THE EFFECT OF THE COLLABORATION OF REFLECTIVE NOTES WITH CALL ON EFL LEARNERS’ WRITING ACCURACY

Fatemeh Soltanpour¹, Mohammadreza Valizadeh²

¹English Department, Kar Higher Education Institute, PO Box 1571, Qazvin, Iran
²English Department, Faculty of Foreign Languages, Gazi University, PO Box 06500, Ankara, Turkey

Abstract:
The aims of this mixed-method action research were: (1) to investigate the effect of feedback provided by MS word processor on EFL learners’ writing accuracy in the context of a university in Iran, (2) to explore whether taking reflective notes (henceforth. RNs) in collaboration with the word processor can modify the effect of such received feedback, and (3) to find out what the participants thought about each treatment. Two intact classes (Advanced Writing) were used, but the classes were randomly assigned to each treatment, called ‘CALL with RNs’ and ‘CALL without RNs’. There were 50 participants altogether who were B.A. English Translation majors. The Straightforward Quick Placement & Diagnostic Test was administered to ensure the participants were homogeneous. Each group received 10-sessions of treatment. Two samples of Task 2 of General Module of IELTS were used for the pretests and posttests. Each essay was scored independently by two raters. The final score consisted of the average score of the two raters. The findings revealed that the feedback provided by MS word processor improved the students’ writing accuracy significantly; however, the ‘CALL with RNs’ group outperformed the other one. To collect qualitative data, semi-structured interviews were conducted with each participant. Generally, the participants in both groups had positive attitudes towards receiving feedback by the word processor. Considering the RNs, the participants also had positive attitudes; nonetheless, there were a few students who didn’t like RNs technique, not because they found it useless, but mainly because of their individual differences.

¹ Correspondence: email: fatemeh.soltanpour@yahoo.com

Copyright © The Author(s). All Rights Reserved.
© 2015 – 2017 Open Access Publishing Group
Keywords: MS Word processor, CALL, reflective notes, feedback, writing accuracy, grammar, mechanics of writing

1. Introduction

As Barkaoui (2007) stated, acquiring L2 writing skill is one of the most difficult skills for L2 learners because it requires the mastery of a variety of linguistic, cognitive, and sociocultural competencies. Some scholars, such as Ferris and Roberts (2001) as well as Hyland and Hyland (2001) stated that, in addition to encouraging L2 learners to engage in writing frequently, they should be provided with helpful and appropriate feedback and support, so, as Meng (2013) pointed out, error correction has had a central position in language teaching. Regardless of the type of feedback that has been offered, some previous studies have proved written corrective feedback (WCF) can improve accuracy in second language learners’ writing performance (see, e.g., Bitchener, 2008; Bitchener & Knock, 2008; Bitchener & Knock, 2010a; Bitchener & Knock, 2010b; Ellis, Sheen, Murakami, & Takashima, 2008; Sheen, 2007; Sheen, Wright, & Moldawa, 2009; Van Beuningen, de Jong, & Kuiken, 2008).

Nevertheless, as the importance of learning second or foreign languages increased around the world, the number of L2 learners also increased in classes; as a result, providing WCF has become a difficult job for L2 teachers because providing the WCF is very time-consuming (Salteh & Sadeghi, 2012; Amrhein & Nassaji, 2010; Corpuz, 2011; Ellis, Loewen, & Erlam, 2006; Ferris, 1999; 2010; Lyster, 2004; Meng, 2013). In this regard, some scholars offered that technology can solve this problem and feedback provided by computers can reduce the teacher’s workload (Chapelle, 2001; Hegelheimer, 2004; Lavolette, Polio, Kahng, 2015; Pica, 1994). Nonetheless, feedback given by a computer could not replace the teacher feedback, but it can be a positive addition to it (Lavolette, et al. 2015).

On the other hand, as most writing instructors frequently observe and Leki, (1992, as cited in Howrey & Tanner, 2008) wrote, the main challenge writing teachers face is ensuring that learners transfer knowledge and skills from one composition to the next, learn from their mistakes, and acquire both writing fluency and accuracy. Despite teacher instructions and required revisions, learners often show little or no improvement in their writing. As Howrey and Tanner (2008) argued, learners often do not learn to take responsibility for their own writing, and often neglect teacher feedback on final drafts.

On the other hand, research has revealed that computer-assisted language learning (CALL) could be effective in helping the L2 learners improve their knowledge...
of English language and the skills relevant to it, (Blake, 2000; Chatel, 2002; Ghanbari, Shamsoddini, & Radmehr, 2015; Ghorbani & Marzban, 2013; Hayati, 2005; Naraghizadeh & Barimani, 2013; Niazi & Pourgharib, 2013; Warschauer & Healey, 1998). Another important issue considering the feedback provided by a computer is what some approaches to SLA, including the interaction approach, usage-based approaches, and sociocultural theory, have suggested and that is feedback on language needs to be immediate to be effective (Polio, 2012).

