A study of the effects of teacher-performed gestures as a means of semantic elaboration on L2 word learning and retention

Ghasemi, Ehsan ✉
Islamic Azad University, Hajiabad Branch, Iran (ehsanghasemi2004@yahoo.com)

Feyzi Behnagh, Reza
University at Albany, SUNY, USA (rfeyzibehnagh@albany.edu)

Abstract

The current study investigates the effects of semantic elaboration operationalized through teacher-performed gestures on L2 vocabulary learning and retention. Forty-two Persian-speaking Iranian university students learning English as a foreign language were randomly selected and assigned to two groups. They were presented with two sets of new English words (determined through a pretest) and were instructed to do their best to learn the words in the two conditions. One set consisted of word-translation pairs (control condition) and the other set consisted of word-translation pairs followed by a video showing the meaning of each word through gestures. Each participant was presented with every experimental word but in two different orders (counter-balancing). Then they were required to complete four recall tasks (i.e. free recall in English, free recall in Persian, cued recall in English, and delayed cued recall in English). The results of data analysis using repeated measures of ANOVA showed a negative effect of using gestures on free recall in English, a positive effect of using gestures on free recall in Persian, and no statistically significant effect of using gestures on cued recall and delayed cued recall in English. Based on the findings some pedagogical implications are provided.

Keywords: teacher-performed gestures; semantic elaboration; free recall; cued recall; delayed cued recall

A study of the effects of teacher-performed gestures as a means of semantic elaboration on L2 word learning and retention

1. Introduction

Research into vocabulary acquisition in second and foreign languages has a long history. Some are studies dealing with active/passive vocabulary repertoire (e.g., Crow & Quigly, 1985; Laufer & Paribakht, 1988) , the order of learning the vocabulary of a language (Dulay & Burt, 1974; Goldschneider & DeKeyser, 2001), corpus analyses for designing high-frequency word lists (e.g., Meara, 1995; Schmitt, 2000; West, 1953), efficient ways of teaching vocabulary (e.g., Folse, 2006; Fraser,1999), and vocabulary learning strategies including direct and incidental methods of vocabulary learning (Batia & Jan, 2001; Brown & Perry, 1991). Moreover, in foreign language teaching methods and programs, teaching vocabulary holds an important part, especially in Grammar-Translation and Direct methods and more recently in lexical approach. Although Audiolingualism and Communicative methodology subordinated the role of vocabulary to structures and notions and functions of language, in the 1980s research into lexis and discourse analysis combined with arguments from psycholinguistics and L1 literacy research to reassert the importance of vocabulary in language learning (Johnson & Johnson, 1999); and one of the concerns of language teachers is still how to get learners to learn SL/FL words.

One related field that has relatively been explored less is the processing of input at the level of words when learning of a foreign language happens. To learn a new word, attention should be directed to both its meaning and form. In at least three stages the new word is then processed. In the first stage, learners pay attention to the structure of the word (form) and encode it in their mind (for example the length of the word, number of letters and syllables, and its visual and acoustic characteristics). Then, semantic structures related to the meaning of the new word is activated or built (in case a given item is totally new and there is no previous concept existing) in the learners' minds. For L2 words these structures are already there and only need to be activated but for L1 words, they may need to be built (Ausubel, 1964; Jiang, 2002; Levelt, 1989). And finally, in the stage of mapping, the encoded item is matched with proper meaning.

It is not possible to teach all new words in the classroom context due to time limits in language programs. A logical substitute for it is identifying and teaching effective vocabulary learning strategies (Oxford, 1990). One of the vocabulary learning strategies which started to receive attention by researchers in the late 1980s and early 1990s is semantic elaboration (sometimes referred to as deep semantic processing). Semantic elaboration is a term which implies focusing more on semantic aspects of the new word and allocation of available resources mainly to its meaning during learning. Focusing attention on structural features in processing words is called structural elaboration.

1.1 Statement of the Problem

The related literature shows positive effects of semantic elaboration on memory for L1 known words and memory for other types of stimuli including sentence recall and text recall (Barcroft, 2004) but for new words (especially in FL learning) that effect is not necessarily positive. Word processing for known words and new words are not the same since for known words semantic and formal structures are already in the mind and just need to be activated but for new words they may need to be built at the time of learning (Giacobbe, 1992; Ringbom, 1983). The resources that are used to focus on semantic features of the new word in FL learning can be allotted to the structural features and matching the form and the meaning of the new word. If this is the case, semantic elaboration during teaching and learning of formal features of new words in a foreign language may have no effect or even a negative effect on the learning and recall of new words.

In EFL classes when instructors want to introduce a new word they usually utilize semantic elaboration techniques such as using pictures, synonyms, asking questions about the word, referring to learners' experience related to the word or using gestures. The existing body of research regarding semantic elaboration in foreign language and second language is relatively less than L1; especially studies where semantic elaboration is not used as a mnemonic device or is not combined with a mnemonic device (Atkinson & Raugh, 1975; Bower & Reitman, 1972).

Moreover, the results of few studies dealing with effects of semantic elaboration on learning of new words in FL/SL and pseudowords are different and sometimes contradictory (e.g. Coomber, Ramstad, & Sheets, 1986; Barcroft, 2002). The reason might be the way researchers attempted to operationalize semantic elaboration (Barcroft, 2002). In the literature, researchers have utilized various techniques to make participants perform semantic elaboration including definition and example matching (Coomber et al., 1986), making a sentence using the target word (Barcroft, 2000), giving a rating to the target word on a scale of pleasantness (Barcroft, 2002), answering a question about the word (Barcroft, 2003), and using varied visual representation of referents for target vocabulary (Sommers & Barcroft, 2013). In the present study, the teacher-performed gesture is used as a form of semantic elaboration. A very common technique among language instructors to draw learners' attention to the meaning of a new word is using gestures. Compared to other investigated techniques, this one seems more authentic and natural in foreign language classes. Besides, the present study is a case of Persian speakers learning English. Since Persian and English utilize different orthography, results of the study may provide new insights in using semantic elaboration in EFL learning and teaching for learners whose languages use different orthographies from English.

1.2 Purpose of the Study

This study aims to address the following questions:

➢ Does the use of teacher-performed gestures (as a form of semantic elaboration) in teaching affect the free recall of new words in English? If yes, what is the nature of that effect?

