The Impact of Cloze Task, Translation, and Back Translation on the Technical Vocabulary Learning and Retention

Zahra Fotovatnia

Islamic Azad University, Najafabad Branch, Iran
fotovatnia@yahoo.com

Fatemeh Amiri

Islamic Azad University, Najafabad Branch, Iran
Fatemeh_872@yahoo.com

Doi:10.5901/mjss.2012.v3n1p137

Abstract

This study focused on the comparison between the effects of translation, back translation, and cloze task used for the learning and retention of technical medical terms in an ESP class.These tasks were selected to comply with Depth of Processing Hypothesis based on which the more cognitively one is engaged in learning a word, the more likely it is to remember it later. Investigating learners' motivation for learning technical vocabulary by these techniques was another concern of this study. In this regard 42 medical female undergraduates attending an ESP class at Islamic Azad University of Najafabad took part in the experiment. They were randomly divided into three groups. The first group filled in the blanks of two English medical texts with technical vocabulary (Cloze task) in two sessions. The second group translated the same medical texts into Persian and the third group translated the same texts from Persian into English (Back translation). All groups took the immediate and delayed posttests. The results showed that the cloze task group(CTG) outperformed the others. In fact, cloze task was found more effective than the other two tasks for technical vocabulary learning, as it required a deeper level of processing.

Keywords: *Cloze task, Translation, Back translation, Technical vocabulary, ESP*

I. Background

English for special purposes (ESP) found its way after the World War II and ESP is the major activity around the world today. It is an enterprise involving education, training and practice that draws upon three major realms of knowledge: language, pedagogy and the students' areas of interest (Robinson, 1991). In such specialized courses the instructors traditionally taught that using glossary to obtain the knowledge of the technical terms will provide the non native reader with what he /she needs particularly in scientific texts. But today using English within academic, professional or workplace environment increases the importance of learning and internalizing technical

vocabulary. Hence, choosing appropriate techniques has a special place in designing ESP courses.

According to Nation (2001), the third level of vocabulary (i.e., technical words) makes up 5% of the running words in specialized texts, and these words consists of words which are of frequent occurrence in a specialized text or subject area, but do not occur at very low frequency in other fields.Chung and Nation (2003, p. 114) mention that "there is considerable research evidence about the nature and coverage of high frequency and academic words, but there has been little investigation of technical vocabulary". Therefore, the most important part of ESP should be considered for teaching technical vocabulary in each field; however, in Iranian ESP classes students memorize technical vocabulary for taking exams. They have many problems in the comprehension of scientific texts.

Catford (1965, p.20) says that translation is "the re placement of textual material in one language (SL) by equivalent textual material in another language (TL)." Tudor (1987) used translation activities with one group of ESP learners in Germany. Two activities were described in which native input material and a variety of translation tasks were used. He found that these activities motivated communicative activities and fostered the acquisition of new language resources.

Back translation, according to Shingenobu (2007), is defined as the original language obtained by translating input into target language and then translating the resulting text back into the original language. On the other hand, if translation is a process from L2 in to L1, back translation will be the translational versions from L1into L2 and by considering the correction of back translation the correction of target language translation can be examined. Douglas and Craig (2007) calls back translation as the technique which most commonly used to check the accuracy of translation in survey research.This procedure is commonly used to test the accuracy of translation in multi country research (Brislin 1970, 1980).Larson (1985) considered the difference between translation and back translation. His theory suggests that a bilingual and skillful person can increase his translation quality by using back translation. In 1997, Pan did a study about using back translation for learning English in China. The results of her study showed that back translation can be used as a way of learning English.

In cloze procedure, described in Taylor (1953), one or several words are removed from a sentence and students should fill in the missing words. That sentence is named the 'stem', and the removed term itself is the 'key' (Higgins, 2006). Steinman (2002) mentions her use of cloze procedure as a teaching instrument which caused students to practice using context clues as a reading strategy. Lombard (1990) says the use of cloze tests in her English second language classes for junior and senior students could solve reading problems of learners and increase their confidence when they received

immediate and satisfactory feedback. Neville (1984), Bachman (1982) and Alderson (2000) contend that cloze procedure can be used to assess reading comprehension. Cloze tasks can be employed for instruction as well as assessment. Valmot (1983) calls instructional cloze as powerful means to meet student needs through selective deletion. Selective deletion is the difference between cloze as instructional and assessment tools. The way of organizing cloze activities are suggested by Rye (1982).He says that these activities should meet the needs of the most students and include content words which are more challenging than other parts of speech. He adds that there should be "inverse relationship between the difficulty of prediction and frequency of occurrence" (p.62).

