning. *Studies in Second Language Acquisition, 34*, 659–685.

Neill, W. T., Valdes, L. A., & Terry, K. M. (1995). Selective attention and the inhibitory control of cognition. In F. N. Dempster & C. J. Brainerd (Eds.), *Interference and inhibition in cognition* (pp. 207-261). San Diego: Academic Press.

Norris, J. M., & Ortega, L. (2000). Effectiveness of L2 instruction: A research synthesis and quantitative meta-analysis. *Language Learning, 50,* 417–528.

Rayner, K. (2009). Eye movements and attention in reading, scene perception, and visual search. *The Quarterly Journal of Experimental Psychology, 62*, 1457–1506.

Robinson, P. (1995). Attention, memory, and the "Noticing" Hypothesis. *Language Learning, 45(2),* 283–331.

Robinson P. (2017). Attention and awareness. In J. Cenoz, D. Gorter, & S. May (Eds.), Language awareness and multilingualism. Encyclopedia of Language and Education (3rd ed.). Springer, Cham.

Robinson, P., Mackey, A., Gass, S., & Schmidt, R. (2012). Attention and awareness in second language acquisition in In S. Gass and A. Mackey (Eds.), The Routledge handbook of second language acquisition (pp. 247-267). New York: Routledge.

Schmidt, R. (1990). The role of consciousness in second language learning. *Applied Linguistics, 11*, 129–158.

Schmidt, R. (2001). Attention. In P. Robinson (Ed.), *Cognition and second language instruction*

(pp. 3–32). Cambridge: Cambridge University Press.

Schwartz, B. D. (1993). On explicit and negative data effecting and affecting competence and linguistic behavior. *Studies in Second Language Acquisition, 15(2),* 147-163.

Szubko-Sitarek, W. (2015). *Multilingual lexical recognition in the mental lexicon of third language users.* Berlin: Springer.

Tomlin, R., & Villa, V. (1994). Attention in cognitive science and second language acquisition. *Studies in Second Language Acquisition, 16(2),* 183-203.

Truscott, J. (1998). Noticing in second language acquisition: A critical review. *Second Language Research, 14*(2), 103-135.

Van der Heijden, A. H. C. (1981). *Short term visual information processing.* London: Routledge & Kegan Paul.

VanPatten, B. (1990). Attending to form and content in the input: An experiment in consciousness. *Studies in Second Language Acquisition, 12(3),* 287–301.

VanPatten, B. (1994). Evaluating the role of consciousness in second language acquisition: Terms, linguistic features & research methodology. *AILA Review, 11,* 27-36.

VanPatten, B. (2004). Input processing in second language acquisition. In B. VanPatten (Ed.), *Processing instruction: Theory, research, and commentary* (pp. 5-32). Mahwah, NJ: Lawrence Erlbaum Associates.

Wong, W. (2001). Modality and attention to meaning and form in the input. *Studies in Second Language Acquisition, 23(3),* 345–368.

About the Authors

Mohammad Nabi Karimi is Associate Professor of Applied Linguistics and Head of the Department of Foreign Languages at Kharazmi University, Tehran, Iran, where he teaches SLA, Second Language Teacher Education, Psycholinguistics, etc. for graduate students. His papers on different language issues appear in prestigious journals like *The Modern Language Journal, Innovation in Language Learning and Teaching* and *System.*

Nahid Fallah is PhD student of Applied Linguistics at Kharazmi University, Tehran, Iran and an invited lecturer at Kashan University, Iran, where she teaches some TEFL courses for undergraduate students.

Ebrahim Zangani (corresponding author) is PhD student of Applied Linguistics at Kharazmi University, Tehran, Iran and an English language teacher in the Ministry of Education. He has published some articles on language issues internationally and co-authored two general English textbooks for university students. His main areas of interest are second language teacher education, language assessment, clinical linguistics and SLA.

Utilizing Textbook Adaptation Strategies: Experiences and Challenges of Novice and Experienced EFL Instructors

Enisa Medeᵉ
Bahcesehir University, Turkey

Şenel Yalçın
Bahcesehir University, Turkey

Abstract

The purpose of this study was to investigate the utilization of textbook adaptation strategies by novice and experienced EFL instructors working in a language preparatory program at a foundation (non-profit, private) university in Turkey. Specifically, the study investigated the self-reported beliefs of the two groups of instructors about textbook adaptation and explored which adaptation strategies were most frequently implemented in their classroom. The study also attempted to compare whether there were differences between the two groups of instructors related to their use of adaptation strategies in their courses and find out the reasons behind their adaptive decisions. The participants were 14 Turkish EFL instructors (7 novices and 7 experienced) offering English courses in an intermediate level English classroom. Data were collected from reflective essays, lesson plans, and semi-structured interviews. The findings revealed that both groups of instructors shared highly positive beliefs about the implementation of adaptation strategies in their courses apart from some differences related to types of strategies and their frequencies. Besides, the adaptive decisions of the participating instructors were closely related to their students, tasks, context, time, and their own beliefs. Based on the obtained findings, the researchers provided some pedagogical implications and suggestions regarding the effective use of textbook adaptation strategies in language preparatory programs.

Key words: textbook adaptation; adaption strategies; novice instructor; experienced instructor; adaptive decisions; English as a foreign language (EFL)

Introduction

Textbooks can be referred to as published materials that help and support language learners to improve their linguistic and communicative abilities (Sheldon, 1987; Ur, 1996). They provide the language input learners require and the necessary language practice. They also provide the content of the lessons serving as a guide for teachers' instruction.

However, when teachers open a page in a textbook, they need to find answers to these questions: "Is the language at the right level? Is the topic/content suitable for the students? Is the sequencing of the lessons logical?" (Harmer, 1998, p. 111). Specifically, teachers need to make sure that textbooks include the essential elements of a language as well as the culture of the taught language referring to the needs of learners, their language proficiency, and their cultural backgrounds.

