The Journal of Teaching Language Skills (JTLS)
4 (4), Winter 2013, Ser. 69/4
ISSN: 2008-8191. pp. 25-51

Elicitation, Recast, and Meta-Linguistic Feedback in Form-Focused Exchanges: Effects of Feedback Modality on Multimedia Grammar Instruction

N. Rashidi *
Associate Professor, TEFL
Shiraz University
email: nrashidi@rose.shirazu.ac.ir

H. Babaie
Ph.D. Student, TEFL
Shiraz University
email: hamed.babaie@gmail.com

Abstract

This research explores the effects of three computer-mediated feedback modalities, that is, elicitation, recast, and meta-linguistics, on the learning of English participial, gerund, and infinitival phrases among Iranian intermediate-level EFL learners. The overriding focus of the present study was to investigate whether different types of feedback given through form-focused computer-human exchanges would produce varying immediate and long-term effects on the participants' incorporation of linguistic forms. To this end, 160 participants were randomly assigned to three equivalent experimental groups. One group then received treatment on the three types of phrases through a tutorial system and multimedia grammar exercises where they received elicitation on the errors they made. The second and the third groups likewise received treatment on the same types of phrases through the same tutorial system, but received either recast or meta-linguistic feedback on their errors. The groups then sat for an immediate and a delayed post-test of grammar two weeks after the experiment. The experiment revealed that meta-linguistic feedback yielded the strongest immediate and sustained effects as compared with those of elicitation and recast. Likewise, while recast produced stronger immediate effects on learning as compared to those of elicitation, its sustained effects were much smaller than those of elicitation and meta-linguistic feedback.

Keywords: Computer-mediated feedback, elicitation, meta-linguistic feedback, multimedia grammar instruction, recast

---

Received: 06/12/2012 Accepted: 02/03/2013

* Corresponding author

1. Introduction

Corrective feedback, it has been argued, plays a beneficial role in facilitating the acquisition of certain L2 forms that are difficult to learn through input alone, including forms that are rare, low in perceptual salience, semantically redundant, do not typically lead to communication breakdown (Long & Robinson, 1998), or that lack a clear form-meaning relationship (DeKeyser, 2005). A growing body of research has also begun to shed light on the emerging relationship between types of corrective feedback and second language learning in face-to-face interaction (Ammar & Spada, 2006; Carroll, 2001; Ellis, Loewen & Erlam, 2006; Loewen & Nabei, 2007; Lyster, 2004; McDonough, 2005). What characterizes corrective feedback is that it offers great potential to draw learners' attention to mismatches between their production and the target-like realization of these hard-to-learn forms. Capitalizing on Schmidt's (1990) Noticing Hypothesis, it has been contended that by drawing learners' attention to mismatches between input and output, or between output and the target-like norm, corrective feedback can facilitate the occurrence of noticing, which is the first step in language building. For learning to occur, second language learners must, then, attend to and notice details and differences between the target language and their interlanguage, and its representation in their production of output (Schmidt, 2001). Corrective feedback, by juxtaposing learning output with input, can assist the acquisition of hard-to-learn forms by increasing the likelihood that they will be noticed. Notwithstanding, a fair amount of controversy exists as to whether different types of corrective feedback affect noticing in the same manner. Indeed, beyond facilitating the noticing of hard-to-learn forms, it has also been suggested that certain types of corrective feedback may also promote L2 processing. Panova and Lyster (2002), for instance, argue that corrective feedback which contains positive evidence about the target language (recasts, for example) may be useful in the internalization of new forms, whereas corrective feedback which does not contain a full reformulation but requires that learners attempt self-repair or output modification may require deeper processing and hence, enhance control of already internalized L2 forms. Feedback modality may, then, affect the quality of noticing and consequently the internalization of target language forms. Yet, the effect of modality on learning outcomes has received scant attention in the literature (Loewen & Erlam, 2006; Sachs & Suh, 2007), which requires further exploration. The main objective of the present study is, then, to explore whether modes of feedback can affect the learning outcome to varying extents.

2. Background

Undoubtedly, grammar accuracy correlates with academic success in the sense that a great deal of classroom work is conducted through oral and written exchanges that require students to strike a balance between the fluency and accuracy of their production. Indeed, failure in harnessing the required accuracy skills in the second language might lead to the emergence of global errors that impede upon communication (Ellis, 2009). This argument thus corroborates the view that, at times, meaning is at the service of form and that in order to avoid communication breakdown, equal importance should be accorded to the role of accuracy while the overriding focus is on communication. Moreover, it is widely recognized that grammar may serve as a "weak interface" between our implicit knowledge of the language or the knowledge built through mere exposure to the salient linguistic features and discourse paradigms of the language, and the knowledge of language which is explicit to the language user, that is, the knowledge that requires both exposure to target language forms and explicit instruction on those forms (Noonan, 2004). The notion of weak interface, then, suggests that certain attention to form is required for promoting accuracy in the second language. In favor of this argument, Ellis (2005) contends that L2 learning is somewhat different from L1 acquisition in that adult learners perceive L2 forms through processes and mechanisms already tuned for the L1. Accordingly, there is great likelihood that in automatic processing of L2 forms, L1 implicit representations conspire, resulting in output fraught with erroneous forms. He further argues that features in the L2 input accessible thanks to recency, frequency, or context are less likely to become intake as their processing is shaped by the L1 mechanisms. This can partly account for why L2 learners need certain amount of explicit instruction in order to achieve target language accuracy in their production. One way of attaining this is through opportunities for learners to engage in form-focused exchanges in the L2 where they receive modified input when their attention can be brought to problematic L2 forms that may lead to communication breakdown.

A growing body of research in SLA has reached a unanimous consensus on the contribution of input enhancement to second language pedagogy (Babaie, 2008; Beatty, 2003; Chapelle, 2001; Gass, 1997; Gass & Torres, 2005; Long, 1996; Pica, 1994; Sharwood Smith, 1993). As long (1996) argues, interactions offer ample opportunities for collaborative efforts such as repetition, confirmation checks, comprehension checks, or clarification requests. These efforts are believed to prove fruitful to language learners to build communication with the instructor. As they struggle to maximize comprehension and negotiate their way through trouble spots, the

teacher-student partnerships are incidentally fine-tuning the second language input so as to render it comprehensible to the learners. What makes such interactions so vital is that at the heart of these exchanges lies the provision of feedback in a variety of forms, including elicitation, recast, explicit correction, explanation, and so on that may help learners notice the *gap* between target language forms and their L2 production. Such feedback may indicate that an error has been committed and hence a provision of the correct target language form or a meta-linguistic explanation about the nature of the error is necessary (Ellis, Loewen, & Erlam, 2006). Learners use feedback to verify their hypotheses about the structural properties of the language (Collentine, 2000; Zamel, 1981). The verification of these hypotheses is an interactive process which may require the provision of metalinguistic feedback or at least recast. The latter can be defined as "a reformulation of all or part of a learner's immediately preceding utterance in which one or more non-target-like (lexical, grammatical, etc.) items is/are replaced by the corresponding target language form(s), and where, throughout the exchange, the focus of the interlocutors is on *meaning,* not language as object" (Long, 2007, p.77).