To investigate the effects of translation, back translation, and cloze task on both learning and delayed recall of technical vocabulary, and choosing the most effective technique in technical vocabulary acquisition, the following questions were formulated.

1. Do cloze task, translation, and back translation affect learners' technical vocabulary development differently?
2. Do the three techniques of vocabulary teaching differ significantly in terms of the retention of technical vocabulary in long-term memory?

To answer the research questions, the following hypotheses were considered.

1. Cloze task, translation, and back translation have the same effect on the learners' technical vocabulary development.
2. Cloze task, translation, and back translation have the same effect on the retention of the technical vocabulary in long-term memory.

2. Methodology

A. Participants and Instruments

The participants were 50 Iranian learners whose first language was Persian. The sample participants were all female students of ESP in the medical department of Islamic Azad University, Najaf Abad branch. Their age range was between18-21. Forty two students comprised the final number of participants in the study. The reason for reducing the number of participants was that after administrating the quick placement test(QPT), which was conducted to ensure the homogeneity of the students , eight students were excluded from the study because they had either extremely high or extremely low score on the test. Those students whose scores were between one standard above and below the mean were chosen for the final data analysis.

The instruments in this study were QPT (version 2), test of target words unfamiliarity, the texts which were chosen *Dynamic Cone-beam Reconstruction* (Montes, 2006) and ("Human Skeleton", nd), immediate and delayed Posttests.

In order to remove the possibility of students' familiarity with to-be-taught words, a test of technical and general vocabulary was used prior to the experiment. This was the test with 30 words. The students were required to write Farsi or English translation of the words, they know. The words with which students were unfamiliar were used in the study and together formed the content of the pretest. Readability of the texts were also calculated, Flesch readability ease for the texts was between30-50, was considered appropriate for college students. Then, the texts were undergone three different adjustments in the line with the objectives of the study. For the cloze task group (CTG), the technical words of the English medical texts were deleted and the word provided at the top of the texts.The English medical texts were translated into Persian for the translation group (TG), and the Persian translation of the medical texts was prepared for back translation group (BTG).

Two vocabulary tests were prepared for two sessions. They tested the participants' acquisition of the words taught in those sessions. The first test included 14 multiple-choice questions and testing 6 technical and 8 general words. The second comprised 16 items which were multiple-choice including 8 technical and 8 general words. Both tests were based on the words instructed those particular days. The reason for presenting the subjects with 6-8 words each session lies in Finocchiaro and Bonomos assertion (1973) that in general, no more than 8 new words should be presented at one time; otherwise, it is not manageable by the students.

A 30-item recognition vocabulary test was used as the delayed posttest. It was constructed to measure the learners' lexical acquisition and recall. The posttest was administered two weeks after the treatment to test retention of words in long-term memory. After preparing the items and before the experiment, the tests were piloted with some students similar to the participants of the study in terms of English background to remove any potential flaws and to find out whether the instructions were comprehensible, whether the allotted time was enough, and whether the distracters were effective. In the piloting stage 50students took the immediate and delayed Posttests. These students had passed that course, so that they studied most of the to-be instructed words before. The reliability of the tests was estimated through KR-21 formula and it was reported 0.73 and 0.74 for the immediate posttests and 0.79 for delayed posttest. To assure the content validity of the tests, two scholars in the field reviewed the tests thoroughly. Then, some modifications were made on the items based on their suggestions in order to alleviate the existing problems. As the result of the revision process, the tests were eventually prepared for the main project.

B. Procedure

A sample of 50 Medical undergraduates was given QPT in order to be homogenized

in terms of their general proficiency level. The allotted time for this administration was 45 minutes. The final participants of the study were 42 who took part in the experiment. Then, the test of target words unfamiliarity was conducted to select the words with which none of the students were familiar with. Then, the treatment which covered sessions two different days were given to the participants. Each session lasted 60 minutes. For two sessions, two medical texts (The human skeleton & the heart and circulation) were selected.

The participants were randomly assigned to three groups. In each group, fourteen female homogenous students were considered as the participants.