Based on these overviews, teachers should be able to select and adapt textbooks to meet the needs their classrooms and individual students. During the process of textbook adaptation, many scholars provided various reasons for teachers' planned decisions on selecting and adapting materials (Ebrahimpourtaher & Hamidi, 2015; Edge & Wharton, 1998; Gabrielatos, 2004; Graves, 2000; Mısırlı, 2010). According to Graves (2000), teachers' beliefs, the implications they draw out of their experiences, and their educational background are among the reasons that lead teachers to adapt course materials. In addition, students' needs, and interests are also determining factors which affect teachers' adaptive decisions. Finally, teaching contexts, testing plans, and student profiles and timetables might impact materials adaptation in language classrooms.

Email: enisa.mede@es.bau.edu.tr, Department of English Language Teaching (ELT), Faculty of Educational Sciences, phone number: +905363957295

Mısırlı (2010) listed more reasons for adapting textbooks. One of them is lack of grammar input which can be a problem when students have a particular goal in their mind such as, preparing for national or international language tests. Lack of communicative activities or lack of balance among language skills may also drive teachers to adapt textbooks. Similarly, discrepancies related to proficiency level and learning styles between the materials, activities, and the textbook users along with the activities are among the reasons for adapting textbooks as well.

Another reason for textbook adaptation is related to the course objectives that have been determined in the educational context; adaptation arises as a need for facilitating the learning process through identifying learning problems so that learners are able to internalize content effectively (Ebrahimpourtaher & Hamidi, 2015). If the materials are adapted to increase learners' awareness, they can be prepared to take their own decisions, control their own learning, and develop their own preferred learning styles. This can result in autonomous learning.

Finally, Edge and Wharton (1998) argued that experienced teachers have the tendency to make changes in textbooks both at the planning stage and while teaching classes in response to student reactions. They also reported that experienced teachers do not usually teach their classes entirely based on what textbooks offer as they can easily interpret the content in a more critical way.

Apart from planned textbook adaptations summarized in the previous section of this study, teachers have been observed to make on-the-spot adaptations as well (Bailey, 1996; Shavelson & Stern, 1981). To exemplify, Shavelson and Stern (1981) stated that the need for adaptation may arise unexpectedly when a teacher realizes that their ordinary style of teaching is not appreciated by students or there is a lack of interest. Similarly, Bailey (1996) indicated that teachers may decide to focus on a specific part of the lesson upon a question from one student, reckoning that other students may also benefit. Students may show enthusiasm about a particular subject, and in this case, teachers may want to make the most of the moment and divert from their plans. Likewise, students' learning styles, their engagement, and equal participation may also be reasons for adaptation which may provide each student with a more effective learning environment and thus, help them to improve their language proficiency.

In brief, it can be said that the job of a teacher is similar to that of a writer (Madsen & Bowen, 1978). Just like a writer who imagines the questions their readers might want to ask; a teacher also needs to meticulously monitor students and respond to all of their voiced and unvoiced needs. To meet these needs and make students more active in the language learning process textbook adaptation is inevitable in any educational context.

Textbook Adaptation Strategies

A close look at the literature illustrates that various adaptation strategies have been suggested (Islam & Mares, 2003; Maley, 1998; McDonough & Shaw, 1993; McGrath, 2002). To begin with, McDonough and Shaw (1993) came up with three basic strategies to be implemented while adapting materials. The first strategy is *adding* which is related to extending the existing material by including more of the same material to increase its quality and effectiveness. The second strategy deleting or *omitting* refers to deleting material on a small or large scale. When the material is reduced quantitatively, it is referred to as *subtracting*, while *abridging* is about the changes made to the methodology. Finally, the *modifying* strategy is examined under two sub-categories, *rewriting* and *restructuring*. Both types of modification involve an internal change that can be implemented in any aspect of the material. While *rewriting* is usually done to cater for students' interests and their backgrounds, *restructuring* happens when the teacher changes the organization of activities to suit the number of students, classroom size, and so on.

Furthermore, the researchers added two more adaptation strategies to be implemented in classroom settings. They indicated the importance of *simplifying* as a type of rewriting activity to aid with student comprehension. Finally, *re-ordering* is the last strategy regarding alterations in the order of the activities within a unit or among the units depending on students' needs and levels.

A very similar list of adaptation strategies was proposed by Maley (1998) including omission, addition, reduction, extension, replacement, re-ordering, and branching strategies. To

exemplify, while *omission* means leaving out things which are not appropriate for the target learners, *addition* is just the opposite. In other words, with this particular strategy the teacher adds some material to meet the needs and interests of the students. Additionally, using the *reduction* strategy makes the activity or material shorter whereas *extension* adds an alternative dimension. Similarly, using the *replacement* strategy helps the teacher replace the material with something more suitable while *re-ordering* aids with the decision to plot a different course through the material than the one the writer intended. Lastly, with the use of the *branching* strategy, the teacher adds options to the existing activity or suggests alternatives.

Looking at the different adaptive strategies suggested in the literature, it is obvious that textbook adaptation is a broad and extensive process which can be utilized in different educational contexts. As stated by Graves (1996), the degree of adaptation may vary with the nature of actual class performance ranging from simple to complex which will have impact on the teachers' adaptive decisions.

Textbook Adaptation in EFL Classrooms

A close review of the literature revealed that teachers' years of teaching experience might lead to differences in terms of their adaptive decisions (Akyel, 1997; Bigelow, 2000; Tsui, 2003). To begin with, Akyel (1997) examined whether there were any differences between experienced and novice EFL teachers in terms of their instructional goals, actions, and thoughts. The results of the study revealed that there were more similarities than differences between the two groups of teachers in terms of their instructional goals and actions. On the contrary, there was a noteworthy difference related to teacher's years of experience and students' actions. Specifically, while experienced teachers were more responsive to students' reactions, novice teachers believed that these reactions affected the flow of their classes negatively.