➢ Does the use of gestures as a form of semantic elaboration in teaching affect free recall of new words in native tongue (here Persian)? If yes, what is the nature of that effect?

➢ Does the use of gestures as a form of semantic elaboration in teaching affect cued recall of new words in English? If yes, what is the nature of that effect?

Dependent variables are free recall in English, free recall in Persian, cued recall in English, and delayed cued recall in English. Independent variables or conditions of the experiments include using gestures along with new words (semantic elaboration) as experimental condition and not using gestures along with new words as control condition.

2. Review of Related Literature

In regard to the effects of semantic and structural elaboration on lexical learning there are at least three important theories in the related literature, namely Levels Of Processing model (LOP), Transfer Appropriate Processing model (TAP), and Type Of Processing-Resource Allocation model (TOPRA) which are discussed here.

2.1 Levels of Processing or Levels Model

One important theory related to the semantic elaboration is "Levels of Processing" (LOP) theory which was proposed by Craik and Lockhart (1972). In their seminal work they questioned the adequacy of previous multi-store models of human memory which were dominated by the concept of stores and information transfers among the stores. They argued that multistore formulation is unsatisfactory in terms of its capacity, coding, and

forgetting characteristics. They relied on the conception of "Depth of Processing" which claims that perception involves the analysis and processing of stimuli at a number of stages. The beginning stages of perception involves the analysis of sensory features (e.g. loudness, pitch, angles lines, etc.) while later stages include pattern recognition and the extraction of meaning (Craik & Lockhart, 1972). Generally, deep processing indicates greater amount of semantic analysis. This hierarchy of processing stages of analysis for perception was adopted in other theories as well (e.g. Sutherland, 1968; Treisman, 1964). Craik and Lockhart (1972) were specifically concerned with memory trace as a byproduct of perceptual analysis. They argue that the persistence of a trace is interrelated to the depth of processing in a way that deeper levels of analysis make more elaborate, longer lasting, and stronger traces. Based on the levels of processing (LOP) framework, semantic processing is beneficial for retention (Craik & Lockhart, 1972; Craik & Tulving, 1975). Some researchers have demonstrated that mature participants instructed to perform a semantic processing task on a list of words demonstrated better retention than participants instructed to perform a non-semantic processing task (e.g., McDaniel & Masson, 1977).

There has been a number of studies supporting the levels of processing (e.g. Atkinson & Raugh, 1975; Bower & Reitman, 1972; Craik & Tulving, 1975; Ellis & Beaton, 1995; Epstein, Philips, & Johnson, 1975; Hyde & Jenkins, 1969; Till & Jenkins, 1973; and Walsh & Jenkins, 1973 among others). Hyde and Jenkins (1969) found that the amount and organization of recall of word lists varied with the type of incidental task performed during presentation of the list. All participants heard a randomized list of high–strength primary word associates. When the incidental task required using word as a semantic unit (rating the word as to its pleasantness), recall and recognition were equivalent to those of control group with no incidental task. When the incidental task involved checking for certain letters or estimating the number of letters in the word, recall and recognition were greatly reduced. Walsh and Jenkins (1973) showed that amount of free recall varied with type and combination of orienting task performed during the presentation of a list of low-frequency English nouns. When the orienting task was semantic (i.e. required the participants to process the meaning of words) recall was significantly higher than that of participants performing nonsemantic orienting tasks. When two orienting tasks were performed serially and one of the tasks was semantic in nature, recall was significantly higher than the recall of groups performing only nonsemantic tasks. When two tasks were nonsemantic, recall was indistinguishable from the performance of participants performing single nonsemantic tasks. Besides, Till and Jenkins' (1973) study revealed that amount and organization of recall of word lists depended on the orienting tasks performed by participants even when the task varied from word to word within a single list.

Levels of semantic processing underwent some adjustments in the following years (Fisher & Craik 1977; Jacoby & Craik1979). In the adjusted versions, for example, it was suggested that there is no difference in the durability of traces made by semantic versus non-semantic (e.g. orthographic and phonological) processing but traces that are made by deeper levels of processing are more distinguishable than those made by shallower levels of processing (see Jacoby & Craik, 1979). However the idea of "levels of processing" never changed in the LOP framework.

A criticism for LOP framework is that it does not explicitly mention why semantic processing is deep, meaningful, and permanent and structural processing is shallow non meaningful and less permanent compared to semantic processing. In other words just semantic related processing seems to be meaningful in LOP. As Moeser (1983) claims "The notion of qualitative differences among levels has never been precisely explained or illustrated by proponents of the levels model" (p. 316). There is a group of studies which have suggested the consideration of some other memory factors such as trace uniqueness, trace congruity, etc. (e.g. Craik & Tulving 1975; Moscovitch & Craik, 1976). Some other researchers have mentioned the need of differentiating levels of processing in the semantic level of analysis in the mind (e. g. Schulman, 1974).

2.2 Transfer Appropriate Processing Model

Morris, Bransford, and Franks (1977) manipulated LOP as a function of acquisition task and the type of recognition test in three experiments. Based on their results they argued that there is a need for reconsideration of

certain assumptions underlying the levels of processing approach. In particular, they questioned arguments that non-semantic or shallow levels of processing are necessarily inferior to deeper levels of processing. They claim that the reason for this so-called inferiority may be due to the inappropriateness of relationship between acquisition task and test type rather than inherent inferiority of the memory traces resulted from shallow levels of processing in conducted studies. They suggested a new framework named "Transfer Appropriate Processing" (or TAP) as a replacement for LOP. This new framework emphasizes that a) the value of particular acquisition activities must be defined relative to particular goals and purposes, and b) assumptions about the quality and durability of the resulting memory traces can be determined relative to the appropriateness of testing situation. In other words, TAP claims that one should not assume that the traces of certain items are less durable or efficient than others because those items were processed at shallower levels. Transfer Appropriate Processing does not assumes any inherent difference in the nature of memory traces resulting from semantic versus non-semantic levels of processing. There are some related studies supporting TAP model (e.g. McDaniel & Kearney, 1984; Moeser, 1983).