Convictions are strong that with the corollaries of Computer-Assisted Language Learning (CALL) approaches to L2 pedagogy finding their way into classrooms, corrective feedback delivered via Synchronous Computer-Mediated Communication (SCMC) may hold great promise for the teaching of L2 forms largely due to great visual saliency of certain forms during written interaction, the amount of processing and planning time afforded by synchronous chat, and the enduring turns (Sauro, 2009). In providing argument for CALL feedback, researchers (DiGiovanni & Nagaswami, 2001; Greenfield, 2003; Warschauer, 1996) criticize teacher's feedback on the ground that passive and less proficient learners do not receive enough opportunities for pushed output following feedback; feedback from teachers is inconsistent and unpredictable. Teachers may fail to realize the full range of all available feedback types, and teachers may be reluctant to encourage learner output during feedback due to time constraints and students' embarrassment (Van den Branden, 1997). However, within a CALL setting, feedback delivered by the computer have been noted to alleviate the aforementioned concerns on the ground that CALL offers opportunities for promoting SLA behavior that enhances learning (Beatty, 2003; Chapelle, 2005); CALL feedback can be more consistent than classroom feedback responses that are sometimes only applicable to an individual (Tsutsui, 2004); a computerized environment has the advantage of making available to the learner a variety of different types of feedback (Brandl, 1995); CALL feedback can save face by lowering the affective filter and avoiding the

psychological anxiety in face-to-face feedback (Torlakovic & Deugo, 2004), and behavior tracking through taking snapshots or creating logs allow the researchers to examine the learning processes in depth (Beaudoin, 2004; Chapelle, 2003; Cowan, Choi, & Kim, 2003; Glendinning & Howard, 2003).

This growing body of evidence, then, suggests that computer-based tutorial courseware may offer great potential for augmenting ESL/EFL courses by providing students with ample opportunities for self-practice and for receiving adequate feedback geared to individuals. Yet different types of feedback may produce varying effects, and little research on Computer-Mediated Feedback (CMF) has explored the varying effects of feedback modality on input enhancement and learner's uptake. Moreover, studies like that of Petersen (2010) have primarily focused on comparing the effects of different modalities of only one type of feedback, for instance, oral vis-à-vis written recasts, on L2 form acquisition. It may be likely that different feedback types contribute to varying qualities of input modification and hence varying degrees of learning success. Due to a paucity of research, however, further studies are required to substantiate such claims.

3. Present Study

This research sought to explore the effects of three types of CMF in multimedia settings – elicitation, recasts, and meta-linguistic – on the learning of three major types of phrases in English, that is, participial, gerund, and infinitive, among Iranian intermediate-level EFL learners. Elicitation can be defined as a type of feedback that prompts learners to self-correct without giving them explanations of errors. Recast, as defined earlier, involves the reformulation of part or all of an erroneous form using acceptable target-like items. Meta-linguistic feedback can be conceptualized as the type of feedback that gives extra explanations on erroneous structures followed by further examples of target-like forms (Heift, 2004). The study, then, aimed to compare both immediate and long-term effects of these feedback types in form-focused computer-human exchanges in an attempt to ascertain which type(s) would prove more beneficial in the learning of L2 forms. Given that the very types of computer-mediated feedback might have had varying effectiveness, the present study purported to inform current design decisions as regards what type(s) of feedback could have served as the best candidate to be incorporated in multimedia tutorial systems. To this end, the present study sought to find an empirically justified answer to the following question:

Is there any statistically significant difference among the use of the multimedia courseware drawing on elicitation feedback, the one using

recast, and the one supplying meta-linguistic feedback on the learning of English phrases by Iranian intermediate-level EFL learners?

4. Method

4.1 Participants

The participants were chosen from among freshmen who were majoring in Teaching English as a Foreign Language (TEFL) at two Iranian universities. To select the site of the study, necessary arrangements were then made with one of the two universities' officials to dedicate a whole computer lab to hosting the participants during the experiment. One hundred and eighty-four students were identified as intermediate-level EFL learners based on the scores they obtained on a UCLES IELTS test of language proficiency. Out of this number, however, 160 candidates were ultimately chosen to take part in the study, given the imposed constraints in terms of the available budget allocated by the research deputies, and the capacity of the computer lab, which could only host a limited number of individuals. For the purpose of the present study, it was hypothesized that intermediate-level students were the most appropriate candidates, as there was little likelihood that they had had prior familiarity with different types of English phrases, but had been equipped with an average vocabulary repertoire and a relative command of grammar to ensure their successful interaction with the computer during form-focused exchanges. The assumption was also based on the researchers' impression of teaching grammar courses to different groups of students during years of instruction, as well as their diagnoses of students' errors when taking grammar tests. It should also be noted that the students did not need to have any prior computer knowledge, as the tutorial system embedded in the main research instruments, that is, the multimedia programs supplying form-focused exchanges, would already provide the prospective participants with instruction on how to interact with the computer. Once the participants were selected, they were then randomly assigned to four equivalent groups under study: one pilot group and three experimental groups. The groups were comprised of a mix of male and female participants.

4.2 Instruments

The main instruments involved three pieces of multimedia courseware authored by one of the researchers. The programs comprised an agent-based multimedia tutorial on English phrases where the agent embarked on providing the participants with explanations of and examples on the very types of phrases in English for approximately 30 minutes. The instruction module was divided into three parts: The first part of the tutorial was

allocated to the teaching of participial phrases, and the second and the third parts were dedicated to the teaching of gerund and infinitival phrases, respectively. The tutorial was then followed by 50 grammar exercises in the multiple-choice format where the participants supplied the blanks with the phrases they had chosen from the alternatives given. The programs monitored the students' progress and would give them feedback on erroneous structures in the form of an elicitation, a recast, or a meta-linguistic explanation. Using ActionScript Programming, an object-oriented programming language for Macromedia's Flash authoring tools, and "shared objects" which act like "cookies", the programs also logged the frequency with which the students did the exercises correctly without receiving feedback, the frequency with which they did the exercises correctly when received feedback, as well as the frequency with which they received feedback, but still did the exercises incorrectly.

Other instruments involved a test of language proficiency based on the UCLES IELTS examination papers that was used to help the researchers to choose the participants of the desired proficiency level; a multimedia pre-test of grammar to measure the participants' prior knowledge of the phrases and to ascertain the groups' homogeneity at the beginning of the study; and two multimedia post-tests of grammar measuring the relative effectiveness of the feedback types immediately and two weeks after the treatment. The pre-test, immediate and the delayed post-test of grammar each comprised 30 different items in the multiple-choice format, which measured the participants' knowledge of the three major English phrases. Each item correctly answered would receive a score of one mark and the total possible would be 30. The students could move the mouse cursor over the correct option and choose it with a single click. A built-in countdown timer too controlled the amount of time the participants spent on answering the questions. The multimedia tests also featured the capability of logging the students' answers for further analyses.