In a similar study, Bigelow (2000) compared the lesson planning of three teachers with varying levels of teaching expertise and looked at how they implemented the lessons in their classrooms. The findings showed that the novice teacher mostly concentrated on students' engagement in class and with the materials. The more experienced teacher, on the other hand, emphasized more on challenging students by simplifying content and the difficulty of tasks. Finally, the teacher with the most experience who was referred to as an expert, was primarily interested in arranging the class in a manner that encouraged students to begin with engagement and then, to gradually work more individually. These results clearly illustrated that years of teaching experience has a crucial role in teachers' adaptive decisions.

Furthermore, Ç oban (2001) conducted a study with 8 novices and 8 experienced teachers which aimed to understand their use of adaptation strategies in their courses. The study also attempted to reveal the rationale behind the teachers' adaptations, and to find if any differences exist regarding their adaptive decisions. The findings obtained from classroom observations and interviews indicated that both groups of teachers used addition as the most frequent adaptation strategy. The majority of these strategies were related to the task itself, and the reasons behind these adaptations were related to teachers' perceptions, students' interests, their needs, and the nature of the tasks. The results also showed that there were no considerable differences between the adaptive decisions made by the two groups of teachers.

Tsui (2003) investigated the pre-active and interactive phases of expert and novice teachers' teaching decisions. The findings of the study revealed that the experienced teachers tried to meet the curriculum objectives more closely and they made adaptations to the related materials in an autonomous manner. However, the novice teachers were hesitant to deviate from the suggested plans. Regarding the length of planning, expert teachers preferred to carry out more long-term planning, whereas novice teachers planned for shorter periods. Expert teachers were much more efficient because they could recall their previous experiences during the process of lesson planning and teaching, while the novices needed more time to make detailed plans before they taught.

Research Questions

Guided by previously reported studies and based on the obtained findings, years of teaching experience is one of the major reasons teachers' adaptive decisions might be affected during classroom practices. To gain more in-depth information on this issue, the adaptive decisions of teachers with different years of expertise should be investigated more closely to find out whether there is any difference between their adaptive decisions and to discover the reasons behind them. The present study, therefore, aims to identify the self-reported beliefs of novice and experienced EFL instructors about adapting textbooks in an intermediate level classroom of a language preparatory program. The study also attempts to examine and compare what textbook adaptation strategies the participants utilize most frequently in their classroom practices. Lastly, the study tries to find out what reasons lead to the instructors' adaptive decisions. To meet these objectives, the following research questions and sub question were addressed in this study:

1. What are the self-reported beliefs of the novice and experienced instructors about textbook adaptation in an intermediate level English preparatory classroom?
2. Which adaptation strategies are most frequently implemented by the two groups of instructors in their classroom?
   2a. Are there any differences between their use of the adaptation strategies?
3. What are the reasons behind their adaptive decisions related to these pre-set categories:
   a. students
   b. time
   c. tasks
   d. context
   e. teacher beliefs

Methodology

Design

For the purposes of this research, a case study was adopted as a research design to enable the researcher to closely examine the data within a specific context (Zainal, 2007). A case study, which is a type of qualitative enquiry, is defined by Yin (2002) as "an empirical inquiry that investigates a contemporary phenomenon within its real-life context, especially when the boundaries between phenomenon and context are not clearly evident" (p. 13). In this sense, the data related to the specific context (i.e., a language preparatory classroom) were collected and analyzed using qualitative research methods.

Participants and Setting

This study was carried out at an English preparatory school of a foundation (non-profit, private) university in Istanbul, Turkey during the second semester of the academic year 2017-2018. In the program, students are tested with a proficiency exam at the beginning of the year in order to determine whether or not their command of English is adequate for the study requirements of their departments. Students who score 60 out of 100 points or above directly start their undergraduate programs at various academic disciplines. However, the ones who receive scores below the average are required to take the language placement test administered by the preparatory school. In this exam, the Common European Framework of Reference for Languages (CEFR/CEF) is taken into consideration to determine the proficiency level of the students. According to the placement exam scores, the students start the preparatory program which lasts between 8 to 18 weeks. The program follows a modular system which offers language skills and grammar courses in accordance with the proficiency level of the students. After the students complete the related modules of the preparatory program, they take an achievement exam. The ones who score above 60 start their undergraduate programs. However, the students who fail have to repeat the preparatory program until they earn the points to pass the exam.

   Considering the preparatory class that the present study was conducted (B2, intermediate

level), New Language Leader Intermediate (http://product.pearsonelt.com/newlanguageleader/) textbook is used as the main source which is accompanied by weekly packs prepared by the instructors.

For convenience, the participants of this study (7 novices and 7 experienced EFL instructors) were chosen from the group of instructors who were teaching intermediate classes in an 18-week long module. Table 1 summarizes the demographic profile of each participant:

Table 1
Demographic Profile of the Participants

Pseudonym	Novice/Experienced (Teaching Experience)	Field of BA and MA (obtained)	Other Qualifications
EI 1	Experienced (22 years)	BA in Linguistics MA in Applied Psychology	-
EI 2	Experienced (22 years)	BA in English Language Teaching	-
EI 3	Experienced (20 years)	BA in English Language Teaching	-
EI 4	Experienced (15 years)	BA in Western Languages and Literature MA in Human Resources Management	-
EI 5	Experienced (11 years)	BA in English Language Teaching MA in Linguistics	Certificate in English Language Teaching (CELTA)
EI 6	Experienced (10 years)	BA in English Language & Literature MA in Educational Management & Planning	Teaching English as a Foreign Language (TEFL)
EI 7	Experienced (9 years)	BA in English Language & Literature MA in American Culture & Literature (on progress)	CELTA
NI 1	Novice (5 years)	BA American Culture & Literature	
NI 2	Novice (5 years)	BA in English Language Teaching MA in English Language Teaching	-
NI 3	Novice (4 years)	BA in English Language & Comparative Literature MA in English Language	-
NI 4	Novice (4 years)	BA in Technical Translation & Teaching Foreign Languages MA in BA in Technical Translation & Teaching Foreign Languages	CELTA
NI 5	Novice (4 years)	BA in Foreign Languages Education BA in Foreign Languages Education (in progress)	TEFL
NI 6	Novice	BA in Teaching English as a	TEFL

	(4 years)	Foreign Language	
		MA in English Literature	
NI 7	Novice (1 year)	BA in Cultural History	CELTA

Note: EI: Experienced Instructor; NI: Novice Instructor

Data Collection

Before the data collection process, permission was obtained from the director of the preparatory school and the level coordinator. Then, all EFL instructors offering B2 level preparatory courses were asked to complete a consent form in which they were provided with the opportunity to participate or not in the study. As a result, the instructors who were willing to take part in this research were ensured with the maintenance of confidentiality regarding the collected data.