McDaniel and Kearney (1984) in their first experiment gave college students three verbal learning tasks to perform with some participants instructed to use a particular verbal or imagined encoding strategy for all three tasks and other participants not instructed to use any particular strategy. They found that the effectiveness of the assigned encoding strategies varied as a function of the learning task. In their second experiment participants were instructed to perform a different encoding strategy for each learning task. Participants who were assigned strategies in a task-appropriate fashion generally recalled more than participants who were assigned the same strategies with the learning tasks in a haphazard fashion. They concluded that optimal learning for a range of tasks can require deployment of several semantic/elaborative strategies in a task-appropriate fashion. Moeser (1983) in two experiments reported that orthographic orienting task (structural elaboration) did not produce poorer retention than semantic orienting task when the orthographic task was presented in such a way to ensure that the list items would be encoded as units and when the test was designed to eliminate the effect of encoding elaboration to positive-response orienting questions. He concluded that the depth-of-processing effect was composed of two components, namely task-demand component (that affects the probability of encoding target items as identifiable units) and trace elaboration to positive-response questions. He argued that the two components can be examined independently of each other to determine the degree to which each contributes to a particular experiment effect.

A similar and related but less discussed framework with emphasis on the type of the test relative to the type of task regarding the effectiveness of a given encoding activity is Task-appropriate processing suggested by McDaniel and Kearney (1984). They claim that their approach, compared to TAP model, is broader and better in capturing the variability in the effectiveness of different types of processing (encoding strategies). They assume that any aspect of a given learning task has an impact (or impacts) on the effectiveness of applied encoding processes aimed at enhancing retention. These aspects include both the task and materials to be learned (p. 371).

2.3 Type of Processing-Resource Allocation Model (TOPRA)

In the same line of research on semantic and structural processing, Barcroft (2000) proposed TOPRA model which is consistent with TAP model. This model focuses on competition for limited cognitive resources when learners try to learn different things at the same time. It predicts that when demands of processing are sufficiently high, semantic elaboration can increase learning rates for the semantic properties of words while decreasing learning rates for the structural properties of words at the same time. This model also claims that structural elaboration can simultaneously increase learning rates for structural properties of new words while decreasing learning rates for the semantic properties of words. The model confirms that with restricted processing resources that learners have, it is difficult to process input for both meaning and form. He conducted some other studies to provide more support for his model (e.g. Barcroft, 2003, 2004, 2006, and 2007). Barcroft (2003) found evidence for a potentially inhibitory effect of semantic elaboration during second language word learning. His study concerned the effects of questions about word meaning during an immediate lexical learning task. English

speaking learners of Spanish were instructed to learn 24 new Spanish words in two conditions while viewing word-translation pairs. For twelve of the words they thought about specific questions related to the meaning of each target word. For the next twelve words, they were told only to do their best to learn the target words. The results of recall tasks showed significantly greater cued-recall scores in the no questions condition than in the questions condition. Barcroft (2004) compared writing new words in sentences with word-translation repetition learning alone. He asked learners of Spanish as a second language to learn 24 Spanish words in one of two conditions while viewing word-translation paired. Results of immediate and delayed posttests on productive vocabulary knowledge showed strong negative effects for sentence writing condition implying that sentence writing can inhibit word from learning during initial stages of L2 lexical acquisition.

In Barcroft (2006) English speaking learners of Spanish as a second language were asked to look at a sequence of word-translation pairs on computer screens. The participants were asked to write 12 target words as they saw them and write nothing for the other 12 target words and just try to learn them. Productive vocabulary learning on immediate and delayed measures was higher in the no-writing condition suggesting that forced output (writing) without access to conceptual aspects can detract from word learning by exhausting processing resources needed to encode new words.

2.4 Performing Gestures and Memory

There are various groups of studies dealing with effects of using gestures on memory which have tried to investigate different aspects of this relationship including memory enhancement, accessibility of words and phrases, facilitated recall, reaction time, word frequency in language use, learning of words and action phrases, and retention. A seminal laboratory study dealing with effects of performing gestures on verbal memory dates back to 1980s. Engelkamp and Krumnacker (1980) in Germany found that if verbal phrases for actions are encoded by self-performed representational actions, their retrieval is better than if the verbal information is only heard or read. Independently, Cohen (1981) conducted a similar study and his findings were highly consistent with the work of Engelkamp and Krumnacker (1980). This effect is called Enactment Effect, Self-Performed Task (SPT) effect, or Action Memory in the related literature. In the same line of investigation, Zimmer, Helstrup, and Engelkamp (2000) investigated the effect of enactment on accessibility in memory. They concluded that enacted verbal materials have a better accessibility in memory, and free recall for enacted items happens with less effort. Spranger, Schatz, and Knopf (2008) in a study of accessibility of enacted items in younger and older adults found high accessibly for both age groups in immediate and delayed free recall tasks. Masumoto, et al. (2006), using magneto-encephalography to find reasons for memory enhancement effect of enactment, observed that enacted phrases had a higher speed of recognition compared to phrases encoded only audiovisually. In a more recent study, Macedonia and Knosch (2011) conducted a research on the effects of using gesture on abstract word learning. They also investigated if learning new words with self-performed gestures facilitates sentence production. They concluded that self-performed gestures produce better memory for abstract words than the condition of encoding words only audiovisually. They also observed more frequency for words presented with gestures in sentence production by participants.

Macedonia and Kriegstein (2012) claim that "the first study on the impact of gestures on memory for verbal information in a foreign language was conducted by Quinn-Allen (1995)". In a study of learning French expressions (as a foreign language) Quinn-Allen (1995) found that learning emblematic gestures with simultaneously presented expressions does lead to greater recall and more retention over time. In another study of the effects of gestures on foreign language learning, Tellier (2008) examined French children that learned EFL (English as a Foreign Language). She asserts that using gestures in presentation of L2 lexical items and specially their reproductions by learners significantly influence the memorization of them as far as active knowledge of vocabulary is concerned. There has been other studies on other aspects of effects of gestures on memory for foreign language words and expressions too (for example Khalili, Rahmany, & Zarei, 2014 among others who studied resolving lexical ambiguity by gestures), all of which confirmed the positive effects of utilizing gestures for foreign language teaching and learning (see Macedonia & Kriegstein, 2012 for a review).

In sum, findings from studies mentioned above and other similar ones show enhancing effects of using gestures on verbal memory. However, in almost all of above-mentioned studies and other similar studies with EFL learners, participants were required to learn and perform gestures in the experimental groups. In regular foreign or second language classes this approach does not seem plausible; it needs special teacher training, syllabus design, and supervision and it may not be applicable to all language learners (for example adults may be reluctant to use gestures to communicate in classes because they may feel insecure in terms of their egos) and all contexts (Scovel, 1978). In regular second language classes, teacher may use gestures to communicate or focus on the message and students are usually not asked to learn and replicate the gestures although they may learn them indirectly and use them unconsciously. The present paper tries to study the condition of regular adult EFL classes in which only teachers use gestures along with audiovisual presentation of items not students.