Each group received a different treatment. They were arranged according to the purpose of the study in the following ways:

1- The first experimental group, CTG, received English medical texts with blanks to fill them with technical vocabulary.
2- The second experimental group, TG, received English medical texts in order to translate into Persian.
3- The third group, BTG, received Persian translation of medical texts in order to translate into English.

In CTG the English medical text in which technical words were deleted, was distributed. They had to fill in the blanks with technical words provided at the top of the text. In TG, English medical text was distributed among the students. They had to translate the text into Persia. In BTG, the Persian translation of the same medical text was distributed among the students. The researcher explained that they were required to translate the text into English.

Any kind of dictionary could be used by each group and the students were allowed to interact with each other. The researcher helped them to remove their problems. After completing the tasks, the students took the immediate posttest which tested the words learnt at that session. The time the students had to answer the questions of each test was 15-18 minutes. The students' overall achievement was assessed by the delayed given posttest after the treatment. It was administrated after two weeks and comprised all the 30 words which were presented in both texts. The estimated time for administration of delayed posttest was 30 minutes.

As mentioned before to ensure the validity and reliability of the tests, they were piloted with another group before administrating them to the experimental groups.

C. Statistical Analysis and Results

One-way ANOVA was performed on QPT, on the immediate and delayed posttests scores. Kruskal Wallis was used on the immediate and delayed posttests scores when one-way ANOVA could not be used. Paired- samples t-test were performed on the three vocabulary test (Posttests) scores achieved by each of the three groups.
In this study, research question one asked whether any of presented techniques, cloze task, translation, and back translation would affect learners' technical vocabulary development differently. This question can be answered by examining the results of immediate posttests which were administrated to assess learning of technical vocabulary. Therefore, Kruscal Wallis Test was used to see whether there were any overall differences among the experimental groups on the immediate posttests for technical vocabulary. First of all, the homogeneity of variances was calculated by Levene statistics which is shown in Table1.

Table 1

Test of Homogeneity of Variances

Technical words, Immediate posttests

Levene Statistic	df1	df2	Sig.
5.343	2	39	.009

Table 1 shows that variances are not equal, p =.009 and one-way ANOVA could not be used, so Kruskal Wallis was run. Table 2 presents the descriptive statistics of immediate posttests for technical vocabulary.

Table 2 Descriptive Statistics of Immediate Posttests for Technical Vocabulary

Descriptives

Technical words, Immediate posttests

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
cloze task	14	13.4286	.75593	.20203	12.9921	13.8650	12.00	14.00
translation	14	10.2857	1.81568	.48526	9.2374	11.3341	7.00	13.00
Back translati	14	10.4286	2.40878	.64377	9.0378	11.8194	7.00	14.00
Total	42	11.3810	2.28412	.35245	10.6692	12.0927	7.00	14.00

The Table shows that the highest mean scores of the immediate posttests for technical vocabulary belongs to CTG (M = 13.4286, SD = .75593) and TG has the lowest mean score (M =10.2857, SD =1.81568).

Table 3 Statistical Kruskal Wallis Test

Test Statistics [a,b]

	Technical words, Immediate posttests
Chi-Square	18.903
df	2
Asymp. Sig.	.000

a. Kruskal Wallis Test

b. Grouping Variable: group

As it is shown, p= .000,which means there was a significant difference among the groups. In order to see which group performed differently on the test, post hoc Tamhane test was run, which revealed significant differences among the groups. Table3 shows the results.

Table 4 Results of Tamhane Test on Mean Differences of the Immediate Posttests for Technical Vocabulary

Multiple Comparisons

Dependent Variable: Technical words, Immediate posttests

Tamhane

(I) group	(J) group	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
cloze task	translation	3.14286*	.52564	.000	1.7548	4.5309
	Back translation	3.00000*	.67473	.001	1.1959	4.8041
translation	cloze task	-3.14286*	.52564	.000	-4.5309	-1.7548
	Back translation	-.14286	.80618	.997	-2.2105	1.9248
Back translation	cloze task	-3.00000*	.67473	.001	-4.8041	-1.1959
	translation	.14286	.80618	.997	-1.9248	2.2105

*. The mean difference is significant at the .05 level.

The result shows that CTG is significantly different from TG and BTG (p = .000 and p = .001 accordingly). That is, the cloze task is better than both translation and back

translation. (MD =3.14286 and MD= 3.00000 accordingly).There is no significant difference between TG and BTG.

