Reading-writing Connections in EAP Courses: Cross-linguistic Summary Protocols

Abbas Emam
Shahid Chamran University of Ahvaz, Iran
Email: sa_ab_emamrood@yahoo.co.uk

Abstract—At higher education settings, summarization skills are essential to academic success. However, research on summarization appear to have been relatively neglected, and consequently "many more are needed … [and] a re-examination of summarization from a reading-writing perspective merits more attention" (Grabe, 2003, pp. 252-253). To this end, 120 EFL students of Shahid Chamran University of Ahvaz, Iran were selected: the three experimental groups were subjected to the task-based language teaching principles in the treatment. Two reading comprehension texts (one as a pre-test, the other a post-test)were assigned to the four groups involved, and they were all required to write up two summary protocols for each text they read, one in English, and another in Persian. In the experimental groups, summary writing was mediated by a particular teaching task while for the control group there was no mediating task. Afterwards, the performance of all subject groups was evaluated based on a cross-linguistic evaluation scale to determine a) if task-based language teaching had any significant effect on learners' reading comprehension and b) if the L1 of the subjects contributed to a better comprehension of the texts they read. Finally, the conclusions and pedagogical implications of the research for EAP courses were highlighted.

Index Terms—reading-writing connections, EAP, reading comprehension, summary writing, mediating task

I. INTRODUCTION

In everyday use of language, we are continually integrating the language skills or switching from one skill to another. It is best to reflect this integration when teaching a second or foreign language (Davies & Pearse, 2000, p. 99). As far as the integration of reading- writing skills is concerned, it is admittedly an old, established technique to ask EAP students of reading comprehension courses to present or produce brief statements of the main ideas in a text or reading passage, either while reading or after completing reading the text. The ability to produce a summary (alternatively known as synopsis, or précis) is referred to in TEFL literature as summarizing, or summary skills, and has been a focus of instruction in the teaching and testing of reading comprehension skills.

At academic settings in particular, summarizing skills or techniques are essential to academic success. Students are usually required to produce study summaries, to complete various types of summary assignments, and to complete tasks that call for the incorporation of a written source material in term papers or any other similar presentation (Johns & Mayes, 1990; Kirkland & Saunders, 1991; Nunan, 2003; Brown, 2004). However, this is not an activity to be expected to be carried out with the same degree of success by all language learners. After all, some more proficient language learners outperform the others and some others, with less background and practice in these skills, lag behind.

Reading-writing connections and interactions can be studied from different perspectives, and with regard to a number of relevant factors and variables. Researchers have elaborated on the numerous approaches to the study of reading–writing connections in EAP courses and have reported on the findings , in theory , research as well as in practice (See, for example Farahzad & Emam, 2010).

This empirical study in its own turn was intended to find answers to the following questions based on a comparison of the content of the source reading texts with the English summary protocols (ESPs) and Persian summary protocols (PSPs) produced under three mediating task conditions.

1. Did the presence of mediating tasks in the three experimental groups, and their absence in the control group lead to any significant difference in reading comprehension among the Iranian EFL learners?

2. On which set of summary protocols did the subject groups outperform the others in terms of the selection of the topic(s) in their summary protocols, the ESPs or the PSPs?

3. On which set of summary protocols did the subject groups outperform the others in terms of the selection of the sub-topic(s) in their summary protocols, the ESPs or the PSPs?

4. Did the summary protocols written in the L1 of the subjects (Persian) manifest a significant difference from those produced in their target language (English) in terms of reading comprehension ability?

II. METHODOLOGY

The cornerstone of this research was to evaluate the effectiveness of the role of the "mediated tasks" of "group discussion", "personalizing", and "clustering" in reading comprehension of EAP students as manifested in their written

summaries (i.e. the interface of reading-writing connections).

A.   Participants

The study involved 120 Iranian EFL, male (30) and female (90), undergraduate students majoring in psychology at Shahid Chamran University of Ahvaz, Iran. The subjects who were in four classes of 30 students each included sophomores who had already passed the basic course of general English. The reason for adopting such a policy in selecting the subjects was to make sure they were sufficiently familiar with strategies, techniques and mechanics (the necessary pre-requisites) needed for carrying out the tasks envisioned for their specific groups.

B.   Treatment

The participants of the study were exposed to special treatment for one semester. Since the main objective of the study was to evaluate the reading comprehension ability of the subjects as reflected in a "Persian written summary task" as well as an "English written summary task", all the four groups were exposed to identical reading texts both during the semester and at the final exam session. However, the mediating tasks between the input (the reading text) and the outcomes (two written summary protocols in each session) were different for each group.

1. The DTG (the discussion-task group)

For the first experimental group, the mediating task to follow reading was a" group discussion" of the content of the text among the members of the group who were subsequently required to write two summaries for the passage they had just read.

2. The PTG (The personalizing-task group)

As far as the second experimental group was concerned, the mediating task to be made use of, after the reading selection was read by the learners, was the "personalizing task". A highly recommended task in communicative language teaching, task-based language teaching, cooperative language learning, and content-based instruction, personalizing is thought to draw on the learners' experiences and opinions and make the teaching material more real and accessible to them.

3. The CTG (The clustering-task group)

The mediating task envisioned for use in the third experimental group was "clustering". At times, referred to as "webbing", "word mapping", "semantic mapping", or "visual brainstorming", "clustering" task is recognized and recommended as an effective reading comprehension and pre-writing teaching/learning activity (Brown, 2001; Nunan, 2003; Broukal, 2004; Rodriguez Kessler, 2006).

4. The NTG (The no- mediating-task group)

However, the fourth group (the control group) received no specific task-based treatment. Learners in this group were simply asked to read the text (the input) and summarize it in Persian and in English within the time limit of the class hour, as they were considered most appropriate for completion of the two protocols.

TABLE 3-1:
GRAPHIC REPRESENTATION OF THE DESIGN OF THE RESEARCH

Groups (n=120)	Input	Mediating Task	Outcome
Grp. 1 (n=30)	Reading text	Group Discussion Task	1.A Persian written summary 2. An English written summary
Grp. 2 (n=30)	Reading text	Personalizing Task	1.A Persian written summary 2. An English written summary
Grp. 3 (n=30)	Reading text	Clustering Task	1.A Persian written summary 2. An English written summary
Grp. 4 (n=30)	Reading text	None	1.A Persian written summary 2. An English written summary

C.   Materials

The main criteria for selecting, analyzing and assigning the reading materials for this specific EAP course (used during the treatment period) were in line with what Nuttal (1996, pp. 170-178) suggests: suitability of the content ,exploitability, readability, variety, and authenticity. To that effect, of the many English reading comprehension textbooks available in the Iranian book market, the present author chose to concentrate on Academic Encounters (Reading , study skills, and writing: content focus, human behavior) by Bernard Seal, published by Cambridge University Press in 1997. The source which consists of 40 authentic reading passages taken from regular North American college textbooks aims to develop reading and study skills among EFL/ESL students.