3. Methodology

3.1 Participants

From accounting students at Islamic Azad University, 42 students were randomly selected (*M* age=19.3 years, age range: 17-28 years). The random selection was to partly ensure that the participants represent the larger population of Persian speaking EFL university students. They agreed to take part in the experiments as volunteers by filling in the consent forms. They had passed two English courses (Prerequisite English and General courses) at the university before the experiments and none of them had English training beyond the English courses that they had passed at the junior high school and the high school. They were randomly divided into two equal groups (21 members each) with the same number of males and females.

3.2 Design

The study was quasi-experimental in nature. We compared two vocabulary learning conditions: a. learning in the presence of teacher-performed gestures (+gesture), and b. learning without teacher-performed gestures (-gesture). In +gesture condition participants were asked to learn words by paying attention to the word-translation pairs and rate the gestures showing the meaning of each word on a 1-5 scale based on how well the gesture represents the meaning of each word. In -gesture condition (control condition) participants were asked to do their best to learn each word after watching the word-translation pairs.

3.3 Experimental Words

Twenty concrete words that could easily be presented and understood through pictures were selected (Appendix A) and divided into two equal groups based on the number of syllables in each word. In order to study words of different length, one-syllable, two-syllable, and three-syllable words were included. To keep a balance between two groups in terms of memory load and processing requirements, in each group there were two one-syllable words, five two-syllable words, and three three-syllable words. Each participant was exposed to both conditions but there were two presentation orders (counterbalancing) for the two groups of participants in order to control the possible bias that could result from the order in which the conditions were provided. Participants of the first section tried to learn words 1-10 in +gesture condition and words 11-20 in –gesture condition. Subsequently, the second group was instructed to learn words 1-10 in –gesture condition and words 11-20 in +gesture condition (Table 1).

Table 1

The conditions and word arrangements for each section

	First section	Second section
Words 1-10	+Gesture	-Gesture
Words 11-20	-Gesture	+Gesture

3.4 Instruments

The following instruments were used to conduct the study: a language background questionnaire was used to ensure that all the participants were learning English as a foreign language and they did not use it outside English classes and to ensure that none of them had the experience of learning another foreign or second language; a consent form for participation in the study was attached to the questionnaire; a test of participants' familiarity with experimental words was used in which they were required to translate 20 English words into Farsi; it was labeled "The pretest" (Appendix A); a gesture evaluation form was utilized on which there was a 1-5 scale for the gesture performed for each word in order to engage participants in semantic elaboration (Appendix B); two sheets of paper that were labeled Post-tests 1-1 and 1-2 were prepared for immediate free recall task in English and immediate free recall task in Persian; one sheet of paper (sheet number 3) was prepared for the purpose of cued recall task in English which was labeled Posttest 1-3 (on this sheet participants were asked to write the experimental words after the picture of each was presented on the screen this time without the word); one sheet of paper (number 4) was provided for delayed cued recall task in English which was labeled posttest 2; a computer program (a power point file) was designed to present the instructions, experimental words and pictures, and gesture videos related to each target word; and finally a computer and a video projector to present the presentation and the testing phases of the study in a timely manner.

3.5 Procedure

The following nine steps were taken in the process of data gathering:

First, participants were asked to fill in the consent form and the language background questionnaire. Then they were asked to turn in the papers. Second, in the test of participants' familiarity with experimental words, participants were asked to translate the experimental words in to Farsi. In order to give all the words equal processing time, words were read one by one for the participants by the experimenter with fifteen-second intervals. Third, participants were instructed about studying phase of the experiment. Instructions were both presented orally and appeared on the screen. They read the following instructions (translated into Farsi): "Two sets of word-translation pairs will be presented to you on the screen. Each set consists of ten pairs. For each set, pay close attention to what you should do. Each set will be presented twice and between each word-translation pair you have eight seconds to do what you are instructed to". There was no time limit at this stage and participants were free to ask any questions if there was any ambiguity. They were required not to take notes during the study phase. Fourth, immediately after the third stage, the words were presented. In +gesture condition the following instruction was presented and read: "Ten word-translation pairs will be presented to you followed by a video showing the meaning of them through gestures. For each, please determine how well the video expresses the meaning of each word on the five-point scale you have. Score one show the lowest level of meaningfulness and score five shows the highest level of meaningfulness of each gesture for you. At the same time do your best to learn each word".

In –semantic condition (control condition), the following instruction appeared on the screen and was read to the participants: "Ten word-translation pairs will be presented to you. Please do your best to learn the words". As mentioned earlier, each set of words were presented twice. In +gesture condition the eight-second time between words started after the gesture presentation was finished. Between two presentations of each set there was a one-minute lapse. After completion of this stage participants turned in the gesture-evaluation form. The fifth stage was the free-recall task in English. Immediately after the fourth stage, two blank sheets of paper (no.1 and no.2) were distributed and participants were instructed as follows: "Please try to remember and write as many words as you can in English based on what you just saw on screen. Please feel free to write as much of the words as you can and if you are not sure about the spelling do not worry, just write what you can remember. Use paper number one". The participants were given six minutes to recall experimental words in English. In the sixth stage, immediately after the free recall task in English, the participants were instructed to remember and write as many words as they could in Persian (free recall in Persian) based on what they saw in the presentation phase. They

were given six minutes to complete the task. They were asked to use paper number two. After the completion of both tasks, the participants were asked to turn in their recall papers. The next stage (seventh) was Posttest 1-3 (cued recall task in English) in which files of experimental pictures were played again, this time without corresponding words and gestures. The sheets of posttest 1-3 (number 3) were distributed. Following instructions were presented: "Write the English word for each picture you see. Try to write as much of the word as you can remember and do not worry about the spelling." In the eighth stage, the participants were given a ten-minute break and they were served refreshments (food and drink). A piece of Persian traditional music was played in the background. This was to introduce a delay in to the experiments to explore the effects of time as a potential factor on the recall. And finally, after the break, Posttest 2 was administered. It was the same as Posttest1-3 and was to examine the effects of delay. The same procedure and instructions as stage seven was employed.