4.3 Procedures

At the onset of the study, the IELTS proficiency test was administered to the prospective participants from whom 184 students were then identified as intermediate-level learners following the rating scale proposed by the UCLES. All the candidates who obtained 4.5 to 5 on the test were identified as a "modest user", which is analogous to learners at an intermediate level of language proficiency. Since more than 160 individuals got the required band score, a simple random sampling was utilized to randomly select the required number of the participants. Since the utilization of simple random sampling was very cumbersome, the researchers drew on a digital

randomizer called SuperCool Random Number Generator[1]. The generator features the caliber to randomize a set of numbers from within a range. Accordingly, for both random selection and random assignment, the software potential was exploited to ease the sampling procedure. To this end, first all the candidates received a number from 1 to 184. The program then randomized the numbers in such a way that all the individuals had an equal chance of being among those 160 participants who would ultimately take part in the study. The program then produced a table where the names of the individuals and their corresponding numbers appeared. For random assignment, the researchers followed the same procedure: First, all the 160 candidates received a number and then the program produced four columns of individuals with 40 candidates in each. Once the required number of the participants was chosen, the randomizer then randomly assigned each of the four columns of individuals to one of the four groups under study.

One group served as a pilot group with whom the multimedia pre- and post-tests of grammar under the study underwent standardization. To standardize the tests, first item statistics of individual test items had to be clearly specified. In so doing, the researchers drew on an item statistics analyzer called Test Analysis Program (TAP) which was developed by Brooks, a professor at Ohio University[2]. An alternative could be SIMSTAT which was developed by Provalis Research Inc[3]. TAP, however, features the capability to mark defective items with an asterisk, that is, the ones whose IF and ID indices are not within the desirable ranges ($0.37 \leq IF \leq 0.63$ & $ID \geq 0.40$). Based on the analysis result, out of 30 items, two items on the pre-test, four items on the immediate post-test, and two items on the delayed post-test were shown to malfunction. These were accordingly excluded, and the participants' papers were rescored.

Next, the construct validly of the tests were established with the help of a factor analysis which was run on SPSS. The statistical program then utilized the Principle Components Extraction technique so as to extract all hypothetical factors whose "eigenvalues" were well beyond unity. Eigenvalue is the amount of variance in the test accounted for by the factors correlating with the test items. Typically, in a good test, only a small number of factors contribute significantly to the total variance of the test, while all other (undesirable, perhaps) factors contribute little or nothing to the total variance. The analysis revealed that as of the pre-test of grammar, around 34% of the test's total variance was accounted for by a single factor which was assumed to represent the construct of interest (knowledge of grammar), while a little amount of it was accounted for by other potential factors involved. This was promising, as the researchers could conclude that the test items highly correlated with the latent construct. As to the immediate and

the delayed post-tests too, the analysis showed that only one factor greatly contributed to the test's total variance. Here, approximately 32% and 37% of the total variance were contributed by the participants' grammatical knowledge for the immediate and delayed post-tests, respectively.

The statistical program also drew up a scree plot of factors against the eigenvalues. Each of the screes represented a hypothetical factor correlating with the test items. The point at which the eigenvalues began to level off could, then, be considered as a cut-off point. As far as the grammar tests were concerned, it was expected, therefore, that only one factor (scree), representing the grammatical competence, correlated highly with the tests items. Fingers 1, 2, and 3 show the loading of factors on individual test's items.

Figure 1. Pre-test scree plot

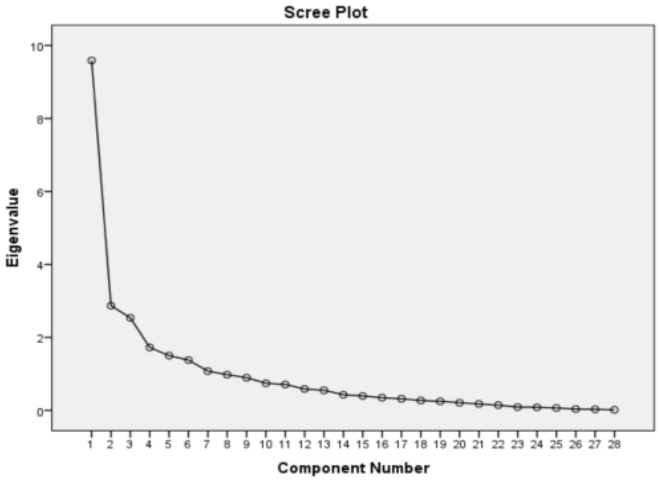


Figure 2. Immediate Post-test scree plot

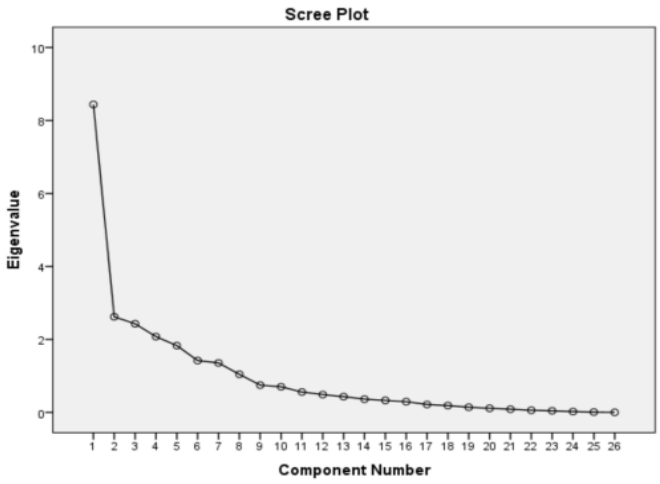
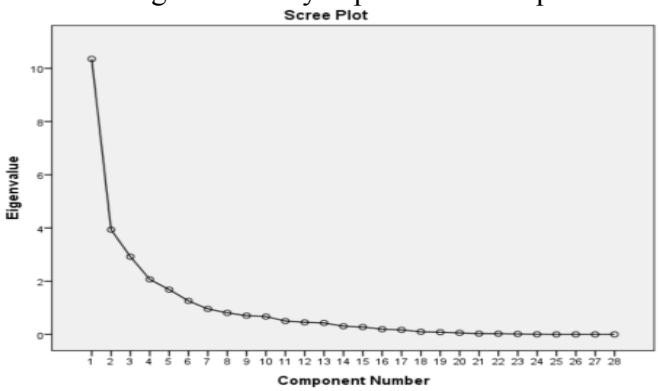


Figure 3. Delayed post-test scree plot



As can be seen in the plots, only one scree with the greatest eigenvalue is on the line with a steep slope. This further corroborated the idea that all the three tests could measure the construct of interest quite well. Once the construct validity of the tests was established, the reliability index of the three tests was calculated through a Cronbach's Alpha. It turned out to be 0.91, 0.88, and 0.92, for the pre-test, immediate post-test, and the delayed post-test, respectively.