In order to, answer the research questions and the sub question of this study, data were collected from reflective essays, lesson plans, and semi-structured interviews administered to the participating instructors. The following section provides detailed information regarding each tool along with the rationale behind their use.

Reflective essays.

To answer the first research question of this study about the self-reported beliefs of novice and experience instructors about textbook adaptation, both groups were asked to write a reflective essay about their general beliefs on adapting textbooks in their classes. The two groups of instructors were also requested to specific provide examples to gain more in-depth information about their adaptive decisions.

Lesson plans.

As the second research question and sub-question aimed to find out which types of adaptation strategies the novice and experienced instructors most frequently used, as well as explore whether there were any differences between their preferences, data were collected from lesson plans. Specifically, the two groups of instructors were asked to prepare a lesson plan related to two pages (76 and 77) of New English File Intermediate textbook. For the purposes of this study, a different textbook was chosen because some of the participating instructors might already have had some familiarity with the main book used in the B2 level preparatory class. This might have had an impact on their lesson planning process. In other words, the familiarity with the textbook could have led to ready-made lesson plans which might have affected the results of this study. Therefore, the two groups of instructors were provided with two pages of a new textbook, and they were asked to prepare a lesson plan using a template designed by the teacher development unit of the preparatory program.

Specifically, the template comprised three sections: aim, procedures, and interaction patterns. In the first section, the instructors had to explain how they plan to start their lesson. In other words, they had to provide a sample lead-in or a warm-up activity related to the content of the lesson. Next, the instructors had to explain the particular aim of their lesson (i.e., introducing new vocabulary). In the second section, the instructors were expected to describe the procedures of their lesson including their planned actions, defining the anticipated problems, and adding predictable responses of the students. Finally, in the last section related to interaction patterns, the participants were asked to indicate the desired interaction relationship, i.e., Teacher-Students, Students-Students, etc. After the two groups of instructors completed these three sections, they were asked to implement their lesson plans for 2 hours (90 minutes) in their classroom.

Semi-structured interviews.

Considering the third research question of this study which addressed the reasons behind the adaptive decisions of the participating instructors, a semi-structured interview was carried out with each participant (both novice and experienced) individually. The interview questions were adapted from the pre-set categories identified in Çoban's (2001) study which aimed to

investigate what adaptive decisions novice and experienced teachers take in their classrooms and reveal the rationale behind their decisions.

The interview included two parts. The first part collected brief information about the **participants' demographic background** in terms of their years teaching of teaching experience, their educational background, and other teaching qualifications. The second part aimed to investigate the reasons which led **to the participating instructors' adaptive decisions. Specifically,** all instructors were prompted with questions related to 5 pre-defined categories: students, time, tasks, context, and teacher beliefs which provided a frame and comprehensiveness for the interview. Each interview lasted approximately for 40-50 minutes. The interviews were audio-recorded and transcribed by the two researchers.

Data Analysis

In this study, the obtained data were analyzed using qualitative techniques. To begin with, to find out the most frequently adaptive strategies used by the novice and experienced instructors and examine whether there were any differences between the two groups, the researchers calculated the frequencies. In addition, the data obtained from the reflective essays and semi-structured interviews were analyzed through pattern coding (Miles & Huberman, 1994). First, the answers given to the open-ended questions were categorized into similar groups; then, these categories were used to determine the main themes. As the next step, the voice recordings of the interviews were converted into transcripts. After reading the transcriptions, the same process of open-ended questions was repeated, and the main themes were determined by coding. Lastly, the main themes were interpreted by the two researchers providing samples from the excerpts and/or quotations of the participating instructors. To identify the degree of agreement between the two researchers, the inter-rater reliability was found to be .86, which indicated a close agreement between the two raters on the general themes.

Finally, the lesson plans and the type of adaptation strategies utilized by novice and experienced instructors were identified referring to the seven strategies (based on McDonough & Shaw, 1993): *adding, deleting, modifying, simplifying, re-ordering, replacing* and *branching*. The novice and experienced instructors were asked to use these strategies in their lesson plans. Based on these pre-set strategies, the preferences of the two groups were reported using frequencies. As the last step, a comparison was made to see if years of teaching experience had **any impact on instructors' adaptive decisions.**

Findings and Discussion

Self-Reported Beliefs of Novice and Experienced EFL Instructors about Textbook Adaptation

The analysis of the reflective essays indicated that the majority of the EFL instructors felt the need to use adaptation strategies in their classroom and that they were highly positive about their adaptive decisions. In other words, the two groups of instructors (both novice and experienced) preferred to use adaptation strategies in their classrooms.