4. Results

None of the participants were able to translate the experimental words in the pretest, which meant the words were new for them. The gesture evaluation results showed an acceptable reliability (Cronbach's alpha: 0.78). The Lexical Production Scoring Protocol (LPSP, Barcroft, 2000), was used to score free recalls and cued recalls in English. The scoring in LPSP depends of the percentage of letters correct or present that are produced by the participants (e.g., Barcroft, 2003; Prince, 2012). For example the score of .25 is given if only one letter is correct (present and at the right place) or if at least 25% but less than 50% of the letters are present. This protocol was used to consider partial knowledge of the English words along with the complete knowledge. Each fully remembered word in Persian received 1 point. Then totals were calculated for both free recall tasks in Persian and English and immediate and delayed cued recall tasks in English in both conditions for each participants. In other words, based on the situation in which the remembered words had been presented, the scores were totaled. Therefore, each participant had two scores for each recall task (+gesture and –semantic).

The SPSS software was utilized for the purpose of data analysis. First, a two way repeated-measures ANOVA was used in which the presence or absence of gestures in the presentation phase and the type of recall task were regarded as independent variables and the scores on recall tasks constituted the dependent variables. Two participants were omitted from the study because they did not complete the experiments. Table 2 represents the mean and standard deviation of eight groups of scores.

Table 2

Descriptive statistics of eight groups of scores

	Mean score	SD	Number
Free Recall in English, +gesture	2.068	1.448	40
Free Recall in Persian, +gesture	6.762	1.790	40
Cued Recall in English, +gesture	2.343	1.850	40
Delayed Cued Recall in English +gesture	1.987	1.494	40
Free Recall in English –gesture	3.368	1.979	40
Free Recall in Persian –gesture	3.762	2.358	40
Cued Recall in English –gesture	2.268	1.894	40
Delayed Cued Recall in English –gesture	2.000	1.718	40

The results of repeated-measures ANOVA indicated that there was a significant main effect of the presence or absence of gestures following vocabulary presentation, $F(1, 39) = 8.39$ $p = .006$. In other words disregarding all other variables, the scores of recall tasks were different for +gesture and –gesture conditions. A significant main effect of the type of recall task was observed too, $F(3, 117) = 99.8$ $p < .001$ which meant the type of recall task affected the scores of recall tasks. Moreover, there was a significant interaction between the type of recall task and the presence or absence of gestures in the presentation, $F(3, 117) = 41.39$ $p < .001$. Tables 3, 4, and 5 represent the estimated marginal means of the above-mentioned main effects.

The results show that there is an outstanding difference in the means between FRE and FRP (Free Recall in

English and Free Recall in Persian) in the +gesture condition (FRP is higher). This difference is much smaller when we compare the same variables (FRE and FRP) in –gesture condition. The mean of CRE (cued recall in English) and the mean of DCRE (delayed cued recall in English) are close in +gesture and -gesture conditions.

Table 3

Estimated marginal means of the presence (1) or absence (2) of gestures

Gesture	Mean	Std. error
1	3.291	0.207
2	2.850	0.250

Table 4

Estimated marginal means of different recall tasks

Recall task	Mean score	Std. error
Free recall in English	2.719	0.221
Free recall in Persian	5.262	0.267
Cued recall in English	2.306	0.279
Delayed cued recall in English	1.994	0.237

Table 5

*Estimated marginal means of Gesture*Recall task*

Condition	Recall task	Mean	Std. error
+Gesture	Free recall in English	2.069	0.229
	Free recall in Persian	6.763	0.283
	Cued recall in English	2.344	0.293
	Delayed cued recall in English	1.988	0.236
-Gesture	Free recall in English	3.369	0.313
	Free recall in Persian	3.762	0.373
	Cued recall in English	2.269	0.300
	Delayed cued recall in English	2.000	0.272

The delay imbedded in the study resulted in smaller recall in DCRE than CRE. The mean of CRE in + gesture condition was higher than FRE but lower than FRP. In contrast, in –gesture condition, CRE was lower than both FRE and FRP. Moreover the means of FRE in +gesture condition was smaller than FRE in –gesture condition. In contrast, FRP in +gesture condition was much higher than FRP in –gesture condition.

Pairwise comparison showed statistically significant differences between mean scores of +gesture and –gesture conditions, free recall in English and free recall in Persian, free recall in English and cued recall in English, free recall in English and delayed cued recall in English, free recall in Persian and cued recall in English, and free recall in Persian and delayed cued recall in English (at the .05 level).

In order to compare effects of the two conditions (-gesture and +gesture) on the scores of different recall tasks, a one way repeated-measures ANOVA was utilized. In other words, the mean scores of +gesture and –gesture conditions of each recall task were compared using one way repeated-measures ANOVA. The results showed a statistically significant difference between the means of +gesture and –gesture conditions of free recall in English, (free recall in –gesture condition was higher), and free recall in Persian, (free recall in +gesture condition was higher). For two other recall tasks no statistically significant difference was observed (see Table 6).

In sum, the results show that performing gestures could have significant effects on free recall in English (research question 1) and free recall in Persian (research question 2). The nature of these effects is not the same. Using gestures negatively affected the scores of free recall in English as a foreign language. In contrast, using gestures positively affected the scores of free recall in Persian as the native tongue. Moreover, using gesture did

not have any statistically significant effect on cued recall and delayed cued recall in English (research question 3).

Table 6

Comparison of +Gesture and –Gesture conditions in three recall tasks

Recall task +gesture condition	Recall task –gesture condition	Mean	Sig.
Free recall in English	Free recall in English	-1.30	0.008
Free recall in Persian	Free recall in Persian	3.00	0.000
Cued recall in English	Cued recall in English	0.10	1.00
Delayed cued recall in English	Delayed cued recall in English	0.03	1.00

5. Discussion and Conclusion

Looking at the results from the 'Levels Of Processing model' perspective (Craik & Lockhart, 1972) and assuming that gestures increase the depth of processing, it seems that giving depth to processing led to an increase in the free recall of words in the native language; however, it negatively affected free recall in the foreign language. Therefore, the results are consistent with the LOP model only for free recall in Persian. Moreover, the LOP model claims a better memory trace for semantically elaborated words. Statistical analysis of the delayed cued recall task showed no difference between semantically elaborated condition and the control group.