D.   Measurement Instruments and Procedure

Three instruments were used in this study:

1. On administering the independent measure of proficiency (the TOEFL) to all subject groups, it was found out that they were sufficiently homogenous with regard to their scores (of 100): DTG = 61.35, PTG=60.5, CTG=62.1, and NTG=61.25.

2. A pretest which consisted of an English reading passage of both medium length (approximately 800 words) and level of readability (Flesch – Kincaid Grade Level 11). The subjects were required to read it first and subsequently write

two summaries(one in Persian and another in English)in 60 minutes(see Appendix I)after being exposed to the particular treatment task envisioned for their particular treatment group. The reason this pre-test was administered was firstly to do with the fact that "summary writing tests writing skills as well as reading skills" (Nuttal, 1996, p.225), and it secondly had to do with the final objective of the research, i.e. to analyze the content of the written summaries that subjects produced as measurement protocols for reading comprehension assessment. As for the rationale of recourse to the Persian language in writing the summary protocol, it is argued that "students who are permitted to use their L1 in responding will explore the text more accurately and thoroughly than those who are restricted to target language responses, [thus] it is important not to dismiss the use of the mother tongue out of hand. Some students may never need to express themselves in the target language. Why should they not respond in the language that most clearly enables them to show that they understand, or to explain where their problems lie?" (Nuttal, 1996, p. 187). Similar opinions have also been expressed by a number of other researchers (e.g., Shohamy, 1985; Lee, 1986, 1987; Wolf, 1993; Hock & Poh, 1997).

3. The third instrument was a post-test which actually served as the final exam of the course. This particular test (corresponding and parallel to the pre-test above) was supposed to serve as the main measurement instrument in the assessment of the reading comprehension ability of the subjects before and after receiving the particular treatments to which they were exposed. The full text of the post-test reading comprehension (Flesch– Kincaid Grade Level 12) is reproduced in Appendix I.

III. DESCRIPTIVE AND INFERENTIAL STATISTICS

The performance of each of the four subject groups (the DTG, the CTG, the PTG, and the NTG), on both English and Persian summary protocols are cross-tabulated below. To make the comparison and contrast of the results easier, in each case tables representing the descriptive indices of performance(Means and SDs) of each particular group on both the pre-test and the post-test and the resultant difference(s) are juxtaposed. Here, the tables are identified as those for ESPs (English Summary Protocols) and their corresponding PSPs (Persian Summary Protocols).Meanwhile, in each pair of tables under the designation "variable" a number of Ps (P1, P2, P3,…p'1, p'2, p'3, …) are introduced, each to stand for a particular point (P) targeted on the English and Persian Evaluation Scales (See Appendix II), and these Ps and P's are taken into account in the design of this dissertation as the dependent variable(s).

To more clarify the descriptive statistical indices-and hence the performance of each paired subject groups identified as ESPs and PSPs-the resultant figures were then further subjected to inferential indices of performance.

To illustrate, as far as the first research question is concerned, the descriptive statistics yielded the following results:

TABLE 1-A:
TOTAL MEANS AND STDs OF SCORES FOR ESPs ACROSS SUBJECT GROUPS ON THE
PRE – TEST AND POST – TEST

Variable	Phase	Statistical Indices → Groups	Means	Std. Deviations	N=120
Total ESPs	Pre – test	DTG	80.67	30.147	30
		CTG	65.41	25.532	30
		PTG	64.88	21.480	30
		NTG	69.78	30.945	30
	Post - test	DTG	83.02	30.888	30
		CTG	88.26	31.640	30
		PTG	68.50	29.918	30
		NTG	60.07	38.996	30
	Mean Differences	DTG	-2.35	32.3875	30
		CTG	-22.85	30.086	30
		PTG	-3.62	26.999	30
		NTG	9.71	36.9705	30

TABLE 1-A':
TOTAL MEANS AND STDs OF SCORES FOR PSPs ACROSS SUBJECT GROUPS ON THE
PRE – TEST AND POST –TEST

Variable	Phase	Statistical Indices → Groups	Means	Std. Deviations	N=120
Total PSPs	Pre – test	DTG	87.96	26.319	30
		CTG	76.41	22.442	30
		PTG	72.43	40.557	30
		NTG	78.56	42.141	30
	Post - test	DTG	86.92	33.818	30
		CTG	85.87	31.655	30
		PTG	85.92	17.242	30
		NTG	63.73	41.358	30
	Mean Differences	DTG	1.04	31.8185	30
		CTG	-9.46	28.0485	30
		PTG	-13.49	30.1495	30
		NTG	14.83	43.7495	30

As it can be seen in the paired tables above, the total differences in mean scores of all Ps in the ESPs indicate that the highest level of performance is observed in the CTG (-22.85) and the lowest level in the NTG (9.71). As for the total differences in mean scores of all P's in the PSPs, it is also clear that the highest level of performance is observed in the PTG (-13.49) and the lowest level in the NTG (14.83). This may tempt one to tentatively conclude that the absence of any mediating tasks in the control group (NTG) led to the lowest level of performance. However, this does not seem to be statistically warranted.
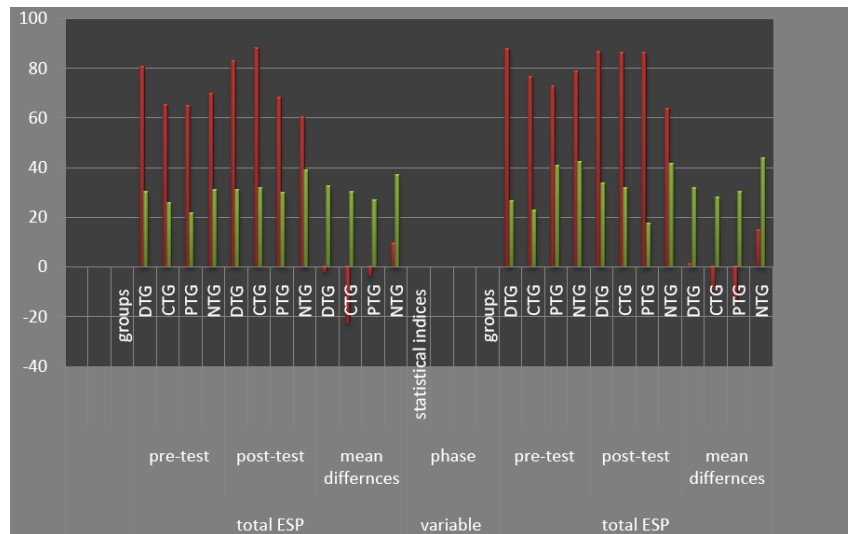
Figure 1: Graphic Representation of the Performance of the Subject Groups on ESPs Vs PSPs

In order to see whether or not the differences among the groups are statistically significant, the one-way ANOVA procedure was run. The results of the ANOVA procedure for the ESPs and the PSPs are given below in tables 3-1-b and 3-1-b'.