Next, the multimedia pre-test of grammar was administered to the other three groups under study, which served as the experimental groups. The pre-test of grammar measured the participants' prior knowledge of the very types of English phrases at the beginning of the experiment. The purpose of pre-testing was twofold: On the one hand, it sought to determine the participants' familiarity with the English phrases. It was hypothesized that a great many EFL learners had difficulty recognizing and producing these types of phrases in different contexts of usage. On the other hand, it sought to ascertain the homogeneity of the groups in regards to the types of phrases to be introduced.

The result of the test revealed that the participants delivered a poor performance on the test and that the difference among the mean scores was not statistically significant ($p \geq 0.05$). The mean scores obtained were $\overline{X}_A$= 14.02, $\overline{X}_B$ = 13.70, and $\overline{X}_C$ = 13.42, which implied that the groups would need to receive treatment on the very types of phrases. To this end, the participants were requested to sit at the computer terminals at Islamic Azad University's computer lab. At the researchers' signal, all the participants inserted the programs' CD into the CD/DVD-ROM drivers and they automatically launched. The programs then embarked on the teaching of the three types of phrases using a computer agent providing explanations on and examples of the instances and contexts in which the phrases were used. In

introducing participial phrases, for instance, the virtual tutor explained that they comprised word groups involving a present participle or past participle, plus any modifiers, objects, and complements, and that they would commonly function as adjectives. The tutorial then gave examples of sample structures in which participial phrases could have appeared; sentences like "The girl *talking to the teacher* is very intelligent." and "The police caught *the young boy stealing a car*." provided examples. As to the gerund phrases, the tutor explained that they were participial forms used in noun functions and hence might have served as the subject of the verb, object of the verb, object of preposition, subjective compliment, and appositives. Again, the tutorial was followed by striking examples, demonstrating the common usage of the phrases. Sentences like "*Her cleaning the house everyday* is not necessary." or "We appreciate *your helping our friends last night*." provided examples. Likewise, to introduce infinitival phrases, the tutor first explained that these types of phrases would begin with an infinitive and could function as nouns, adjectives, and adverbs. The tutorial then gave examples of instances where theses phrases served the aforementioned functions; sentences like "To finish her shift without spilling another pizza into a customer's lap is Michelle's only goal tonight." or "The best way to survive Dr. Peterson's boring history lectures is a sharp pencil to stab in your thigh if you catch yourself drifting off." provided examples.

The instruction was then followed by 50 grammar exercises where the participants were required to supply the blanks with appropriate phrases from the options given. Concurrently, the programs monitored the students' progress where the participants' choice was first processed and then matched with the appropriate option in the programs' database. In case there was a mismatch between the participants' output and the correct option in the programs' database, a window would pop up providing the students with one of the aforementioned types of computer-mediated feedback. The programs were different in the type of feedback they supplied the students with. One experimental group received feedback through a multimedia program providing elicitation feedback on the subjects' errors. This type of CMF would use indirect cues to prompt learners to self-correct their errors. For instance, as of participial phrases, elicitation could involve letting the learners know which action had taken place before a second action. Similarly, to provide feedback on erroneous gerund phrases produced by the learners, the elicitation feedback would cue the participants which of the phrases in the sentence served as the subject or the object of the structure. Likewise, the elicitation CMF would prompt the participants to correct erroneous infinitival phrases by helping them recognize which phrase described the purpose for the action in the main clause. As an example, in

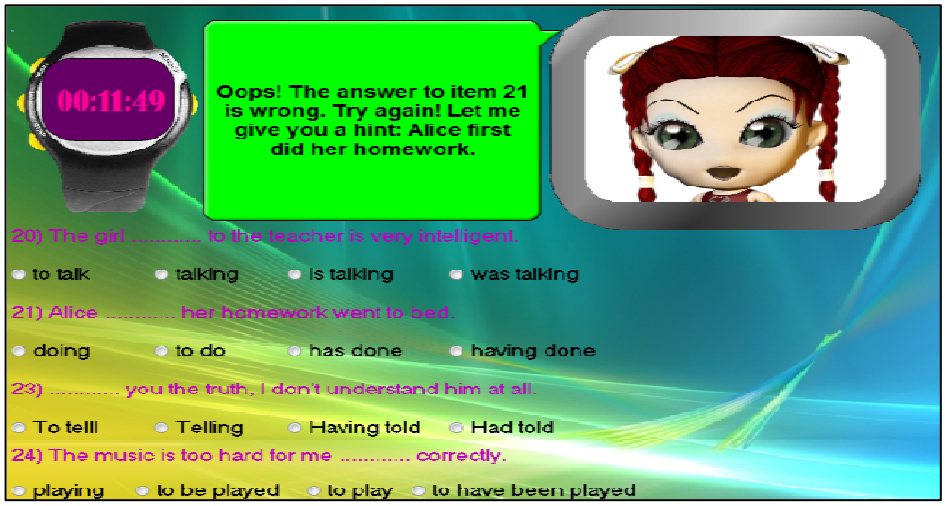
case the students were unable to supply the blank with the correct verb form in the sentence "He rejoiced …… him again.", where an infinitive is indeed required, the elicitation feedback would let the participants know the verb here was used to give reason(s) for why the person rejoiced. These cues were, then, assumed to persist in students' memory, thus aiding in the retrieval of the correct form of the phrases.

The second experimental group likewise received feedback through a similar condition, but received recast as feedback. Here, the correct form of the phrase was given by the program in order to further persist in the participants' memory. Indeed, what distinguishes elicitation from recast is that in elicitation students receive indirect cues by the computer as a prompt to remember the correct form whereas in recasts the whole phrase is built by the program. In a similar vein, the third experimental group received feedback through a multimedia program providing meta-linguistic explanations on the types of errors made. This type of feedback would give extra explanations on and examples of the phrases being introduced. Indeed, what distinguishes elicitation from meta-linguistic feedback in multimedia settings is that while the former gives clues to the learners to self-correct their errors and the students do not receive any extra explanations on the errors made, the latter provides the participants with further explanations as to when to use the very types of phrases, as well as with example structures where the common usage of the phrases is evident. In antithesis to recast, however, neither elicitation nor meta-linguistic feedback would provide the participants with direct clues as to the correct forms of the phrases. Figure 4 shows the tutorial screen and Figures 5, 6, and 7 illustrate the inherent differences among the three types of CMF feedback.

Figure 4. The tutorial screen

As can be seen in Figure 4, the robot guide is introducing the gerund phrase, supplying the participants with example structures where the phrases could be used. The tutorial lasted for approximately 30 minutes.