The results of the analysis of free recall in English provide support for TOPRA model (Barcroft, 2000) as well, since the mean score of free recall in –gesture condition was significantly higher than +gesture condition. It may be argued that paying attention to gestures used the cognitive resources that could otherwise be allotted to encode word forms. In other words, semantically oriented task of watching and scoring gestures won the competition of using the limited cognitive resources of participants when learning new words. This trend was not supported by the results of statistical analyses of free recall in Persian, cued recall in English and delayed cued recall in English.

Besides, part of the findings (more specifically free recall in English and Persian) show some evidence for TAP model (Morris et al., 1977). It supports TAP in that in the encoding phase (presentation), participants in +gesture condition are required to pay attention and score the gestures on evaluation form; a task that has no relation to testing phase in which they are required to recall the words in English. The participants might have performed better in free recall in English which is basically a structure-focused task if instead of watching gestures and evaluating them, they had been asked to pay attention to the number of consonants and vowels (structural properties), for instance, in the presentation task (Moeser, 1983). In free recall in Persian task, it could be argued that in presentation task, participants did not need to encode Persian words in the first place since they already knew the concepts and words (Barcroft, 2004). They just needed to activate them, a task for which gestures following words played a facilitative role. A later version of TAP model recognizes a distinction between known words (here Persian) and new words (here English) when considering the role of a variable on memory for words (Barcroft, 2002).

5.1 Previous Studies and Present Findings

The positive effect of using gestures on memory for words in Persian (participants' native language) is consistent with several previous studies which found that semantic elaboration positively affects memory for known words (e.g. Bower & Reitman, 1972; Hyde & Jenkins, 1969; Ross, 1981; Schulman, 1974; Tresselt & Mayzner, 1960). Barcroft (2000) found that free recall in a known language (English there) was grater in the semantic elaboration condition (operationalized through making pleasantness rating about word referents) compared to structural elaboration condition. However, McDaniel and Kearney (1984) found that semantic orienting conditions resulted in less vocabulary recall than uninstructed condition for a known language. Barcroft (2007) found that forced output (word writing) inhibited word learning by exhausting limited processing

resources available for word learning which is consistent with the findings of present study regarding free recall in English (foreign language) when participants were required to score the gestures in presentation phase.

The findings are also consistent with Sommers and Barcroft (2013) who found that semantic elaboration (operationalized through token variability in presentation phase) produced negative effects on L2 vocabulary learning in contrast to previous studies which found positive effects of word form variability (emphasizing on structural properties of words) on L2 vocabulary learning. In studies on Iranian EFL learners, Khalili et al. (2014) found a positive effect for using gestures on resolving English lexical ambiguity; however, the gestures were both teacher and student performed which may not provide an appropriate basis for comparison. In another study, Ahmadi (2014) found no significant positive significant effect for elaboration on retention of EFL vocabulary compared to the control condition (structural elaboration).

5.2 Pedagogical Implications and future research

The findings of the current study have implications for teaching foreign languages. They are as follow:

First, in accordance with TOPRA model and the results of the free recall tasks in the present study which support this model, language learners have limited information processing resources. Therefore, teachers should recognize this limitation and consider it when they teach. They should not expect students to process several pieces of new information simultaneously.

Besides, language teachers should recognize a distinction between semantic and structural properties of words when they are focusing on vocabulary, and deal with them separately. In other words, when teachers intend to teach new words, they should draw learners' attention to one aspect of words at a time. For example, they may first emphasize the degree of formality of specific word, the synonyms, antonyms, distinctive semantic properties of the word in the target language compared to native language, matching the word to its picture, and using gestures to clarify its meaning, all of which refer to semantic properties of the words with different degrees. Then, they may draw learners' attention to structural properties of the word such as pronunciation, pitch and stress patterns, number of syllables, and arrangement of vowels and consonants. They may change the order but it does not seem appropriate to use a combination of these two stages.

Measurement is another issue. There should be a match between how teachers present materials and what the test (e.g. quizzes, midterm and final examinations) measures. In other words, if the presentation of materials puts emphasis on semantic properties of new words, the test should not ask about formal characteristics. For example, asking students to write the definition of a new word is a semantically-oriented task which should be used if in teaching the emphasis has been on the meaning of the target word. In contrast, asking learners to write the target word in the exam is basically structurally oriented. Moreover, the same notion of match between encoding and testing tasks may be applied in writing EFL and ESL vocabulary learning textbooks, more specifically in the compilation of exercises related to each unit. Exercises should be tuned to the way the materials are presented.

Additionally, foreign and second language teachers seem to be biased in using semantic elaboration techniques. As Sommers and Barcroft (2013) claim in spite of mounting body of evidence showing that semantically-oriented tasks do not facilitate L2 word form (i.e. structure) learning, many L2 instructors and other language learning professional do not consider this basic pedagogical guideline to be intuitive. Therefore there should be a reconsideration of this standpoint. And finally, we should draw a distinction between the role of teacher-and-student-performed gestures (self-performed) and teacher-only-performed gestures. In the reviewed body of literature there are numerous evidence supporting the positive role of the former in learning both first language and second language however the findings of the present study shows, that at least with regards to formal properties of L2 words (manifested in free recall and cued recall task), the effect is negative.

Future research may consider the effects of cultural differences in using gestures for L2 vocabulary learning

and teaching especially in terms of the frequency and adoption. Some cultures use more hand and face gestures in oral communication and for another culture in a different geographical context that frequency of gestures may not be natural. Kavakli and Nasser (2012), for instance, explored the effects of cultural differences of Latin Americans and Anglo-Celtics in gesture-based interfaces and found that Anglo-Celtics use more hand gestures in longer periods than Latin Americans in describing objects. Therefore, the question would be what will happen if L2 teachers are advised not to use gestures when they are teaching new L2 (in a gesture-neutral setting) vocabulary across different cultures, as the results of the present study suggest a negative effect for using gestures in presentation phase on free and cued recalls in L2.