TABLE 1-B:

THE RESULTS OF TOTAL ANOVAS ON ESP$_s$ POST – TESTS

Source of Variance	Sum of Squares	df	Mean Squares	F - ratio	Sig
Between Groups	15159.426	3	5053.142	4.624	.004
Within Groups	126756.17	116	1092.726		
Total	141915.60	119			

TABLE 1-B':

THE RESULTS OF TOTAL ANOVAS ON PSP$_s$ POST– TESTS

Source of Variance	Sum of Squares	df	Mean Squares	F - ratio	Sig
Between Groups	11415.548	3	3805.183	3.665	.014
Within Groups	120451.19	116	1038.372		
Total	131866.74	119			

The results of total ANOVA on ESPs post-tests as reflected in their F-ratio (4.624) and significance (0.004) belonging to the means of the groups indicate that because the level of significance is lower than %5,it follows that there is a statistically significant difference among the subject groups. To locate the differences between the groups, the post-hoc Tukey's HSD was applied resulting in table 1-c. A similar survey of the results of the total ANOVAs on PSPs post-tests with F-ratio (3.655) and statistical significance 0.014 belonging to the means of the groups indicate that here again because the level of significance is lower than %5,it can be inferred that there is a statistically significant difference among the subject groups of PSPs as well. To locate the differences between the groups, the post-hoc Tukey's HSD was applied which resulted in table 1-c'.

TABLE 1-C:

THE RESULTS OF TUKEY`S HSD ON TOTAL ESP$_s$ POST – TESTS

Groups	Mean	DTG	CTG	PTG	NTG
DTG	-2.35	-	-	-	*
CTG	-22.85	-	-	-	*
PTG	-3.62	-	-	-	-
NTG	9.71	*	*	-	-

TABLE 1-C':

THE RESULTS OF TUKEY`S HSD ON TOTAL PSPs POST – TESTS

Groups	Mean	DTG	CTG	PTG	NTG
DTG	1.04	-	-	-	*
CTG	-9.46	-	-	-	*
PTG	-13.49	-	-	-	*
NTG	14.83	*	*	*	-

The differences among the ESPs and PSPs thus confirmed, one more procedure was required to demonstrate on which set did the subject groups outperformed the other, on the ESPs or on the PSPs? This demanded applying the T-test to determine the statistical significance of the difference between the means on the two paired sets of scores. Once the T-test was applied to the two paired sample groups (Tables 1-d and -1-e, below), it turned out that the

difference in means between the two sets (ESPs vs. PSPs) were 16.9056 vs. 18.4778 with significance levels of 0.043 and 0.043,respectively.This clearly signifies the higher performance of the subject group producing PSPs. Thus, as far as these particular sets of groups are concerned, the first research question is safely answered positively in favor of the experimental groups in which particular tasks were used.

TABLE 1-D:
PAIRED SAMPLES STATISTICS

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Total ESPs PTs	74.96	120	34.534	3.152
	Total PSPs PTs	80.61	120	33.288	3.039

TABLE 1-E:
PAIRED SAMPLE TEST

	Paired Differences Mean	t	df	Sig. (2-tailed)
Total ESPs PT – Total PSPs PT	-5.650	-1.914	119	.041

TABLE 2-A:
MEANS AND STD$_S$ OF SCORES FOR ESP$_2$ ACROSS SUBJECT GROUPS ON THE PRE – TEST AND POST – TEST

Variable	Phase	Statistical Indices Groups	Means	Std. Deviations	N=120
P$_2$	Pre – test	DTG	15.56	4.004	30
		CTG	14.30	4.584	30
		PTG	12.81	4.966	30
		NTG	13.27	7.311	30
	Post - test	DTG	19.31	13.194	30
		CTG	18.91	6.184	30
		PTG	16.24	5.139	30
		NTG	12.93	7.494	30
	Mean Differences	DTG	-3.75	10.489	30
		CTG	-4.61	6.384	30
		PTG	-3.43	6.0525	30
		NTG	0.34	9.4025	30

TABLE 2-A':
MEANS AND STD$_S$ OF SCORES FOR PSP$_2$ ACROSS SUBJECT GROUPS ON THE PRE – TEST AND POST – TEST

Variable	Phase	Statistical Indices Groups	Means	Std. Deviations	N=120
P$_2'$	Pre – test	DTG	16.43	4.492	30
		CTG	14.99	3.951	30
		PTG	16.69	20.315	30
		NTG	14.98	13.340	30
	Post - test	DTG	20.46	14.003	30
		CTG	21.51	11.688	30
		PTG	20.59	8.718	30
		NTG	12.51	7.687	30
	Mean Differences	DTG	-4.03	10.4975	30
		CTG	-6.52	8.8195	30
		PTG	-3.9	16.5165	30
		NTG	2.47	12.2635	30

Regarding the second research question, the figures contained in Tables 2-a vs. 2-a' representing the differences in mean scores of all Ps in ESP2 demonstrate that the highest level of performance in selection of the topic(s) of the post-test reading passage belongs to the CTG (-4.61), and the lowest level to the NTG (0.34). As for the differences in mean scores of all P's in PSP2, it is also evident that the highest level of performance is observed in the CTG (-6.52) and the lowest level in the NTG (2.47).
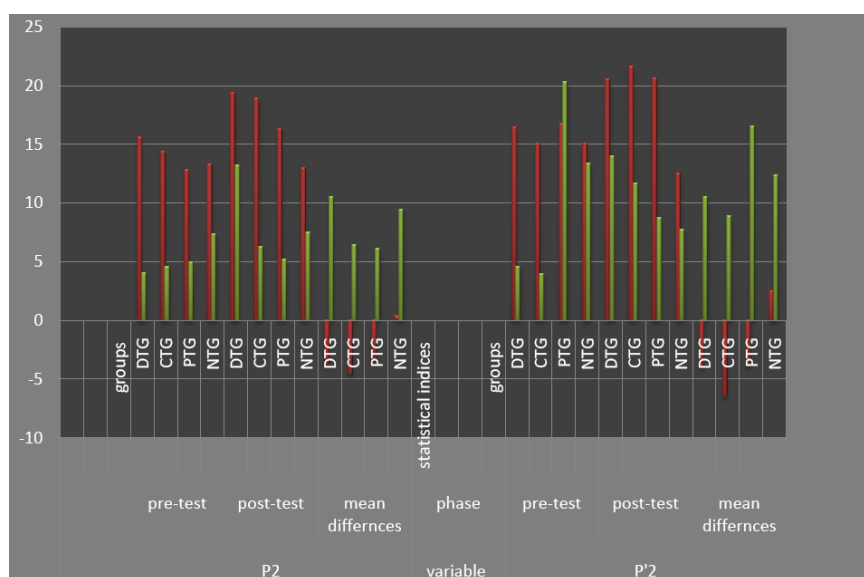