In their reflective essays, the participants stated that they felt the need to adapt their textbooks to meet the course **objectives. This finding was in accordance with Tsui's (2003) study** which revealed that meeting the course objectives is one of the primary reasons for textbook adaptation. Excerpt 1 taken from the reflective essay of an experienced instructor supports this finding:

> *Excerpt 1: [...] I feel the need to use adaptation strategies in my lesson quite often. It helps me meet the course objectives. (Experienced instructor 1, Reflective Essay data, 28th March 2017)*

Apart from the course objectives, the instructors expressed that they prefer to use adaptive strategies to meet the needs of their learners who vary in their learning styles and **interests. Similarly, Bailey (1996) and Mısırlı (2010) emphasized the importance of learning** styles as a need for adaptation. Excerpt 2 which was shared by a novice instructor is an example of this finding:

Excerpt 2: [...] Adaptation answers a certain need that is derived from learner's learning styles and interests. It helps me to meet the needs of my students (Novice instructor 7, Reflective Essay data, 29ᵗʰ March 2017)

Furthermore, both groups of instructors perceived classroom interaction and collaboration to be helpful in terms of facilitating learning and increasing student interest. These results were in line with the findings of Ç oban (2001) who showed that instructors adapted their textbooks to promote student interaction in their classrooms. Considering this finding, an experienced instructor said:

Excerpt 3: My preference is choosing activities focusing on learner interaction and collaboration. Therefore, I exploit and adapt textbooks. (Experienced instructor 2, Reflective Essay data, 26ᵗʰ March)

Types of Adaptation Strategies Utilized by Novice and Experienced EFL Instructors

In order to investigate the types of adaptation strategies utilized by the novice and experienced instructors, and find out the differences between their preferences, the lesson plans were analyzed by the two researchers. Table 2 reports the frequencies of the strategies used by the two groups of participants:

Table 2
The Types and Frequencies of Textbook Adaptation Strategies Utilized by Novice and Experienced EFL Instructors

Textbook Adaptation Strategies	A	D	M	Rep	Re-o	S	B	Total
EIs	18	34	31	9	8	-	-	100
NIs	23	29	31	2	1	-	-	76

Note: EIs: Experienced Instructors; NIs: Novice Instructors
A: Adding; D: Deleting; M: Modifying; Rep: Replacing; Re-o: Re-ordering; S: Simplifying; B: Branching

As shown in Table 2, both experienced and novice instructors utilized a considerably high number of adaptive strategies while designing their lesson plans (EI: 100, NI: 76). Edge and Wharton (1998) found similar results showing that teachers feel flexible to use adaptive strategies in their classroom practices.

On the other hand, the experienced and novice instructors differed in their preferences related to the types of adaptation strategies they used in their courses. To exemplify, while adding was utilized 23 times by novices, the experienced instructors used this strategy 18 times in their classroom. In other words, novice instructors tended to make more additions to the existing activities when compared to the experienced ones. Specifically, the additions were usually made as warm-up activities to activate students' schemata and raise their interest. This strategy was also implemented to practice newly-taught language such as vocabulary and grammar. The findings also indicated that some novice instructors added kinesthetic or online interactive games to their lesson plans.

Furthermore, for the deleting strategy, the results contrasted with those from the adding strategy. The experienced instructors utilized this particular strategy 5 more times than novices (EI: 34, NI: 29). The activities which were deleted from the textbook pages included vocabulary sections and free practice for new vocabulary structures.

Modifying was the most popular strategy among novice instructors and it was the second most popular among experienced instructors; it was used 31 times by each group. The modifications made by the participants included changes in the classroom mode, changes in the

information delivered, changes in the content of speaking tasks, etc. The most commonly modified activity was related to the practice of newly-learned grammatical structures. Even though this activity was designed as a pair-work speaking activity, 3 of the novice instructors modified it into a group-work, which increased the amount of student interaction. As for the experienced instructors, they preferred to delete it completely or replace it with their own activity (see Figure 1).
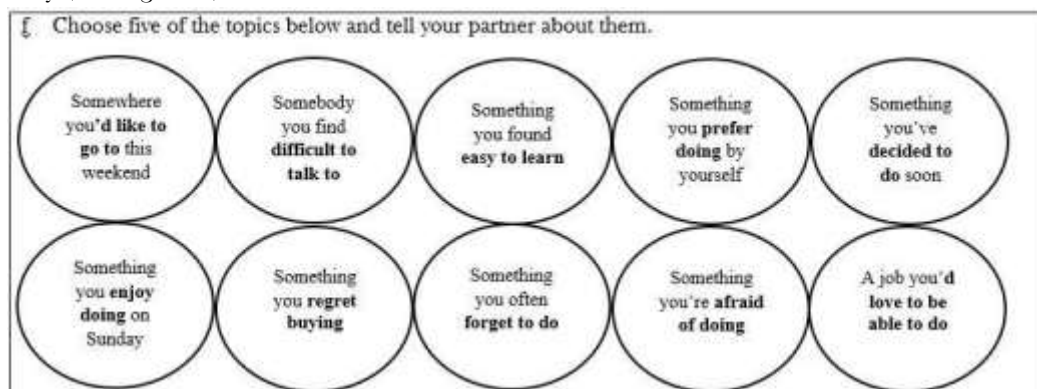

Figure 1. A Frequently Modified Activity. From Oxenden, C., & Latham-Koenig, C. (2007). *New English File Intermediate* (p. 77). Oxford: Oxford University Press.

Considering other adaptation strategies, replacing was not used as commonly as the first three strategies mentioned in the previous sections of this study. Experienced instructors were observed to replace the activities more frequently than novices (EI: 9, NI: 2). Almost half of the replacing was related to activities involving free practice of vocabulary and grammar.

The obtained results for re-ordering strategies showed some similarities with the replacing strategy which was used more frequently by the experienced instructors (EI: 8, NI: 1). To illustrate, the experienced instructors changed the order of activities while practicing vocabulary and speaking. Finally, even though simplifying and branching were among the strategies given to the participants, none of the participants used them in their lesson planning.

Based on these findings, it is obvious that both novice and experienced instructors preferred to use adaptation strategies in their classroom practices. This clearly indicates that there is a tendency towards adapting textbooks in English language preparatory classrooms.