6. References

Ahmadi, M. (2014). Semantic and structural elaboration in L2 vocabulary learning and retention. *Procedia-Social and Behavioral Sciences, 98*, 109-115. http://dx.doi.org/10.1016/j.sbspro.2014.03.395

Atkinson, R. C., & Raugh, M. R. (1975). An application of the mnemonic keyword method to the acquisition of a Russian vocabulary. *Journal of Experimental Psychology: Human Learning and Memory, 104,* 126-133. http://dx.doi.org/10.1037/0278-7393.1.2.126

Ausubel, D. P. (1964). Adults versus children in second-language learning: Psychological considerations. *Modern Language Journal, 48*, 420–424. http://dx.doi.org/10.1111/j.1540-4781.1964.tb04523.x

Barcroft, J. (2000). *The effects of sentence writing as semantic elaboration on allocation of processing resources and second language acquisition.* Unpublished doctoral dissertation, University of Illinois at Urbana-Champaign.

Barcroft, J. (2002). Semantic and structural elaboration in L2 lexical acquisition. *Language Learning, 52,* 323-363. http://dx.doi.org/10.1111/0023-8333.00186

Barcroft, J. (2003). Effects of questions about word meaning during L2 Spanish lexical learning. *The Modern Language Journal, 87*(3)*, 546-561. http://dx.doi.org/10.1111/1540-4781.00207

Barcroft, J. (2004). Effects of sentence writing in L2 lexical acquisition. *Second Language Research, 20*(4), 303-334. http://dx.doi.org/10.1191/0267658304sr233oa

Barcroft, J. (2006). Can writing a new word detract from learning it? More negative effects of forced output during vocabulary learning. *Second Language Research, 22*(4), 487-497. http://dx.doi.org/10.1191/0267658306sr276oa

Barcroft, J. (2007). Effects of opportunities for word retrieval during second language vocabulary learning. *Language Learning, 57*(1), 35-56. http://dx.doi.org/10.1111/j.1467-9922.2007.00398.x

Batia, L., & Jan, H. (2001). Incidental vocabulary acquisition in a second language: The construct of task-induced involvement. *Applied Linguistics, 22*, 1-26. http://dx.doi.org/10.1093/applin/22.1.1

Bower, G. H., & Reitman, J. S. (1972). Mnemonic elaboration in multi-list learning. *Journal of Verbal Learning and Verbal Behavior, 11,* 478-485. http://dx.doi.org/10.1016/S0022-5371(72)80030-6

Brown, T. S., & Perry, L. P. (1991). A comparison of three learning strategies for ESL vocabulary acquisition. *TESOL Quarterly, 25*(4), 655-670. http://dx.doi.org/10.2307/3587081

Cohen, R. L. (1981). On the generality of some memory laws. *Scandinavian Journal of Psychology, 22*, 267-281. http://dx.doi.org/10.1111/j.1467-9450.1981.tb00402.x

Coomber, J. E., Ramstad, D. A., & Sheets, D. R. (1986). Elaboration in vocabulary learning: A comparison of three rehearsal methods. *Research in the teaching of English, 20,* 281-293

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. http://dx.doi.org/10.1016/S0022-5371(72)80001-X

Craik, F. I. M., & Tulving, E. (1975). Depth of processing and the retention of words in episodic memory research, *Journal of Experimental Psychology: General, 104,* 268-294. http://dx.doi.org/10.1037/0096-3445.104.3.268

Crow, J. T., & Quigly, J. R. (1985). A semantic field approach to passive vocabulary acquisition for reading comprehension. *TESOL Quarterly, 19*(3)*, 479-513. http://dx.doi.org/10.2307/3586275

Dulay, H. C., & Burt, M. K. (1974). Natural sequences in child second language acquisition. *Language Learning,*

24(1), 37–53. http://dx.doi.org/10.1111/j.1467-1770.1974.tb00234.x

Engelkamp, J., & Krumnacker, H. (1980). Imaginale und motorische Prozesse beim Behalten verbalem Materials.Zeitschrift für experimentelle und angewendte. *Psychologie, 27*, 511–533.

Ellis, N., & Beaton, A. (1995). Psychologistic determinants of foreign language vocabulary learning. In B. Hardy (Ed.), *Lexical issues in language learning* (pp. 107-165). Ann Arbor, MI: Benjamins.

Epstein, M. L., Philips, W. D., & Johnson, S. J. (1975). Recall of related and unrelated word pairs as function of processing level. *Journal of Experimental Psychology: Human Learning and Memory, 104,* 149-152. http://dx.doi.org/10.1037/0278-7393.1.2.149

Fisher, R. P., & Craik, F. I. M. (1977). Interaction between encoding and retrieval operations in cued recall. *Journal of Experimental Psychology: Human Learning and Memory, 3*, 701-711. http://dx.doi.org/10.1037/0278-7393.3.6.701

Folse, K. S. (2006). The effect of type of written exercise on L2 vocabulary retention. *TESOL Quarterly, 40*, 273-293. http://dx.doi.org/10.2307/40264523

Fraser, C. A. (1999). Lexical processing strategy use and vocabulary learning through reading. *Studies in Second Language Acquisition, 21*, 225-241. http://dx.doi.org/10.1017/s0272263199002041

Giacobbe, J. (1992). A cognitive view of the role of L1 in the L2 acquisition process. *Second Language Research, 8*, 232–250. http://dx.doi.org/10.1177/026765839200800304

Goldschneider, J., & DeKeyser, R. M. (2001). Explaining the "natural order of L2 morpheme acquisition" in English: A meta-analysis of multiple determinants. *Language Learning, 51*(1), 1–50. http://dx.doi.org/10.1111/1467-9922.00147

Hyde, T. S., & Jenkins, J. J. (1969). The differential effects of incidental tasks on the organization of recall of a list of highly associated words. *Journal of Experimental Psychology, 82,* 472-481. http://dx.doi.org/10.1037/h0028372

Jacoby, L. L., & Craik, F. I. M. (1979). Effects of elaboration of processing at encoding and retrieval: Trace distinctiveness and recovery of initial context. In L. S. Cermak & F. I. M. Craik (Eds.), *Levels of Processing in Human Memory* (pp. 1-21). Hillsdale, NJ: Lawrence Erlbaum Associates.

Jiang, N. (2002). Form–meaning mapping in vocabulary acquisition in a second language. *Studies in Second Language Acquisition, 24*, 617-637. http://dx.doi.org/10.1017/s0272263102004047

Johnson, K., & Johnson, H. (1999). *Encyclopedic dictionary of applied linguistics.* Oxford: Blackwell Publishers. http://dx.doi.org/10.1111/b.9780631214823.1999.x

Kavakli, M., & Nasser, K. 2012. Impacts of culture on gesture-based interfaces: A case study of Anglo-Celtics and Latin Americans. *International Journal on advances in life science, 4*(3)*,* 100-110.