Figure 2: Graphic Representation of the Performance of Subject Groups on ESP2 vs. PSP2

Nevertheless, in order to see whether or not the differences among the groups are statistically significant, the one-way ANOVA procedure was run. The results of the ANOVA procedure for the ESP2 and PSP2 are given below in Tables 2-b and 2-b'.

TABLE 2-B:
THE RESULTS OF ANOVA ON ESP₂ POST – TEST

Source of Variance	Sum of Squares	df	Mean Squares	F - ratio	Sig
Between Groups	780.367	3	260.122	3.528	.017
Within Groups	8552.044	116	73.725		
Total	9332.411	119			

TABLE 2-B':
THE RESULTS OF ANOVA ON PSP₂ POST – TTEST

Source of Variance	Sum of Squares	df	Mean Squares	F - ratio	Sig
Between Groups	1585.104	3	528.368	4.518	.005
Within Groups	13565.919	116	116.948		
Total	15151.022	119			

According to table 2-b, the results of ANOVA procedure on mean differences of ESP2 post-test as reflected in the F-ratio of 3.528 and significance level of 0.017 (sig<%5) reveals that because the level of statistical significance is lower than %5, it can logically be argued that there is a statistically significant difference among the subject groups. To locate where the differences between the groups lie, the post-hoc Tukey's HSD was applied resulting in table 2-c. A similar examination of the results of ANOVA on PSP2 post-test with F-ratio of 4.518 and statistical significance 0.005 belonging to the means of the groups indicate that in this case too because the significance level is lower than %5, it can be asserted that there is a statistically significant difference among the subject groups of PSP2. To locate the differences between the groups, the Tukey's HSD was applied which led into Table-2-c'.

TABLE 2-C:
THE RESULTS OF TUKEY`S HSD ON ESP₂ POST – TEST

Groups	Mean	DTG	CTG	PTG	NTG
DTG	-3.75	-	-	-	*
CTG	-4.61	-	-	-	*
PTG	-3.43	-	-	-	-
NTG	0.34	*	*	-	-

TABLE 2-C':
THE RESULTS OF TUKEY`S HSD ON PSP₂ POST – TEST

Groups	Mean	DTG	CTG	PTG	NTG
DTG	-4.03	-	-	-	*
CTG	-6.52	-	-	-	*
PTG	-3.9	-	-	-	*
NTG	2.47	*	*	*	-

The differences between the ESP2 and PSP2 thus confirmed one more procedure was required to prove on which set did the subject group outperformed the other, on the ESP2 or on the PSP2. This demanded applying the T-test to determine the statistical difference between the means on the two paired sets of scores. Once the T-test was applied to the two paired sample groups (Tables 2-d and 2-e, below), it was discovered that the difference in means between the two sets (ESP2 vs. PSP2) were 16.85 vs. 19.42 with significance levels of 0.048 and 0.049. This indicates that there is a significant difference in performance between ESP2 vs. PSP2. Thus, as far as these particular sets of groups are concerned, the second research question as to the selection of the topic(s) is answered in favor of the subject groups in PSP2.

TABLE 2-D:
PAIRED SAMPLE STATISTICS

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	ESP2(PT)	16.85	120	8.856	.808
	PSP2(PT)	18.77	120	11.284	1.030

TABLE 2-E:
PAIRED SAMPLE TEST

	Paired Differences Mean	t	df	Sig. (2-tailed)
Pair 1 Past Q2 - Past N2	-1.917	-2.261	119	.026

TABLE 3-A:
MEANS AND STDₛ OF SCORES FOR ESP₃ ACROSS SUBJECT GROUPS ON THE PRE – TEST AND POST – TEST

Variable	Phase	Statistical Indices → Groups	Means	Std. Deviations	N=120
P₃	Pre – test	DTG	14.54	4.109	30
		CTG	12.50	5.227	30
		PTG	11.32	4.772	30
		NTG	12.79	6.446	30
	Post - test	DTG	13.87	4.962	30
		CTG	18.13	6.230	30
		PTG	15.83	7.183	30
		NTG	12.28	6.030	30
	Mean Differences	DTG	0.67	5.7655	30
		CTG	-5.63	7.5185	30
		PTG	-4.51	7.8275	30
		NTG	0.51	8.238	30

TABLE 3-A':
MEANS AND STDₛ OF SCORES FOR PSP₃ ACROSS SUBJECT GROUPS ON THE PRE – TEST AND POST – TEST

Variable	Phase	Statistical Indices → Groups	Means	Std. Deviations	N=120
P₃'	Pre – test	DTG	15.08	4.328	30
		CTG	16.00	4.883	30
		PTG	12.99	4.446	30
		NTG	13.12	7.969	30
	Post - test	DTG	20.30	8.657	30
		CTG	18.30	5.621	30
		PTG	18.84	5.075	30
		NTG	13.32	7.597	30
	Mean Differences	DTG	-5.22	7.7425	30
		CTG	-2.3	6.252	30
		PTG	-5.85	5.7605	30
		NTG	-0.2	9.783	30