Reasons behind the Adaptive Decisions of Novice and Experienced EFL Instructors

The third research question of this study attempted to reveal the reasons behind the adaptive decisions of the novice and experienced instructors. The following section describes the findings under the pre-set categories: student-related reasons, teacher-related reasons, time-related reasons, task-related reasons, context-related reasons, and reasons related to the teachers' beliefs.

Student-related reasons.

When the novice and experienced instructors were asked whether they needed to adapt their textbooks for student-related reasons, they affirmed that factors such as student attention, student proficiency level, student needs, and classroom interaction often prompted them to make adaptations regardless of their teaching experience.

The two groups of participants stated that their adaptive decisions were related to student interests which helped them to increase student participation and attention. Excerpt 4 was made by a novice instructor during the interview:

> *Excerpt 4: [...] My students get easily bored. To attract their attention, I prefer to adapt the activities such as changing partners, groups and places in the textbook. (Novice instructor 1, Interview data, 9ᵗʰ April 2017)*

The proficiency level of students was the second most frequently mentioned reason

during the interviews. Both novice and experienced instructors said that using textbook adaptation strategies helped students to understand the lesson better as illustrated in Excerpt 5 from an experienced teacher's interview:

> Excerpt 5: [...] I simplify or modify the activities from the textbook according to the proficiency level of my students to help their understanding. (Experienced instructor 6, Interview data, 7ᵗʰ April 2017)

Students' needs were another common reason that was mentioned during the interviews and also written in the reflective essays as displayed in Excerpt 6 shared by a novice instructor supports this finding:

> Excerpt 6: [...] Students' needs are very important. I omit, delete or simplify the activities in the textbook considering what they need in the lesson. (Novice instructor 6, Interview data, 9ᵗʰ April 2017)

Classroom interaction and student collaboration were also provided as reasons to use adaptation strategies in English classrooms. The experienced and novice instructors emphasized in their interviews that utilizing textbook adaptation in their lessons helps them to increase collaboration among students. Considering this finding, a novice instructor said:

> Excerpt 7: [...] I adapt the textbook to make sure the students interact with each other and share their ideas. (Novice instructor 4, Interview data, 7ᵗʰ April 2017)

According to the findings gathered from the interviews and reflective essays, both novice and experienced instructors shared similar reasons for their adaptive decisions. A possible explanation for this finding might be that the instructors predominantly consider their students when they plan their lessons, and they try to design their lessons to meet students' learning needs and interest in the course. All these findings agree with the study conducted by Bailey (1996) who confirmed that teachers changed their activities and modes of instruction when they wanted to increase students' interest. Similar to the findings of the present study, simplifying the difficulty level of activities to meet students' proficiency level is another adaptation reason found in Bigelow's (2000) study. As for the classroom interaction, the gathered findings in this study were parallel to those of Ç oban (2001), which revealed that teachers tend to use adaptation strategies to enhance collaboration among students.

Time-related reasons.

For the purposes of this study, both novice and experienced instructors shared time-related reasons for utilizing textbook adaptation strategies in their classroom practices. Specifically, the deleting technique was mostly employed by the two groups. All of the instructors clearly stated that they had to delete certain activities due to time constraints and pacing as displayed in the novice teacher's excerpt below:

> Excerpt 8: [...] With such a loaded weekly program that must be followed, lack of time and pacing seem to be the major problems that forces me to omit some exercises. (Novice instructor 4, Interview data, 8ᵗʰ April 2017)

This finding is in accordance with the finding of Bigelow (2000) supporting the fact that teachers give importance to time management during their lesson plans. This clearly shows that instructors feel responsible for covering the existing program on time.

Task-related reasons.

Another common reason identified for adaptation was tasks not being relevant to the learning styles of the students. The participant instructors indicated that the given tasks were not related to the students' learning styles. One of the experienced instructors shared the following comment in the interview:

> *Excerpt 9: [...] Some tasks are not related to my students' learning styles. Therefore, I adapt the activities to make sure they are appropriate for them. (Experienced instructor 6, Interview data, 9ᵗʰ April 2017)*

The two groups of instructors also stated that some tasks were very repetitive, and the students felt bored. This led them to use the modifying strategies as shared by an experienced instructor below:

> *Excerpt 10: [...] I have to modify the activities as they are very repetitive, and the students get bored easily. (Experienced instructor 7, Interview data, 7ᵗʰ April 2017)*

In brief, task-related were among the important reasons for the implementation of the adaptive strategies both by the novice and experienced EFL instructors. The instructors highlighted the importance of students' learning styles, and they also tried to avoid repetition in the tasks to gain students' attention. These findings show that material adaptation helps to increase students' attention and their motivation in language classrooms.

Context-related reasons.

Context-related reasons were perceived crucial by the instructors for their adaptive decisions as well. Specifically, both novice and experienced instructors added more exam-related activities related to the reading skills as shared in Excerpt 11 of a novice instructor:

> *Excerpt 11: [...] I prefer the type of material that we generally test in exam; I always try to add one or two reference questions when we cover reading. (Novice instructor 2, Interview data, 8ᵗʰ April 2017)*

Considering the exams, the participants emphasized the influence of pacing on adapting materials. Excerpt 12 which was share by one of the experienced instructors clarifies this point:

> *Excerpt 12: [...] Pacing is very important in our program as we have exams. In order to catch up with the schedule, I change the order of the activities or replace them with my activities from different sources. (Experienced instructor 4, Interview data, 7ᵗʰ April 2017)*

To summarize, context-related reasons were considered crucial for the adaptive decisions of novice and experienced instructors. This finding might be related to modular system of the preparatory program which accommodates pre-scheduled and centralized exams. All classes are required to fulfill the necessary learning objectives before the pre-scheduled exams. The importance of exams was also listed as a reason for adapting materials by Graves (2000) and Yan (2007) who explicitly stated that teachers have to consider the test-oriented nature while planning their lessons.

Reasons related to teacher beliefs.