Figure 5. Elicitation in multimedia grammar exercises

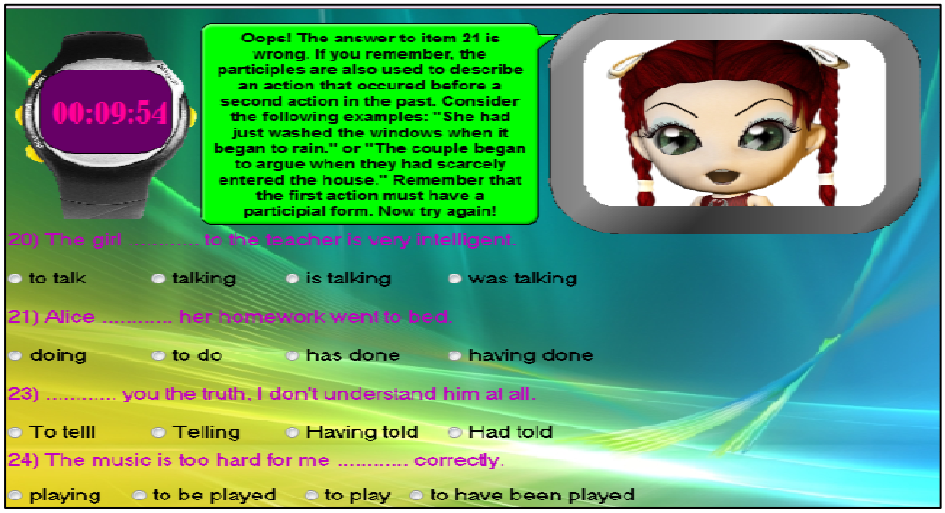


As can be seen in Figure 5, this type of feedback is describing which action had occurred before a second action in the past. It does not, however, directly guide the participant toward the correct form. No further example structure is also given.

Figure 6. Recast in multimedia grammar exercises

As shown by Figure 6, the whole structure is reformulated through recasting to help the learner to self-correct in an indirect way. Again, no further explanations or examples are given.

Figure 7. Meta-linguistic feedback in multimedia grammar exercises



As can be seen in Figure 7, as opposed to elicitation and recast, further explanations are given on the student's errors, together with example structures which indirectly guide the learner to correct his error. Like elicitation, no direct clue is offered. The feedback would be given to the participants only when they did the exercises incorrectly. Moreover, the feedback would be given only once, that is, in case the students failed to do the exercises correctly a second time, they would no longer receive feedback, and their answers would be logged. One rationale was that the researchers had to ensure that all the three groups would receive the same amount of feedback during the experiment. As the participants were doing the exercises, the frequency with which they did the exercises correctly with or without the provision of feedback, as well as the frequency with which they did the exercises incorrectly when received feedback were also logged by the programs and saved in their profiles on their PCs' hard drives. These frequencies were logged so that the researchers could then ascertain the extent to which the three types of computer-mediated feedback had proved fruitful in helping the participants to self-correct. The frequency tables in the "Results and Discussion" section demonstrate these frequencies.

The participants spent approximately 60 minutes doing the exercises after which they were given a short break for refreshment. They were not, however, allowed to leave the experiment setting. Next, an immediate post-

test of grammar comprising a different set of items in the multiple-choice format was administered to the three groups under study so that the researchers could determine and compare the immediate effects of these feedback types, given through form-focused computer-human interactions, on the learning of English phrases among the participants. The multimedia post-test of grammar, however, did not provide the participants with any type of feedback as they approached the test items. Like the multimedia pre-test, the immediate post-test of grammar comprised 30 items and it would automatically calculate and log the students' right and wrong answers in their profiles and save them in shared objects on their PCs' hard drives for later analysis. The result of the immediate post-test appeared in the "Results and Discussion" section. In a similar vein, to examine the long-term effects of the three types of CMF feedback, a delayed post-test of grammar comprising 30 different items was also administered to the three groups under study two weeks after the experiment. As the efficiency of these feedback types in form-focused exchanges could be best captured by examining the degree to which their effectiveness had been sustained over time, the result of the delayed podst-test of grammar would be of great consequence. The result of this test, too, appeared in the "Results and Discussion" section.

5. Results and Discussion

The present study sought to address the following question: "Is there any statistically significant difference among the use of the multimedia courseware drawing on elicitation feedback, the one using recast, and the one supplying meta-linguistic feedback on the learning of English phrases by Iranian intermediate-level EFL learners?" The experiment showed that the three types of CMF feedback yielded varying effects, thus leading to varying degrees of learning success. Given that the three groups under study were homogeneous at the beginning of the experiment and that the multimedia programs controlled the amount of instruction across the three groups, the difference in the gains on the two post-tests can be largely accounted for by the immediate and sustained effects of these feedback types. The frequency tables below, under the rubric of "descriptive statistics of feedback types", further corroborate this idea, as the mean frequency with which participants in the three groups did the exercises correctly without receiving feedback was approximately the same across the three groups.

Table 1. Descriptive statistics of feedback types

Frequency Table A = Feedback Type: Elicitation			
Frequency / Student	Number of times students did the excersises correctly without receiving feedabck	Number of times students did the excercises correctly once received feedback	Number of times students received feedback, but still did the excercises incorrectly
1	10	35	5
2	15	25	10
3	20	29	1
4	20	25	5
5	20	30	0
6	15	22	13
7	9	37	4
8	12	28	10
9	12	25	13
10	20	20	10
11	14	30	6
12	30	20	0
13	23	27	0
14	24	18	8
15	24	12	14
16	42	8	0
17	19	23	8
18	23	25	2
19	19	30	1
20	45	5	0
21	12	32	6
22	18	30	2
23	20	23	7
24	17	33	0
25	14	26	10
26	20	23	7
27	14	36	0
28	26	24	0
29	34	16	0
30	30	16	4
31	23	22	5
32	22	28	0
33	15	35	0
34	20	30	0
35	15	30	5
36	19	30	1
37	8	32	10
38	10	34	6
39	10	36	4
40	28	20	2
Mean	19.78	25.75	4.48

Frequency Table C = Feedback Type: Meta-Linguistic			
Frequency / Student	Number of times students did the excersises correctly without receiving feedabck	Number of times students did the excercises correctly once received feedback	Number of times students received feedback, but still did the excercises incorrectly
1	22	28	0
2	20	30	0
3	14	36	0
4	16	30	4
5	23	27	0
6	28	22	0
7	20	30	0
8	13	37	0
9	25	20	5
10	23	27	0
11	15	35	0
12	20	30	0
13	9	41	0
14	20	25	5
15	15	32	3
16	34	16	0
17	20	30	0
18	18	32	0
19	13	36	1
20	26	24	0
21	22	18	10
22	18	30	2
23	14	28	8
24	16	34	0
25	38	10	2
26	16	34	0
27	24	20	6
28	42	6	2
29	22	28	0
30	21	27	2
31	10	27	13
32	35	10	5
33	12	38	0
34	20	30	0
35	18	32	0
36	27	23	0
37	28	22	0
38	19	30	1
39	12	38	0
40	10	40	0
Mean	20.45	27.83	1.73