With respect to the third research question, the figures contained in tables 3-a vs. 3-a' representing the differences in

mean scores of all Ps in the ESP3 reveal that the highest rate of performance in selection and identification of subtopic(s) of the reading passage belongs to the CTG (-5.63), and the lowest rate of performance to the NTG (-0.2). As for the differences in mean scores of all P's in the PSP3, it is clear that the highest level of performance is observed in the PTG (-5.85) and the lowest level in the NTG (-0.51).
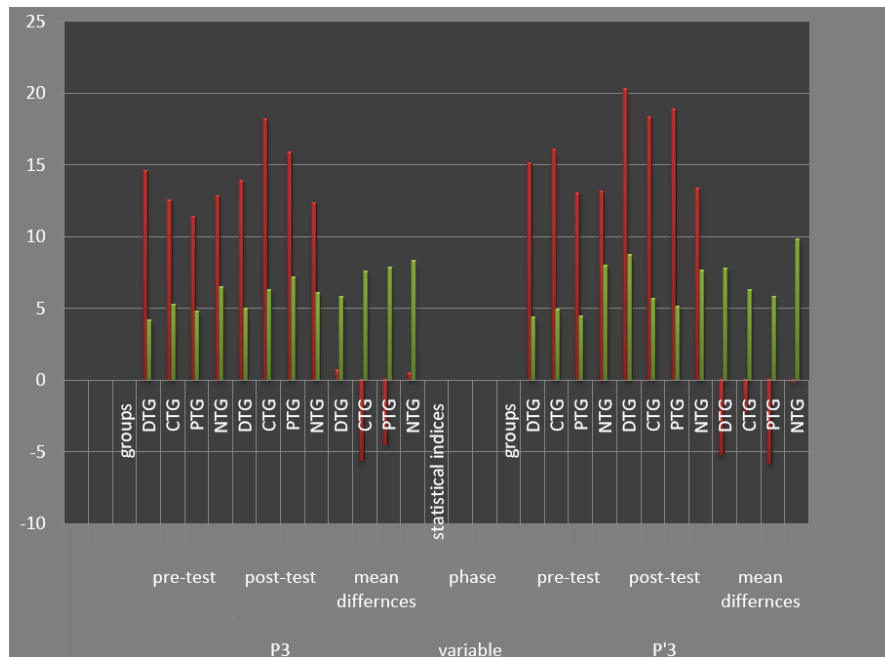

Figure 3: Graphic Representation of the Performance of Subject Groups on ESP3 vs. PSP3

However, in order to see whether or not the differences among the groups are statistically significant, the one-way ANOVA procedure was run. The results of the ANOVA procedure for the ESP3 and PSP3 are given below in Tables 3-3-b and 3-3-b'.

TABLE 3-B:
THE RESULTS OF ANOVA ON ESP$_3$ POST − TEST

Source of Variance	Sum of Squares	df	Mean Squares	F - ratio	Sig
Between Groups	576.122	3	192.041	5.074	.002
Within Groups	4390.452	116	37.849		
Total	4966.574	119			

TABLE 3-B':
THE RESULTS OF ANOVA ON PSP$_3$ POST − TEST

Source of Variance	Sum of Squares	df	Mean Squares	F - ratio	Sig
Between Groups	827.832	3	275.944	5.809	.001
Within Groups	5510.426	116	47.504		
Total	6338.258	119			

According to Table 3-b, the results of ANOVA procedure on mean differences of ESP3 post-test as reflected in the F-ratio of 5.074 and the significance level of 0.002 (sig<%5) indicate that because the level of statistical significance is lower than %5,it can logically be argued that there is a statistically significant difference among the subject groups. To locate where the differences between the groups lie, the post-hoc Tukey's HSD was applied resulting in Table 3-c. A similar examination of the results of ANOVA on the PSP3 post-test with F-ratio of 5.809 and the statistical significance of 0.001 belonging to the means of the groups reveal that in this case because the significance level is lower than %5, it can be asserted that there is a statistically significant difference among the subject groups of PSP3. To locate the differences between the groups, the post-hoc Tukey's HSD was applied which resulted into Table 3-c'.

TABLE 3-C:
THE RESULTS OF TUKEY`S HSD ON ESP$_3$ POST − TEST

Groups	Mean	DTG	CTG	PTG	NTG
DTG	0.67	-	-	-	-
CTG	-5.63	*	-	-	*
PTG	-4.51	-	-	-	-
NTG	0.51	-	*	-	-

TABLE 3-C':
THE RESULTS OF TUKEY`S HSD ON PSP$_3$ POST − TEST

Groups	Mean	DTG	CTG	PTG	NTG
DTG	-5.22	-	-	-	*
CTG	-2.3	-	-	-	*
PTG	-5.85	-	-	-	*
NTG	-0.2	*	*	*	-

The difference between the ESP3 and PSP3 thus confirmed, one more procedure was required to prove on which set did the subject groups outperformed the other, on the ESP3 or on the PSP3. This demanded applying the T-test to determine the statistical difference between the means on the two paired sets of scores. Once the T-test was applied to the two paired sample groups (Table3-d and 3-e, below) it was found that the difference in means between the two sets (ESP3 vs. PSP3) were 15.03 and 17.69 with significance levels of 0.003 and 0.003. This indicates that there is a

significant difference in performance between ESP3 vs. PSP3. Thus, as far as these particular paired sets of groups are concerned, the third research question as to the selection of the sub-topic(s) is answered in favor of the subject groups in the PSP3.

TABLE 3-D:
PAIRED SAMPLE STATISTICS

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	ESP3(PT)	15.03	120	6.460	.590
	PSP3(PT)	17.69	120	7.298	.666

TABLE 3-E:
PAIRED SAMPLE TEST

	Paired Differences Mean	t	df	Sig. (2-tailed)
ESP3(PT) – PSP3(PT)	-2.664	-3.926	119	.000

TABLE 4-A:
TOTAL MEANS AND STD$_S$ OF SCORES FOR ESP$_S$ ACROSS SUBJECT GROUPS ON THE PRE – TEST AND POST – TEST

Variable	Phase	Statistical Indices Groups	Means	Std. Deviations	N=120
P$_5$	Pre – test	DTG	80.67	30.147	30
		CTG	65.41	25.532	30
		PTG	64.88	21.480	30
		NTG	69.78	30.945	30
	Post - test	DTG	83.02	30.888	30
		CTG	88.26	31.640	30
		PTG	68.50	29.918	30
		NTG	60.07	38.996	30
	Mean Differences	DTG	-2.35	32.3875	30
		CTG	-22.85	30.086	30
		PTG	-3.62	26.999	30
		NTG	9.71	36.9705	30

TABLE 4-A':
TOTAL MEANS AND STD$_S$ OF SCORES FOR PSP$_S$ ACROSS SUBJECT GROUPS ON THE PRE – TEST AND POST – TEST

Variable	Phase	Statistical Indices Groups	Means	Std. Deviations	N=120
P$_5$'	Pre – test	DTG	87.96	26.319	30
		CTG	76.41	22.442	30
		PTG	72.43	40.557	30
		NTG	78.56	42.141	30
	Post - test	DTG	86.92	33.818	30
		CTG	85.87	31.655	30
		PTG	85.92	17.242	30
		NTG	63.73	41.358	30
	Mean Differences	DTG	1.04	31.8185	30
		CTG	-9.46	28.0485	30
		PTG	-13.49	30.1495	30
		NTG	14.83	43.7495	30

To answer the fourth and in fact the last research question, a careful look at Tables-4-a vs.4-a' above representing the differences in mean scores of all Ps in the ESPs reveals that the highest level of performance is observed in the CTG (-6) and the lowest level in the NTG (1.03). As for the differences in mean scores of all P's in the PSPs, it is evident that in this set of groups the highest performance belongs to the CTG (-4.19) and the lowest to the NTG (0.14).