The second research question asked whether any of presented techniques of vocabulary teaching differ significantly in terms of the retention oftechnical vocabulary in long-term memory.This question can be answered by examining the results of delayed posttest administered to assessretention of the meaning of technical vocabulary. In order to investigate whether there were any overall differences among the experimental groups on the delayed posttest for technical vocabulary descriptive statistics were calculated. Table 5 shows the results.

Table 5 Descriptive Statistics of Delayed Posttest for Technical Vocabulary

Descriptives

technical words, delayed posttest

	N	Mean	Std. Deviation	Std. Error	5% Confidence Interval fo Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
cloze task	14	12.4286	1.34246	.35879	11.6535	13.2037	11.00	14.00
translation	14	9.5714	1.45255	.38821	8.7328	10.4101	7.00	12.00
Back translatio	14	9.0000	2.66025	.71098	7.4640	10.5360	5.00	13.00
Total	42	10.3333	2.40595	.37125	9.5836	11.0831	5.00	14.00

Table 5 shows that CTG obtained highest mean score on the delayed posttest (M =12.4286, SD = 1.34246) and BTG the lowest (M =9.0000, SD =2.66025).This means that the close task was the most effective technique which was used to recall the technical vocabulary. Then, a one-way ANOVA was run in order to see whether the differences among the mean scores were statistically significant. The following table (Table 6) shows the results.

Table 6 Results of One-way ANOVA on the Delayed Posttest for Technical Vocabulary among Groups

ANOVA

technical words, delayed posttest

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	94.476	2	47.238	12.896	.000
Within Groups	142.857	39	3.663		
Total	237.333	41			

The results show that there is a significance difference among groups, $F(2.39) = 12.896$, $p = .000$. In order to know which group performed differently on the test, post hoc Scheffe test was run, which revealed marked differences among the groups. Table 7 shows the results.

Table 7 Results of Scheffe Test on Mean Differences of the Delayed Posttests for Technical Vocabulary

Multiple Comparisons

Dependent Variable: technical words, delayed posttest
Scheffe

(I) group	(J) group	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
cloze task	translation	2.85714*	.72339	.001	1.0162	4.6980
	Back translation	3.42857*	.72339	.000	1.5877	5.2695
translation	cloze task	-2.85714*	.72339	.001	-4.6980	-1.0162
	Back translation	.57143	.72339	.734	-1.2695	2.4123
Back translation	cloze task	-3.42857*	.72339	.000	-5.2695	-1.5877
	translation	-.57143	.72339	.734	-2.4123	1.2695

*. The mean difference is significant at the .05 level.

The result of Scheffe test shows that CTG is significantly different from TG and BTG ($p = .001$ and $p = .000$ accordingly). That is, the cloze task is better than both translation and back translation. (MD =3.42857 and MD= 2.85714 accordingly).However, no significant difference was observed between TG and BTG.

Table 8

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	technical words, delayed posttest	10.3333	42	2.40595	.37125
	Technical words, Immediate posttests	11.3810	42	2.28412	.35245

Table 8 shows that the mean of three groups decreases after two weeks.

Table 9

Paired Samples Test

		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference Lower	Upper			
Pair 1	technical words, delayed posttest - Technical words, Immediate postte	1.04762	1.14663	.17693	1.40493	-.69031	-5.921	41	.000

Table 9 shows that there is a significant difference between the two means on immediate and delayed posttests. This result suggests that three groups performed better on the immediate posttests.

3. Discussion and Conclusion

The result of Kruskal Wallis test analysis showed that there was a significance difference among groups on the immediate posttests and according to the result of post hoc Tamhane test, CTG outperformed of the other groups. The result of one-way ANOVA analysis from the delayed posttest also suggested that there was a significance difference among three groups, and by considering Scheffe Test, CTG performed better on the delayed posttest. To explain why CTG outperformed the other two groups in learning and retention of technical vocabulary, one might refer to peculiar features of this task.