Frequency Table B = Feedback Type: Recast			
Frequency / Student	Number of times students did the exercises correctly without receiving feedabck	Number of times students did the exercises correctly once received feedback	Number of times students received feedback, but still did the exercises incorrectly
1	24	26	0
2	34	16	0
3	30	20	0
4	22	28	0
5	22	28	0
6	20	30	0
7	28	22	0
8	28	22	0
9	20	30	0
10	20	30	0
11	16	34	0
12	25	25	0
13	23	27	0
14	23	27	0
15	23	27	0
16	42	8	0
17	22	28	0
18	20	30	0
19	12	38	0
20	14	36	0
21	14	36	0
22	20	30	0
23	9	41	0
24	12	38	0
25	19	31	0
26	30	20	0
27	20	30	0
28	20	30	0
29	18	32	0
30	45	5	0
31	27	23	0
32	26	24	0
33	18	32	0
34	28	22	0
35	15	35	0
36	17	33	0
37	20	30	0
38	14	36	0
39	20	30	0
40	14	36	0
Mean	21.85	28.15	0.00

As can be seen in the frequency tables, out of 50 grammar exercises, the mean of the times the experimental groups, receiving elicitation, recast, and meta-linguistics feedback, did the exercises correctly are $\overline{X}_A = 25.75$, $\overline{X}_B = 28.15$, and $\overline{X}_C = 27.83$, respectively. This exceeds the mean of the times they did the exercises correctly without receiving feedback ($\overline{X}_A = 19.78$, $\overline{X}_B = 21.85$, & $\overline{X}_C = 20.45$), as well as the mean of the times they received feedback, but still did the exercises incorrectly ($\overline{X}_A = 4.48$, $\overline{X}_B = 0.00$, & $\overline{X}_C = 1.73$). Given that all the three groups received the same amount of instruction, the difference in the means can be attributed to the varying effects of the computer-mediated feedbacks. It is also evident that the mean frequency with which the participants received feedback and did the exercises correctly is much greater the mean frequency with which they received feedback and still did the exercises incorrectly. This mean frequency is 0.00 for recast; this is typical, as the type of computer-mediated feedback given in the form of recasts under this study involved the provision of the correct forms of the phrases. Yet the mean frequency for the other two types of CMF is also low, thus implying that feedback has the potential to supplement instruction and hence increase its efficiency through form-focused exchanges.

Table 2 shows the result of the immediate post-test. As can be seen, the difference among the mean scores is statistically significant ($p \leq 0.05$). According to Table 3, the significance value for ANOVA analysis is smaller than the alpha value ($p \leq 0.05$); the groups were heterogeneous. To investigate where the difference is, one can have a glimpse at the result of the Scheffé's test (See Table 4). It is apparent that all the three means are statistically different and that one can contend that the three types of feedback had varying effects on the students' learning despite the fact that the participants obtained a higher mean on the immediate post-test as compared with the pre-test means. It is also evident that the third experimental group (Group C) who received meta-linguistic feedback outperformed the other two groups, and the one who received recast (Group B) outperformed the one who received elicitation (Group A).

Table 2. Descriptive statistics of the immediate posttest

Groups [a]	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
Group A	40	19.8000	2.85729	.45178	18.8862	20.7138	15.00	26.00
Group B	40	22.0500	2.45941	.38887	21.2634	22.8366	17.00	26.00
Group C	40	24.1250	2.06544	.32657	23.4644	24.7856	18.00	26.00
Total	120	21.9917	3.03342	.27691	21.4434	22.5400	15.00	26.00

[a] Group A = Elicitation, Group B = Recast, Group C = Meta-Linguistic

Table 3. ANOVA results of the posttest

ANOVA	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	374.317	2	187.158	30.385	.000
Within Groups	720.675	117	6.160		
Total	1094.992	119			

Table 4. Result of the scheffé's test

(I) Experimental Groups	(J) Experimental Groups	Mean Difference (I-J)	Std. Error	Sig.	Scheffé's *Test* Lower Bound	Upper Bound
A	B	-2.25000*	.55496	.000	-3.6260	-.8740
	C	-4.32500*	.55496	.000	-5.7010	-2.9490
B	A	2.25000*	.55496	.000	.8740	3.6260
	C	-2.07500*	.55496	.001	-3.4510	-.6990
C	A	4.32500*	.55496	.000	2.9490	5.7010
	B	2.07500*	.55496	.001	.6990	3.4510

* $p \leq 0.05$

Likewise, Tables 5, 6 and 7 show the results of the delayed posttest.

Table 5. Descriptive statistics of the delayed post-test

Descriptives	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean Lower Bound	Upper Bound	Minimum	Maximum
Group A	40	19.5500	2.71699	.42959	18.6811	20.4189	15.00	26.00
Group B	40	17.8000	2.10250	.33243	17.1276	18.4724	16.00	25.00
Group C	40	23.1750	2.51036	.39692	22.3721	23.9779	18.00	28.00
Total	120	20.1750	3.31450	.30257	19.5759	20.7741	15.00	28.00

Table 6. ANOVA results of the delayed posttest

ANOVA	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	601.250	2	300.625	49.815	.000
Within Groups	706.075	117	6.035		
Total	1307.325	119			

Table 7. Result of the scheffé's test

(I) Experimental Groups	(J) Experimental Groups	Mean Difference (I-J)	Std. Error	Sig.	Scheffé's *Test* Lower Bound	Upper Bound
A	B	1.75000*	.54931	.008	.3880	3.1120
	C	-3.62500*	.54931	.000	-4.9870	-2.2630
B	A	-1.75000*	.54931	.008	-3.1120	-.3880
	C	-5.37500*	.54931	.000	-6.7370	-4.0130
C	A	3.62500*	.54931	.000	2.2630	4.9870
	B	5.37500*	.54931	.000	4.0130	6.7370

As can be seen in Table 7, there was a slight decrease in the effects of elicitation and meta-linguistic feedback over time; nevertheless, the third experimental group (Group C) again, who received meta-linguistic feedback, outperformed the other two groups. Furthermore, the experimental group who received elicitation (Group A) outperformed the one (Group B) who received recast in form-focused exchanges. There was indeed a dramatic decrease in the effects of recast across the two post-tests, as the mean of Group B dropped from 22.05 to 17.80. This implies that in antithesis to the immediate post-test result, which favors the use of recast in computer-human exchanges, the effects of elicitation might be sustained longer, thus leading to better retrieval of the correct forms of English phrases.

One explanation for meta-linguistic feedback leading to higher gains on both tests is that it might have provided students with a rich, elaborate environment in which the richness of explanations and contextual cues might have led to more efficient learning and hence better storage of information. Indeed, what makes meta-linguistic feedback even more efficient is that it might lead to more successful uptakes on the part of the students. The idea is favored by Loewen (2007) when he argues that successful uptake can be conceptualized as the incorporation of the linguistic information provided in the feedback into the learners' current interlanguage. Successful uptake might, then, correlate highly with efficient storage and then retrieval of linguistic information on the leaner's part. In line with the same argument, Babaie and Khalili (2010) make the case that efficient encoding and retrieval of linguistic information are highly correlated with contextual richness when elaboration on linguistic information can persist in individuals' working memory, and consequently, lead to encoding of information in their long-term memory for long-term retrieval.