Figure 3-4: Graphic Representation of the Performance of the Subject Groups on ESP4 vs. PSP4

Nevertheless, in order to see whether or not the differences among the groups are statistically significant, the one-way ANOVA procedure was run. The results of the NOVA procedure for the ESPs and PSPs are given below in Tables-4-b and 4-b'.

TABLE 4-B: THE RESULTS OF TOTAL ANOVAs ON ESPs POST – TESTS					
Source of Variance	Sum of Squares	df	Mean Squares	F - ratio	Sig
Between Groups	15159.426	3	5053.142	4.624	.004
Within Groups	126756.17	116	1092.726		
Total	141915.60	119			

TABLE 4-B': THE RESULTS OF TOTAL ANOVAs ON PSPs POST- TESTS					
Source of Variance	Sum of Squares	df	Mean Squares	F - ratio	Sig
Between Groups	11415.548	3	3805.183	3.665	.014
Within Groups	120451.19	116	1038.372		
Total	131866.74	119			

According to Tables 4-b, the results of total ANOVA procedure on the mean differences of the ESPs post-tests as reflected in the F-ratio of 6.624 and significance level of 0.004 (sig<%5) reveals that because the level of significance is lower than %5, it can reasonably be inferred that there is a statistically significant difference among the subject groups. To locate the differences between the groups, the post-hoc Tukey's HSD was applied resulting in Table 4-c. A similar examination of the results of ANOVA on PSPs post-tests with F-ratio of 3.665 and the statistical significance 0.014 belonging to the means of the groups indicate that in this case too because the significance level is lower than %5, it can be concluded that there is a statistically significant difference among the subject groups of PSPs. To locate the differences between the group, the Tukey's HSD was applied which resulted in Table 4-c'.

TABLE 4-C: THE RESULTS OF TUKEY'S HSD ON TOTAL ESPs POST – TESTS					
Groups	Mean	DTG	CTG	PTG	NTG
DTG	-2.35	-	-	-	*
CTG	-22.85	-	-	-	*
PTG	-3.62	-	-	-	-
NTG	9.71	*	*	-	-

TABLE 4-C': THE RESULTS OF TUKEY'S HSD ON TOTAL PSPs POST – TESTS					
Groups	Mean	DTG	CTG	PTG	NTG
DTG	1.04	-	-	-	*
CTG	-9.46	-	-	-	*
PTG	-13.49	-	-	-	*
NTG	14.83	*	*	*	-

The differences between the total ESPs and total PSPs thus confirmed, one more procedure was required to demonstrate on which set did the subject groups outperformed the other, on the ESPs or on the PSPs. This demanded applying the T-test to determine the statistical difference between the means on the two paired sets of scores. Once the Tukey test was applied to the two paired sample groups (Tables 4-d and 4-e, below), it was found that the difference in means between the two sets (ESPs vs. PSPs) were 16.9056 vs. 18.4778 with significance levels of 0.043 and 0.043, respectively. This clearly indicates that there is a significant difference in performance between the total ESPs vs. PSPs. Thus, as far as these particular paired sets of groups are concerned, the fourth research question regarding the degree of difference in reading comprehension ability of the ESPs vs. the PSPs (as reflected in their total difference indices) it can safely be claimed that the summary protocols produced in the L1 of the subjects (Persian) demonstrated a higher degree of reading comprehension ability vs. those protocols written in the TL (English). As a result, the last research question is answered positively in favor of the groups writing their summaries in their L1 (Persian).

TABLE 4-D:
PAIRED SAMPLES STATISTICS

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Total ESPs PTs	74.96	120	34.534	3.152
	Total PSPs PTs	80.61	120	33.288	3.039

TABLE 4-E:
PAIRED SAMPLE TEST

	Paired Differences Mean	t	df	Sig. (2-tailed)
Total ESPs PT – Total PSPs PT	-5.650	-1.914	119	.041

IV. DISCUSSION

A. Research Findings Corresponding to Research Question 1

Overall, the experimental groups (DTG,CTG, and PTG) outperformed the control group (NTG) .This implies that the groups in which particular mediated-tasks were used did much better (in varying degrees) on both ESPs and PSPs in comparison with the group in which no mediated-tasks were used between reading the passage and the written summary protocols that ensued. One can also argue that particularly in PSPs, gains were significantly higher than those for the ESPs and this is again confirming the viewpoint that summarizing in one's L1 tends to result in a better performance than summarization in an L2. Furthermore, the efficacy of a task-based language teaching approach to contributing to improving reading comprehension is also confirmed.

B. Research Findings Corresponding to Research Question 2

With respect to the difference in performance of the same sample population on ESPs vs. PSPs (regarding the identification of the topics in the reading passage), it was found that the difference in means between the two sets (ESP2 vs. PSP2) were 16.85 vs. 18.77 with significance level of 0.026. This indicated that there was a significant difference in performance between ESP2 vs. PSP2.Thus,as far as the particular sets of groups were concerned, the second research question as to the selection of the topic(s) was answered in favor of the subject groups in PSP2.This is a finding in support of an argument put forward by Nuttal (1982, p. 187) to the effect that students who are permitted to use their L1 in responding will explore the text more accurately and thoroughly than those who are restricted to target language responses.

C.   Research Findings Corresponding to Research Question 3

The difference in performance of the same sample population on ESPs vs. PSPs (on identification of the subtopic(s) in the reading passage assigned), it was discovered that the difference in means between the two sets (ESP3 vs. PSP3) were 15.03 vs. 17.69 with significance level of 0.000. This indicated that there was a significant difference in performance between what happened on the ESP3 vs. PSP3. As a result, with respect to these particular sets of subject groups, the third research question as to the selection of the subtopics was answered in favor of the subject groups in PSP3.This is again a finding confirming the fact that in identifying the subtopics too performing in one's mother tongue contributes to better performance than performing in a target language.

D.   Research Findings Corresponding to Research Question 4

As far as the last research question regarding the degree of difference in reading comprehension ability of the same sample population on the ESPs vs. PSPs is concerned (as reflected in their total mean difference indices) the following findings need to be taken into account. The difference in means between the two sets (ESPs vs. PSPs) were 47.96 vs. 80.61 with the significance level of 0.041. This clearly indicates that the summary protocols produced in the L1 of the subjects (Persian) demonstrated a higher degree of reading comprehension ability vs. those protocols written in the L2 (English). As a result, the last research question is answered positively in favor of the groups writing their summaries in their L1 (Persian).