As for the last category about he reasons for textbook adaptation, all instructors were asked if they were driven to make adaptations based on their own beliefs and understanding. The great majority reported students' engagement and their performance were highly influential for their adaptive decisions as displayed in the Excerpt13 and Expert 14 taken from the interviews:

> *Excerpt 13: [...] I try to add different activities during my lesson to increase production and make sure my students actively participate in the lesson. (Experienced instructor 4, Interview data, 7ᵗʰ April 2017)*

> *Excerpt 14: [...] I usually adapt textbooks by adding an activity to increase learner involvement and create an atmosphere where they can use the language in different contexts. (Novice instructor 7, Interview data, 10ᵗʰ April 2017)*

Students' interest was another important reason when the instructors shared their beliefs

as professionals regardless of their years of teaching experience. This finding is supported in Excerpt 15:

> *Excerpt 15: [...] My students get bored quickly. That's why I try to adapt activities to attract their interests by modifying or adding. (Experienced instructor 3, Interview data, 9ᵗʰ April 2017)*

As a final remark, Figure 2 illustrates how the findings of the reflective essays and semi-structured interviews overlap and support each other. The results of the reflective essays represent instructors' self-reported beliefs about why textbook adaptation is considered as a rewarding practice, while the semi-structured interview results list the reasons behind the instructors' adaptive decisions.
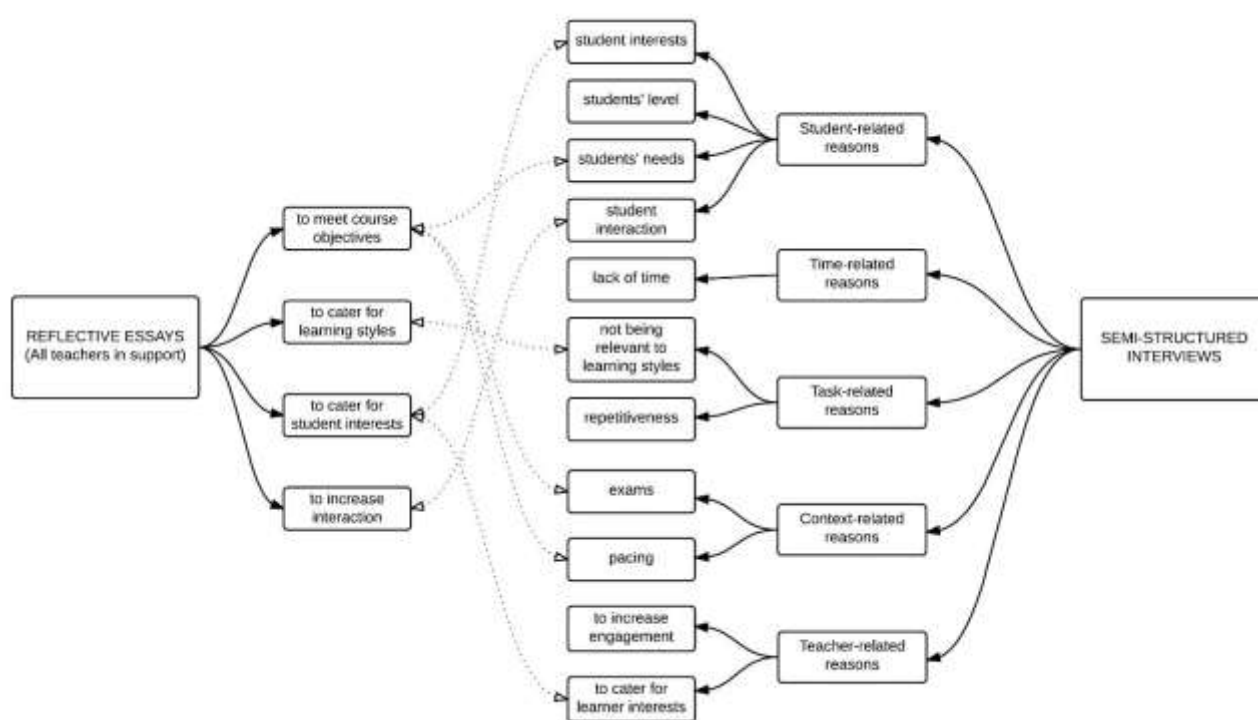

Figure 2. EFL Instructors' Self-Reported Beliefs about Textbook Adaptation and the Reasons for their Adaptive Decisions

According to Figure 2, it is obvious that reasons resulting from the understanding and beliefs of novice and experienced instructors were dominant while implementing adaptation strategies in the preparatory classroom. The two groups stated that certain adaptations were made to increase the language performance and engagement of their students. This finding might be related to the fact that students learn more effectively when they are actively involved in the lesson which was highlighted in Bailey's (1996) study as well.

Conclusion

The present study aimed to shed light on the self-reported beliefs of experienced and novice instructors about utilizing adaptation strategies in intermediate (B2 level) English classes of a preparatory program. The findings revealed that both group of instructors prefer to adapt activities in their classroom practices apart from some differences related to the types strategies used and their frequency. Regardless of their years of teaching experience, all of the participating instructors adapted their activities to meet the students' needs and interest in the course.

Considering the gathered data, this research has some pedagogical implications to be taken into consideration. First of all, all instructors had an opportunity to use various adaptive strategies in their lesson which helped them to gain more experience on utilizing textbook adaptation strategies in language classrooms. Therefore, the results of this study can be used to design an in-service training program to be offered both to the experienced and novice EFL teachers. Besides, the material designers can take the adaptation reasons provided in this study as important clues to develop materials for intermediate level English classes. In a similar vein, textbook writers may draw a good number of implications related to what reasons cause teachers to adapt their textbooks. This may help them to design materials to cater for different language needs, interest and objectives.

On the other hand, this study includes some limitations. One of them is related to sample size which was particularly small due to the heavy workload of the other participants. A larger number of participants could provide more generalizable results. Selection of the sample is another limitation of this study. For the purposes of this study, the instructors were selected only from one classroom (intermediate, B2 level). Different results could emerge if instructors teaching different proficiency levels were selected as participants in this study. Therefore, all these findings should be taken as suggestive rather than definitive.