Still another justification is that the researchers contend that in meta-linguistic feedback, there are two modalities of elaboration. On the one

hand, the very type of feedback provides further explanations on linguistic form to be incorporated into leaner's interlanguage; on the other hand, it provides learners with further example structures where the linguistic form could appear. Accordingly, there are two modes of elaboration: explanations on errors, and examples on common usage of the linguistic form. As of elicitation and recast, only one modality exists: in elicitation, for instance, the CMF feedback can describe the verb's action in phrases so as to help learners identify their errors and correct them; however, no further example structures are given. In recast, only the correct form of the phrase or linguistic form is given with no explanations on learners' errors or example structures providing more contexts of usage. Overall, it can be argued that elaboration modality might play a key role in the efficiency of CMF feedback in terms of their immediate and long-term effects on learning linguistic structures.

The study further reveals that recast might produce stronger immediate effects as compared to those of elicitation, but its effects might not sustain as long as those of elicitation. As shown by the findings, the participants who received recast through form-focused exchanges obtained a higher mean on the immediate post-test as compared with the ones who received elicitation. Notwithstanding, the experimental group who received elicitation outperformed the one who received recast on the delayed post-test. The stronger immediate effects of recast can be explained by the idea that when learners are given the correct form of the linguistic forms, the form may temporarily persist in their working memory.

The form may, then, reside in the working memory and hence, whenever it appears in the test items, it would readily be picked up by the learners. Yet since the students receive no explanation for the errors they make, the provision of recast may not necessarily lead to more efficient, more meaningful learning of the linguistic forms. Indeed, as Loewen (2007) aptly puts, learners might temporarily respond to the feedback they receive, but this does not mean that they have also successfully incorporated the linguistic information. Based on what he claims, it can, then, be argued that learners also reap benefits from explicit explanations on their errors which may lead to more efficient learning and hence storage of correct structures. While the provision of correct forms may aid in transient storage and short-term retrieval of linguistic structure, this by no means ensures efficient incorporation of the forms into learners' interlaguage.

What distinguishes elicitation from recast is, then, the fact that inherent in their difference is the explanation on learners' errors given by elicitation where learners' awareness of the nature of their errors might lead to more successful uptake and hence greater sustained effects on their learning. This

also suggests that explanation on errors, relative to, the mere reformulation of linguistic forms as is evident in recast, is more beneficial to learners' incorporation of those forms into their linguistic knowledge. In sum, it can be argued that explanations on learners' errors are more consequential in successful incorporation of linguistic forms than the mere provision of example structures or the correct forms of structures and that the integration of explanations and example structures, as is the case with meta-linguistic feedback, is even more crucial in efficient encoding of linguistic structures.

Given the findings of the present study, it is suggested, therefore, that teachers or teachers as designers draw heavily on either elicitation or meta-linguistic feedback for their long-term effects on learning and retrieval of English phrases. Agent-based multimedia courseware combined with CMF feedback might, then, offer great promise for efficient grammar instruction among EFL or perhaps ESL learners. The integration of tutorials and accompanying multimedia grammar exercises can provide learners with opportunities for further exposure to salient linguistic features and discourse paradigms of the language through engaging learners in computer-human form-focused exchanges. The study further suggests that the provision of CMF has the potential to increase the perceptual salience of linguistic forms and that different types of feedback might increase this saliency to varying degrees. Regrettably, however, research on CMF feedback is sparse and hence future studies should corroborate such views.

6. Conclusion and Implications

The present study compared the immediate and long-term effects of three major types of computer-mediated feedback, that is, elicitation, recast and meta-linguistic, on the learning of three main types of English phrases, that is participial, gerund, and infinitival, among Iranian intermediate-level EFL learners. The experiment revealed feedback modality mattered and it exerted a potential influence on the outcome of learning English phrases. The experiment showed that while recast produced a strong immediate effect on learning when compared to that of elicitation, its potential effect significantly dropped over time. On the other hand, there was a slight decrease in the effects of elicitation and meta-linguistic feedback across the two tests, with elicitation producing a longer-term effect on learning as compared with that of recast. Yet it was shown, both the immediate and long-term effects of meta-linguistic feedback were comparatively higher than those of recast and elicitation. The study, then, favors the use of meta-linguistic feedback where a combination of examples and supplementary explanations could add to the elaboration of information presented. It is suggested, therefore, that this type of feedback be incorporated in the

instruction of grammar so that further elaboration on the types of structures being introduced could enhance the visual saliency and hence more efficient noticing and encoding of information in memory.

7. Recommendations for the Follow-Up

This study explored the effects of three feedback modalities on the learning of three major types of phrases in English. A major limitation of the present study is that only the participants' recognition ability in recognizing the correct forms of the phrases was focused on. During form-focused exchanges, once the feedback was provided by the computer, the participants were required to choose the correct forms of the phrases from the alternatives given. Accordingly, their successful incorporation of linguistic forms into their current linguistic system was only measured through their recognition ability. One rationale for why the researchers confined the testing situation to the participants' recognition ability through multiple-choice items was that the scoring mechanism utilized by the computer was unable to recognize degree of responses produced by the participants. To measure learners' production of linguistic forms, one may, then, need to do more sophisticated programming where the computer draws on a cue-based interpretative algorithm to score suppletion-type items. In this case, future researchers may also need to build a large database of possible answers to have the computer compare learners' production with the best match in its database. Still another suggestion for follow-up studies is that the present research focused on the effects of CMF feedback on the learning of English phrases. Different types of structures in English, including but not limited to conditionals, causative structures, restrictive and non-restrictive clauses, and so on, could be introduced through multimedia tutorial systems and their incorporation into students' linguistic system can then be explored through the provision of different types of CMF feedback in such environments. Future studies may also explore the effects of other types of CMF on the learning of grammatical structures. Chief among them are "clarification" which takes the form of "Try again!" in multimedia environments and do not provide learners with either an explanation on their errors or the correct forms of their production, and "repetition" which is manifested by a rising intonation and a question mark appearing at the end of the learners' production to signal that they have made errors (Heift, 2004).

References

Ammar, A., & Spada, N. (2006). One size fits all? Recasts, prompts and L2 learning. *Studies in Second Language Acquisition*, *28*, 543-574.

Babaie, H. (2008). On the effects of help options in MCALL programs on the listening comprehension of EFL learners. *Journal of TELL*, Tarbiat Modares University, *2*(6), 27-47.

Babaie, H., & Khalili, M. (2010). Pictorial, textual, and picto-textual glosses in e-reading: A comparative study. *Canadian Journal of English Language Teaching (ELT)*, *3*(4), 1-11.