In sum, of the four research questions formulated above, it can be stated that according to the findings available:

1. Across the subject groups in the study, the experimental groups (in which particular mediated tasks were used between the reading passage and the written summary protocols) outperformed the control group.

2. Across the subject groups in the study, the experimental groups whose summary protocols were written in Persian outperformed the other groups whose summary protocols were written in English.

V.   CONCLUSION

Based on the research findings mentioned above, from particular groups of EAP subjects and with a particular elicitation technique (cross-linguistic summary protocol) and under particular task conditions, a number of conclusions may seem warranted. First of all, the facts and figures related to the first research question highlight the facilitating effect of the mediating tasks in improving reading comprehension of a source text. In addition, it was empirically established that the L1 of the foreign language learner is not necessarily a hampering factor in improving reading comprehension among the adult second/foreign language learner in EAP courses. Furthermore, as far as the integration of reading and writing skills aimed at assessing reading comprehension of EAP students is concerned, an integrative orientation in teaching and learning English and in syllabus design was sufficiently justified. The two skills, if made use of appropriately, can be mutually supportive.

Looked at from another perspective, the outperformance of the subjects on the PSPs (and their relative poor performance on the ESPs) may suggest that we need to make a reasonable distinction between the "reading comprehension ability" and the "language proficiency" of the learners. To clarify the point, failing to write a well-developed summary protocol in the TL, by and in itself, does not imply insufficient reading comprehension ability on the part of the foreign language learners. They may not be well-versed in productive aspects of the TL (text length, syntactic complexity, lexical familiarity etc.); however, they may demonstrate that their reading comprehension ability is not flawed as might be expected in such cases. This conclusion appears to be in line with the view expressed by Koda (2005, p. 249) to the effect that "language-production skills generally lag behind comprehension capabilities, and task performance, if required in L2, can be seriously constrained by production, rather than comprehension liabilities." In another study, American college students studying Spanish produced, on average , longer and more accurate protocols in their L1 (Lee,1986) than a comparable group who conveyed their recall in L2 (Carrell,1983).

As such, post reading response construction subsumes comprehension, and it does not seem unlikely that successful achievement can be a reliable indicator of comprehension competence. Unsuccessful performance, on the other hand, reveals little about comprehension per se, because failed efforts, Koda (2005, p. 249) argues can be attributed to comprehension or to production limitations. In sum, language selection for auxiliary non-text elements is a unique problem in L2 reading assessment, particularly when the intent is gauging comprehension rather than language proficiency. Therefore, appearing incompetent in expressing oneself in L2 writing can not be taken as the sole determinant in reading comprehension. The current production-based techniques of reading comprehension assessment

with their preoccupation with production (e.g. summary protocol) apparently fail to fully tap the reading comprehension ability of L2 learners.

A corollary of the same issue is that based on the research findings reported here, it appears that in EAP courses, as far as reading comprehension in L2 is concerned, language learners typically tend to look at a text as a "vehicle for information", not as a "linguistic object". To them, the focus is predominantly on the information contained in the text rather than concentrating on every details of any linguistic significance. As a result, recourse to the L1 in reading comprehension in any L2 (whether receptively or productively) is not a phenomenon to frown upon.

VI. PEDAGOGICAL IMPLICATIONS OF THE RESEARCH

The pedagogical implications from this research are hoped to be of potential and/or practical value to the practitioners involved in EAP, in particular to those dealing with the instruction of reading comprehension skill or writing. Theoreticians in the relevant field may also find some useful hints and clues to boost further research. More specifically, the findings of the research are assumed to contribute to a better and more profound understanding of the following topics as well as to addressing them in curricular planning: a. task-mediation, b. skills-integration, c. summarization, d .cross-linguistic approach to EAP, e. content-based, and ESAP instruction.

VII. LIMITATIONS OF THE RESEARCH

The present study suffered from a number of limitations, some arising from the complexity and dynamics of the nature of reading-writing interconnections, and others from the limitations probably imposed by the conceptual framework adopted, sampling procedure, variables selected for the study, instrumentation, etc. Such factors clearly restrict the generalizability of the results. Some of these factors can be listed as follows:

To measure the reading comprehension ability of the EFL learners, summary protocols are not the only diagnostic tools for researchers to rely on. Other task formats, from discrete-point response modes, such as matching, true-false, multiple-choice and cloze deletion items, to global response modes such as recall protocols are among the other tools serving researchers. The present study aimed to concentrate on analyzing summaries as "products" of reading-writing connections; therefore, the "processes" involved in this respect were consciously excluded from the research design, because that can be the subject of another independent study.

Because of the time- limitation the researcher faced with, it is believed that a longitudinal study of the same issue may arrive at results partly or totally different from those obtained here. The number of participants had to be limited to around 120 to enable the researchers to take into account the problems of manageability and practicality. Therefore, care must be taken as to the comprehensiveness of the conclusions reached because they may not be applicable to a wider representative population.

APPENDIX I THE FULL TEXT OF THE PRE-TEST AND POST-TEST READING COMPREHENSION PASSAGES

A.   The Full Text of the Pre-test Reading Comprehension Passage

WHAT IS STRESS?

The term stress has been defined in several different ways. Sometimes the term is applied to stimuli or events in our environment that make physical and emotional demands on us, and sometimes it is applied to our emotional and physical reactions to such stimuli. In this discussion, we will refer to the environmental stimuli or events as stressors and to emotional and physical and physical reactions as stress.

Many sorts of events can be stressors, including disasters such as hurricanes or tornadoes, major life events such as divorce or the loss of a job, and daily hassles such as having to wait in line at the supermarket when you need to be somewhere else in ten minutes. What all these events have in common is that they interfere with or threaten our accustomed way of life. When we encounter such stressors, we must pull together our mental and physical resources in order to deal with the challenge.

How well we succeed in doing so will determine how serious a toll the stress will take on our mental and physical well-being.

REACTING TO STRESSORS

The Canadian physiologist Hans Seyle has been the most influential researcher and writer on stress. Seyle has proposed that both humans and other animals react to any stressor in three stages, collectively known as the general adaptation syndrome. The first stage, when the person or animal first becomes aware of the stressor, is the alarm reaction. In this stage, the organism becomes highly alert and aroused, energized by aburst of epinephrine. After the alarm reaction comes the stage of resistance, as the organism tries to adapt to the stressful stimulus or to escape from it. If these efforts are successful, the state of the organism returns to normal. If the organism cannot adapt to continuing stress, however, it enters a stage of exhaustion or collapse.