Apart from certain limitations, the present study offers some recommendations for further research. First, this study investigated the use of textbook adaptation in one particular (intermediate, B2 level) preparatory language classroom. Future research could be conducted with a larger sample from different proficiency levels groups to obtain comparative and in-depth results. The actual practices of the participating instructors regarding their adaptive decisions could not be observed in this study due to the instructors' loaded schedule. A different study integrating classroom observations might provide an opportunity for analyzing teachers' in-class practices in detail.

To conclude, the present study attempted to make contribution to the field of language education providing some evidence about the implementation of adaptations strategies by novice and experienced instructors in a language preparatory classroom. The findings revealed important pedagogical implications and recommendations that may be of benefit for practitioners, material designers, and textbook writers in terms of the tole of textbook adaptation in English language preparatory programs.

References

Akyel, A. (1997). Experienced and student EFL teachers' instructional thoughts and actions. *Canadian Modern Language Review, 58*(4), 677-692. doi: 10.20880/kler. 2010.38.233.

Bailey K. M. (1996). The best laid plans: teachers' in-class decisions to depart from their lesson plans. In K. M. Bailey & D. Nunan (Eds.), V*oices from the language classroom* (pp. 15-40). Cambridge: Cambridge University Press.

Bigelow, L. (2000). Class planning strategies of expert and novice teachers. *The Sloping Halls Review, 7*, 79-88. Retrieved from http://repository.cmu.edu/cgi/viewcontent.cgi?article=1058&context=shr

Botelho, M. (2003). *Multiple intelligences theory in English language teaching: An analysis of current textbooks, materials and teachers' perceptions* (Master's Thesis). Ohio University, Ohio. Retrieved on August 20, 2016 from: https://etd.ohiolink.edu/

Çoban, Z. Z. (2001). *Novice and experienced English language teachers' use of textbook adaptation strategies at Gazi University* (Unpublished Master's Thesis). Gazi University, Turkey.

Ebrahimpourtaher, A. & Hamidi, E. (2015). Authenticity and adaptation in EFL materials. *International Archive of Applied Sciences and Technology, 6*(2), 34-39.

Edge, J., and S. Wharton. (1998). Autonomy and development: Living in the materials world. In B. Tomlinson (Ed.), *Materials development in language teaching* (pp. 295–311). Cambridge: Cambridge University Press.

Garinger, D. (2002). Textbook selection for the ESL classroom. *Center for Applied Linguistics*

Digest 2(10). Retrieved from http://www.cal.org/resource-center/briefs-digests/digests/(offset)/105

Graves, K. (2000). *Designing language courses: A guide for teachers.* Boston: Heinle & Heinle.

Harmer, J. (1998). *How to teach English.* Harlow, UK: Longman.

Islam,C. & Mares, C. (2003). Adapting classroom materials. In B. Tomlinson (Ed.), *Developing materials for language teaching* (pp. 86-103). London: Continuum.

Madsen, H. S., & Bowen, J. D. (1978). *Adaptation in language teaching.* Rowley Mass: Newbury House.

Maley, A. (1998). Squaring the circle-reconciling materials as constraint with materials as empowerment. In B. Tomlinson (Ed.), *Materials development in language teaching* (pp. 279-294). Cambridge: Cambridge University Press.

McDonough, J. and Shaw, C. (1993). *Materials and methods in ELT.* Oxford: Blackwell.

McGrath, I. (2002). *Materials evaluation and design for language teaching.* Edinburgh: Edinburgh University Press.

Mısırlı, S. (2010). Materials in TEFL: A discussion of what lies behind them and implications. *Humanising Language Teaching, 12*(4). Retrieved from http://www.hltmag.co.uk/aug10/sart04.htm

Miles, M. B. & Huberman, A. M. (1994). *Qualitative data analysis* (2nd ed.). Thousand Oaks, CA: Sage Publications.

Shavelson, R., & Stern, P. (1981). Research on teachers' pedagogical thoughts, judgments, decisions and behavior. *Review of Educational Research, 51*(4), 455-498. doi: 10.3102/00346543051004455

Sheldon, L. (1987). *ELT textbooks and materials. Problems in evaluation and development.* Oxford: Modern English Publications.

Tsobanoglou, S. (2008). *What can we learn by researching the use of textbooks and other support materials by teachers and learners?* (Master's Thesis). University of Nottingham, Nottingham.

Tsui, A. B. M. (2003). *Understanding expertise in teaching: Case studies of second language teachers.* Cambridge: Cambridge University Press.

Ur, P. (1996). *A course in language teaching.* Cambridge: Cambridge University Press.

Yan, C. (2007). Investigating teachers' materials adaptation. *Humanising Language Teaching, 9*(4), 1-15. Retrieved from http://www.hltmag.co.uk/jul07/mart01.htm

Yin, R. K. (2002). *Case study research: Design and methods.* Thousand Oaks, CA: SAGE Publications.

Zainal, Z. (2007). Case study as a research method, *Journal of Kemanusiaan, 9*(1), 1-6.

About the authors:

Enisa Mede is an Assistant Professor and the chair at the Faculty of Educational Sciences, Department of English Language Teaching (ELT), Bahcesehir University, Istanbul, Turkey. She has been offering Teaching Language Skills, Teaching English to Young Learners and Teaching Practicum courses at the undergraduate level and Syllabus Design, Program Evaluation in Language Education and Second Language Acquisition courses at the graduate level. Her chief research interests are program design and evaluation in language education, first/second language development in young learners and bilingual education.

Şenel Yalçın is an EFL instructor at the English Language Preparatory School, Bahcesehir University, Istanbul, Turkey. She has been teaching English for seven years. She graduated from the Department of English Language Teaching (ELT) and received her Master's Degree in English Language Education. Her research interests are material design and evaluation, teacher education and second language acquisition.