Beatty, K. (2003). *Teaching and researching computer-assisted language learning*. London: Longman.

Beaudoin, M. (2004). Educational use of databases in CALL. *Computer Assisted Language Learning, 77*(5), 497-516.

Brandl, K. K. (1995). Strong and weak students' preferences for error feedback options and responses. *The Modern Language Journal, 79,* 194-211.

Carroll, S. (2001). *Input and evidence: The raw material of second language acquisition. Amsterdam*: John Benjamins.

Chapelle, C. A. (2001). *Computer applications in second language acquisition: Foundations for teaching, testing and research*. Cambridge: Cambridge University Press.

Chapelle, C. A. (2003). *English language learning and technology.* Amsterdam: John Benjamins.

Chapelle, C. A. (2005). Interactionist SLA theory in CALL research. In J. Egbert & G. M. Petrie (Eds.). *CALL research perspectives* (pp. 53-64). Mahwah, NJ: Lawrence Erlbaum Associates.

Collentine, J. (2000). Insights into the construction of grammatical knowledge provided by user-behavior tracking technologies. *Language Learning & Technology, 5*(2), 44-57.

Cowan, R., Choi, H. E., & Kim, D. H. (2003). Four questions for error diagnosis and correction in CALL. *CALICO Journal*, *20*, 451-463.

DiGiovanni, E. & Nagaswami, G. (2001). Online peer review: an alternative to face-to-face? *ELT Journal*, *55* (3), 263-272.

DeKeyser, R. M. (2005). What makes learning second language grammar difficult? A review of issues. Language Learning, *55*(1), 1-25.

Ellis, N. (2005). At the interface: dynamic interactions and explicit and implicit language knowledge. *Studies in Second Language Acquisition*, *27*(1), 305-352.

Ellis, R. (2009). Corrective feedback and teacher development. *L2 Journal*, *1*, 3-18.

Ellis, R., Loewen, S., & Erlam, R. (2006). Implicit and explicit corrective feedback and the acquisition of 12 grammar. *Studies in Second Language Acquisition*, *28*, 339-368.

Gass, S. M. (1997). *Input, interaction, and the second language learner*. Mahwah, NJ: Lawrence Erlbaum Associates.

Gass, S. M., & Torres, M. J. (2005). Attention when? An investigation of the ordering effect of input and interaction. *SSLA*, *27*, 1-31.

Glendinning, E., & Howard, R. (2003). Lotus ScreenCam as an aid to investigating student writing. *Computer Assisted Language Learning, 16*, 31-46.

Greenfield, R. (2003). Collaborative e-mail exchange for teaching secondary ESL: A case study in Hong Kong. *Language Learning and Technology*, *7* (1), 46-70.

Heift, T. (2004). *Corrective feedback and learner uptake in CALL. ReCALL, 16*(2), 416–431.

Long, M. H. (1996) The role of linguistic environment in second language acquisition. In W. C. Ritchie & T. K. Bhatia (Eds.). *Handbook of Second Language Acquisition* (pp. 413-468). San Diego, CA: Academic Press.

Loewen, S. (2007). The prior and subsequent use of forms targeted in incidental focus on form. In H. Nassaji, & S. Fotos (Eds.). *Form Focused Instruction and Teacher Education: Studies in Honour of Rod Ellis* (pp. 101-116). Oxford: Oxford University Press.

Loewen, S., & Erlam, R. (2006). Corrective feedback in the chatroom: An experimental study. *Computer Assisted Language Learning 19*(1), 1-14.

Loewen, S., & Nabei, T. (2007). Measuring the effects of oral corrective feedback on L2 knowledge. In A. Mackey (Ed.). *Conversational interaction in second language acquisition: A collection of empirical studies* (pp. 361-377). Oxford: Oxford University Press.

Long, M. H. (2007). *Problems in SLA*. Mahwah, NJ: Lawrence Erlbaum.

Long, M. H., & Robinson, P. (1998). Focus on form: Theory, research and practice. In C. Doughty, & J. Williams (Eds.). *Focus on form in second language acquisition* (pp. 15-41). Cambridge: Cambridge University Press.

Lyster, R. (2004). Differential effects of prompts and recasts in form-focused instruction. *Studies in Second Language Acquisition*, *26*, 399-426.

McDonough, K. (2005). Identifying the impact of negative feedback and learners' responses to ESL question development. *Studies in Second Language Acquisition*, *27*, 79-103.

Noonan, F. J. (2004). Teaching ESL students to "notice" grammar. *The Internet TESL Journal*, *10*(7), 1-10.

Panova, I., & Lyster, R. (2002). Patterns of corrective feedback and uptake in an adult ESL classroom. *TESOL Quarterly, 36*(4), 573-595.

Petersen, K. A. (2010). *Implicit corrective feedback in computer-guided interaction: Does mode matter?* (Doctoral dissertain). Retrieved from ProQuest Dissertations and Theses Database. (AAT 3450858).

Pica, T. (1994). Research on negotiation: What does it reveal about second-language learning conditions, processes and outcomes? *Language Learning*, *44*, 493-527.

Sachs, R., & Suh, B., (2007). Textually enhanced recasts, learner awareness, and L2 outcomes in synchronous computer-mediated interaction. In A. Mackey (Ed.). *Conversational interaction in second language acquisition: A collection of empirical studies* (pp. 197-227). Oxford: Oxford University Press.

Sauro, S. (2009). Computer-mediated corrective feedback and the development of L2 grammar. *Language Learning & Technology*, *13*(1), 96-120. Retrieved October 18, 2011, from http://llt.msu.edu/vol13num1/sauro.pdf

Sharwood Smith, M. (1993). Input enhancement in instructed SLA. *Studies in Second Language Acquisition*, *15*, 165-179.

Schmidt, R. (1990). The role of consciousness in second language learning. *Applied Linguistics, 11*(2), 129-158.

Schmidt, R. (2001). Attention. In P. Robinson (Ed.). *Cognition and second language instruction* (pp. 3-32). Cambridge: Cambridge University Press.

Torlakovic, E., & Deugo, D. (2004). Application of a CALL system in the acquisition of adverbs in English. *Computer Assisted Language Learning*, *17*(2), 203-235.

Van den Branden, K. (1997). Effects of negotiation on language learners' output. *Language Learning*, *47*, 589-636.

Warschauer, M. (1996). Comparing face-to-face and electronic discussion in the second language classroom. *CALICO Journal*, *13*(2), 7-26.

Zamel, V. (1981). Cybernetics: A model for feedback in the ESL classroom. *TESOL Quarterly*, *15*, 139-150.

Notes:

[1] The randomizer can be downloaded via:
http://www.supercoolbookmark.com/download/supercoolrandom104.zip

[2] TAP can be retrieved via:
http://www.ohio.edu/people/brooksg/downloads/tap.exe

[3] SIMSTAT can be downloaded through:
http://www.pro\valisresearch.com/simstat.php