The cloze task stimulates metacgnitive awareness which means in this activity reflecting upon one's own thinking process occurs (Burley, Brown, & Saunders et al, 1985; Dewitz, Carr & Patberg, 1987). Increasing metacognition in competing cloze activity can be observed when the students argue with each other when they deliberate on their choices for deleted words. They put themselves in the framework of authors' mind and discover the writing strategies that were used in the text, this awareness creates important linkage between reading and written expression (Anderson & Rubano, 1991). In completing cloze activity, the readers are more aware of the meaning and use reading skills like searching and scanning consciously, because the reader need to replace deleted words (Steinman, 2002). Gunning (1998) also claims that in cloze procedure, the words are deleted, so the readers are forced to pay close attention to meaning; and they have to comprehend what they read. Consequently, cloze task is an effective task and most researchers argue that an effective task should induce a deep level of processing of the new words, a high degree of elaboration, or richness of encoding. This claim is related to Craik and Lockhart's (1972) levels of

processing depth theory, according to which the chance that some piece of new information will be stored into long-term memory is determined by the depth with which it is initially processed. It is impossible to decide which of the tasks requires deeper processing. Only aftercomparing amount of the learners' vocabulary learning in the two conditions, it can be concluded that the task which resulted in better vocabulary learning, required a deeper level of processing. Therefore cloze task requires a deeper level of processing.

References

Anderson, P. M., & Rubano, G. (1991). *Enhancing aesthetic reading and response*. Urbana, Illinois: National Council of Teachers of English.*Applicationfor college readers.* Paper presented at the annual meeting of the BooksInc.

Brislin, & Richard W. (1970), "Back Translation for Cross-Cultural Research," *Journal of Cross-CulturalPsychology, 1* (3), 185–216.

Burley, J., Brown, B. G., & Saunders, B. L. (1985). *Metacognition: Theory and Application for college readers.* Paper presented at the annual meeting ofthe International Reading Association.(ED266428)

Catford, J. C. (1965): A Linguistic Theory of Translation. An Essay in Applied Linguistics.London: OUP.

Chung, T.M., Nation, I.S.P.(2003). Technical vocabulary in specialised texts. *Reading ina Foreign Language,15* (2), 103–116.

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

Dewitz, P., Carr, E. M. & Patberg, J. P. (1987). Effects of inference training on comprehension and comprehensionmonitoring. *Reading ResearchQuarterly, 22*(1), 99-121.

Douglas,P.,&Craig,C.S.(2007). Collaborative and Iterative Translation: An Alternative Approach to BackTranslation. *Journal of International Marketing,2*(1), 30-43

Finochiaro, M.,& Bonomo, M.(1973). The foreign language learner: Guide for teachers. New York: Regents Publishing Company. Inc.

Gunning,T.G. (1998). *Assessing and Correcting: Reading and Writing Difficulties.* Needham Heights:Allyn & Bacon.

Humanskeleton. (n.d).Retrieved March 21, 2012 from http://www.personal.une.edu.au

Larson, M.L.(1986). *A Guide to Cross-Language Equivalence*, *Meaning-Based Translation,* University Press of America Inc.

Lombard, J. V. (1990). How to Improve Language Proficiency by Means of ClozeTesting. *Journal for Language Teaching, 24*(3) 21-29.

Montes,P.(2006). *Dynamic Cone-beam Reconstruction for perfusion Computed Tomography*(Doctoral dissertation).Retrieved from http://www.ub.uni-hedelberg.de/archiv/7020

Nation, I. S. P. (2001). *Learning vocabulary in another language*. Cambridge: Cambridge University Press.

Neville, M. (1984). National Assessment of English. *Teaching English, 17*(3),26-30.

Newmark, P. (1988b). *A Text Book of Translation.* London: Prentice Hall International (UK) Ltd.

Pan,W.G.(1997).*The Chinese-English Contrastive Outline*. Beijin: Beijing Language and CultureUniversity Press.

Robinson, P. (1991). *ESP Today: a Practitioner's Guide.* Hemel Hempstead:Prentice Hall

Rye, J. (1982). *Cloze procedure and the teaching of reading.* Portsmouth, NH: Educational

Shigenobu,T.(2007). Evaluation and usability of back translation for intercultural communication.*Springer-Verlag Berlin Heidelberg*.4560,259-265

Taylor, W.L.(1953). Cloze procedure:a new tool for measuring readability. *Journalism Quarterly 30,415-433.*

Tudor,L.(1987).Using translation in ESP.*ELT Journal Volume 41/4*. Oxford: Oxford University Press.

Valmont, W. J. (1983). Cloze deletion patterns: How deletions are made makes abig difference.