Seyle developed his model of the general adaptation syndrome as a result of research with rats and other animals. In rats, certain stressors, such as painful tail-pulling, consistently lead to the same sorts of stress reactions. In humans, however, it is harder to predict what will be stress- full to a particular person at a particular time. Whether a particular

stimulus will be stressful depends on the person's subjective appraisal of that stimulus. How threatening is it? How well have I handled this sort of thing in the past? How well will I be able to handle it this time? For one person, being called upon to give a talk in front of a class is a highly stressful stimulus that will immediately produce such elements of an alarm reaction as a pounding heart and a dry mouth. For another person, being called on to give a talk is not threatening at all, but facing a dead-line to complete a term paper is extremely stressful. In humans, more-over, the specific stress reaction is likely to vary widely; some stressful situations give rise predominantly to emotions of fear, some to anger, some to helplessness and depression

B. *The full Text of the Post-test Reading Comprehension Passage*

SIMILARITY

If people could construct heir mates, the mates would look, act, and think very much like themselves. By being attracted to people like ourselves, we are in effect validating ourselves, saving to ourselves that we are worthy of being liked, that we are attractive. Although there are exceptions, we generally like people who are similar to ourselves in nationality, race, ability, physical characteristics, intelligence, attitudes, and so on. We are often attracted to mirror images of ourselves.

THE MATCHING HYPOTHESIS

If you were to ask a group of friends, "To whom are you attracted?" they would probably name very attractive people; in fact, they would probably name the most attractive people they know. But if we would find that they go out with and establish relationships with people who are quite similar to themselves in terms of physical attractiveness. Useful in this connection is the matching hypothesis, which states that although we may be attracted to the most physically attractive people, we date and mate with people who are similar to ourselves in physical attractiveness. Intuitively, this too seems satisfying. In some cases, however, we notice discrepancies; we notice an old person dating an attractive younger partner or an unattractive person with a handsome partner, in these cases, we will probably find that the less attractive partner possesses some quality that compensates for the lack of physical attractiveness. Prestige, money, intelligence, power, and various personality characteristics are obvious examples of qualities that may compensate for being less physically attractive.

ATTITUDE SIMILARITY

Similarity is especially important when it comes to attitudes. We are particularly attracted to people who have attitudes similar to our own, who like what we like, and who dislike what we dislike. The more significant the attitude, the more important the similarity. For example, it would not make much difference if the attitudes of two people toward food or furniture differed (though even these can at times be significant), but it would be of great significance if their attitudes toward children or religion or politics were very disparate. Marriages between people with great and salient dissimilarities are more likely to end in divorce than are marriages between people who are very much alike.

Generally, we maintain balance with ourselves by liking people who are similar to us and who like what we like. It is psychologically uncomfortable to like people who do not like what we like or to dislike people who like what we like. Our attraction for similarity enables us to achieve psychological balance or comfort. The person who lies what we like in effect tells us that we are right to like what we like. Even after an examination it is helpful to find people who wrote the same answers we did.

Notice the next time you have an examination how you prefer the company of others who have given the same answers as you!

APPENDIX II

(a)The Evaluation Scale of English Summaries

points

1. Identification of the main idea(s) of the reading text targeted. (0)⟶ (30)

2. Identification of the subtopic(s) of the reading text-targeted. (0) ⟶ (30)

(b)

مقیاس ارزشیابی خلاصه نوشت فارسی

امتیاز داده شده

(30) ⟵ 1 - شناسائی موضوع/ موضوعات اصلی متن خوانده شده

(0)

(30) ⟵ 2 - شناسائی موضوع/ موضوعات فرعی متن خوانده شده

(0)

REFERENCES

[1]   Broukal, M. (2004). Weaving it together 4 (Connecting reading and writing). Boston: Thomson (Heinle).
[2]   Brown, H. D. (2001). Teaching by principles (An interactive approach to language pedagogy (2nd edition). New York: Longman (Pearson Education).
[3]   Brown, H. D. (2004). Language assessment (Principles and classroom practices). New York; Pearson Longman.
[4]   Carekkm P. L. (1983). Three components of background knowledge in reading comprehension. *Language Learning*, 33, 183-207.
[5]   Davies, P. & Pearse, E. (2000). Success in English teaching. Oxford: Oxford University Press.
[6]   Farahzad,F. & Emam A. (2010). Reading–writing connections in EAP courses: implications and application. *Journal of Language Teaching and Research*. 1(5), 596-604.
[7]   Grabe, W. (2003). Reading and writing relations: second language perspectives on research and practice, In Kroll, B.(Ed.). *Exploring the dynamics of second language writing* (pp. 242-262). Cambridge: Cambridge University Press.
[8]   Hock, S.T., & Poh, C.L. (1997).The performance of a group of Malay-Medium students in an English reading comprehension test. *RELC Journal*, 10, 81-89.
[9]   Johns, A., & Mayes, P. (1990). An analysis of summary protocols of university ESL students. *Applied Linguistics*, (11), 253-271.
[10]   Kirkland, M. R. & Saunders, M. A. (1991). Maximizing student performance in summary writing: Managing cognitive load: *TESOL Quarterly*, 25, (1),
[11]   Koda, K. (2005). Insights into second language reading (a cross-linguistic approach). Cambridge: Cambridge University Press.
[12]   Lee, J.F. (1986). On the use of the recall task to measure L2 reading comprehension. *Studies in Second Language Acquisition*, 12, 25-41.
[13]   Lee, J.F. (1987). Background knowledge and L2 reading. *The Modern Language Journal*, 70, 350-354.
[14]   Nunan, D. (Ed.) (2003). Practical English language teaching. New York: McGraw Hill.
[15]   Nuttall, C. (1996). Teaching reading skills in a foreign language. Oxford: Macmillan Heinemann.
[16]   Rodriguez, K. E. (2006). Contemporary issues and decisions (Reading, writing, and thinking in today's world). Longman (Pearson Education).
[17]   Seal, B. (1997). Academic encounters (Reading, study skills, and writing). Cambridge: Cambridge University Press.
[18]   Shohamy, E. (1985). A practical handbook in language testing for the second language teacher. Tel Aviv, Israel: School of Education. Tel Aviv University.
[19]   Wolf, D.F. (1993). A comparison of assessment tasks used to measure foreign language reading comprehension. *The Modern Language Journal*, 77(4), 473-488.

Abbas Emam holds a PhD in TEFL from Allameh Tabatabai University, Tehran, Iran. He has been teaching at the English Department of Shahid Cahmaran University of Ahvaz, Iran since 2001. His main field of interest is translation studies, and has been extensively writing and translating on different aspects of translation issues in Iranian magazines and journals, e.g. Motarjem (The Translator), The Iranian Journal of Translation Studies.