Khalili, M., Rahmany, R., & Zarei, A. A. (2014). The effects of using gestures on resolving lexical ambiguity in L2. *Journal of language Teaching and Research, 5*(5), 1139-1146. http://dx.doi.org/10.4304/jltr.5.5.1139-1146

Laufer, B., & Paribakht, T. S. (1998). The relationship between passive and active vocabularies: Effects of language learning context. *Language Learning, 48*(3), 365-391. http://dx.doi.org/10.1111/0023-8333.00046

Levelt, W. J. M. (1989). *Speaking: From intention to articulation*. Cambridge, MA: MIT Press.

Macedonia, M., & Knosch, T. R. (2011). Body in mind: How gestures empower foreign language learning. *Mind, Brain, and Education, 5*(4), 196-211. http://dx.doi.org/10.1111/j.1751-228X.2011.01129.x

Macedonia, M., & Kriegstein, K. V. (2012). Gestures enhance foreign language learning. *Biolinguistics, 6*(3-4)*,* 303-416.

Masumoto, K., Yamaguchi M., Sutani, K., Tsuneto, S., Fujita, A., & Tonoike, M. (2006). Reactivation of physical motor information in the memory of action events. *Brain Research*, *1101*(1), 102–109. http://dx.doi.org/10.1016/j.brainres.2006.05.033

McDaniel, M. A., & Kearney, E. M. (1984). Optimal learning strategies and their spontaneous use: The importance of task-appropriate processing. *Memory and Cognition, 12*(4), 361-373. http://dx.doi.org/10.3758/BF03198296

McDaniel, M. A., & Masson, M. E. (1977). Long-term retention: When incidental semantic processing fails.

Journal of experimental Psychology: Human Learning and Memory, 3, 270-281. http://dx.doi.org/10.1037/0278-7393.3.3.270

Meara, P. (1995). The importance of an early emphasis on L2 vocabulary. *The Language Teacher, 19*(2), 8–11.

Moeser, S. D. (1983). Levels of processing: qualitative differences of task-demand differences*? Memory and Cognition*, *11*, 316-323. http://dx.doi.org/10.3758/BF03196978

Morris, C. D., Bransford, J. D., & Franks, J. J. (1977). Levels of processing versus transfer appropriate processing. *Journal of Verbal Learning and Verbal Behavior, 16,* 519-533. http://dx.doi.org/10.1016/S0022-5371(77)80016-9

Moscovitch, M., & Craik, F. I. M. (1976). Depth of processing, retrieval cues, and uniqueness of encoding as factors in recall. *Journal of Verbal Learning and Verbal Behavior, 15*, 447-458. http://dx.doi.org/10.1016/S0022-5371(76)90040-2

Oxford, R. (1990). *Language Learning Strategies: What every teacher should know.* Rowley, MA: Newbury House.

Prince, P. (2012). Towards an instructional program for L2 vocabulary: Can a story help? *Language and Technology, 16*(3)*,* 103-120.

Quinn-Allen, L. (1995). The effects of emblematic gestures on the development and access of mental representation of French expressions. *The Modern Language Journal, 79*(4), 521-529. http://dx.doi.org/10.1111/j.1540-4781.1995.tb05454.x

Ringbom, H. (1983). Borrowing and lexical transfer. *Applied Linguistics, 4*, 207–212. http://dx.doi.org/10.1093/applin/4.3.207

Ross, B. H. (1981). The more, the better? Number of decisions as a determinant of memorability. *Memory and Cognition, 9*, 23-33. http://dx.doi.org/10.3758/BF03196948

Schmitt, N. (2000). *Vocabulary in language teaching*. Cambridge: Cambridge University Press.

Schulman, A. (1974). Memory for words recently classified. *Memory and Cognition, 2*, 47-52. http://dx.doi.org/10.3758/BF03197491

Scovel, T. (1978). The effect of affect on foreign language learning: A review of anxiety research. *Language learning, 28*(1), 129-142. http://dx.doi.org/10.1111/j.1467-1770.1978.tb00309.x

Sommers, M., & Barcroft, J. (2013). Effects of referent token variability on L2 vocabulary learning. *Language Learning, 63*(2)*,* 186-210. http://dx.doi.org/10.1111/lang.12007

Spranger, T., Schatz, T.R., & Knopf, M. (2008). Does action make you faster? A retrieval-based approach to investigating the origins of the enactment effect. *Scandinavian Journal of Psychology, 49*(6)*,* 487-495. http://dx.doi.org/10.1111/j.1467-9450.2008.00675.x

Sutherland, N. S. (1968). Outlines of a theory of visual pattern recognition in animals and man. *Proceedings of the Royal Society. Series B, 171,* 297-317. http://dx.doi.org/10.1098/rspb.1968.0072

Tellier, M. (2008). The effect of gestures on second language memorisation by young children. *Gesture*, *8*(2), 219-235. http://dx.doi.org/10.1075/gest.8.2.06tel

Till, R. E., & Jenkins, J.J. (1973). The effects of cued orienting tasks on the free recall of words. *Journal of Verbal Learning and Verbal Behavior, 12*, 489-498. http://dx.doi.org/10.1016/S0022-5371(73)80029-5

Treisman, A. (1964). Monitoring and storage or irrelevant messages in selective attention. *Journal of Verbal Learning and Verbal Behavior, 3*, 449-459. http://dx.doi.org/10.1016/S0022-5371(64)80015-3

Tresselt, M. E., & Mayzner, M. S. (1960). A study of incidental learning. *Journal of Psychology*, *50,* 339–347. http://dx.doi.org/10.1080/00223980.1960.9916451

Walsh, D. A., & Jenkins, J. J. (1973) Effects of orienting tasks on free recall in incidental learning: "Difficulty," "effort," and "process" explanations. *Journal of Verbal Behavior, 12*, 481-488. http://dx.doi.org/10.1016/S0022-5371(73)80028-3

West, M. (1953). *A general service list of English words*. London: Longman

Zimmer, H. D., Helstrup, T., & Engelkamp, J. (2000). Pop-out into memory: a retrieval mechanism that is enhanced with the recall of subject-performed tasks. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26*(3), 658-670. http://dx.doi.org/10.1037/0278-7393.26.3.658