The efficacy of CALL can be supported by various second language learning theories. As Warschauer (2005) explained, Vygotsky's sociocultural theory can be applied to CALL. According to Vygotsky (1981, as cited in Warschauer, 2005), via mediation or the incorporation of tools or meditational means, the entire flow and structure of mental functions will alter. Warschauer (2005) mentioned computers as an example of those meditational means. CALL can also provide interactive learning environment and collaborative writing in the L2 learning, qualities that are in line with the social learning aspect of the Vygoskyian’s theory (Warschauer, Turbee, & Roberts, 1996) as well as Ellis’s (1999) perspective of interactionist SLA; Based on Chapelle (2005), interaction in CALL takes place not only in face-to-face conversation between the learners and teacher as well as peers but also electronically over a computer, i.e., between the learner and the computer. As Long (1996) and Pica (1994) stated, learning takes place through learners’ negotiation of meaning to make connections between form and meaning.

Furthermore, Li (2000) claimed that CALL can stress the role of social and affective factors in L2 learning. The affective factor could be in line with the flow theory. The flow concept was developed by Schiefele and Csikszentmihalyi (1994) as well as Csikszentmihalyi (1997a, 1997b, 1997c, 1997d, all cited in Egbert, 2005) and was defined as a psychological state that results in optimal experience. Abbot (2000) commented that flow experiences may help the L2 learners be more motivated and thus their language learning can be more effective; such experience might help learners to persevere in their language studies. Trevino and Webster (1992) noted that computer software can provide “varied, novel, and surprising stimuli” (p. 543) that makes the learners willing to continue; this feeling could lead to the flow experience. In addition, As Benson (2001) mentioned, in the case of CALL, there is “an assumption that technology can provide learners with the kinds of support they need in order to develop skills associated with autonomy” (p. 140). Abbott (2000) found contexts that support autonomy create more favorable conditions for flow than controlled contexts and tasks. As Egbert (2005) indicated, applying CALL causes the user to perceive a sense of control over the task that is being done, and as a result, based on Thanasouls (2000), such sense of control is important to
language learning. Ghani and Deshpande (1994) had also found that computer activities can support flow because they provide users with both challenge and control.

Moreover, Pennington (2004) as well as Pennington and Brock (1992) (both sources cited in Levy, 2009) stated that the word processor has become one of the most widely accepted technologies for writing since the early 1980s. As it facilitates the flexible manipulation of text, it enables the user to draft and redraft the composition easily and produce a professional product. Barrass (1995) stated that word-processing can help the L2 writers writing in all four stages of composition: thinking, planning, writing, and revising. During the writing process, a word-processor will automatically format text; may provide a choice of fonts; inserts running heading and page numbers; may enable one to check spelling, syntax, and grammar; may provide advice on the choice of words and on the use of words, and may provide a thesaurus. (p. 97)

According to Thomas (2004), word processing programs allow editing and formatting, spelling, and grammar checking; therefore, as Sergeant (2001) stated, typing a text word by word may help students remember and improve the knowledge of words and their spelling as well as the grammatical structures; it may also be more fun than copying a text using pen and paper. Moreover, Kenning (1996) said that word processors can be intrinsically supportive of cognitive and metacognitive autonomy because they encourage their users to consider their compositions critically and try to make improvements. Since 1980s, some research has been done into the effect of the word-processor on the writing of the L2 learners and positive satisfactory results have been revealed (see e.g, Barzegar, Fattahi Bafghi, & Allami, 2011; Bernhardt, Wojahn, & Edwards, 1989; Darus, Ismail, & Ismail, 2008; Jafarian, Soori, Kafipour, 2012; Lam & Pennington, 1995; Lee, 2004, Li, 2006; Li & Cumming, 2001; Owston, Murphy, & Wideman, 1992, 1997).

In addition to the above-mentioned issues, based on the constructivist view of language learning, students learn in problem-solving environments that challenge their prior knowledge and encourage them to reflect on the differences between their own knowledge and that of the course. Such an approach to learning can be applicable in the composition classroom (Nelson, 2002). According to constructivists, learning always involves both analyzing and transforming new information (O’ Donoghue & Clarke, 2010). As Duffy and Cunningham (1996, as cited in Callele, 2008) stated, within constructivist approaches, students are considered active participants in their learning. Students need to be involved in reviewing content in a variety of ways and at different times to acquire understanding of a subject or lesson, and teachers are responsible for providing them with such opportunities (Driscoll, 1994, as cited in Razak & Asmawi, 2004). Thus, because learners are in charge and control of what, when, and how they
learn (Driscoll, 2000; Hannafin, 1992, as cited in Callele, 2008), they need to be aware of their own thinking and learning processes (Driscoll, 1994, as cited in Razak & Asmawi, 2004). When learners are in charge and control of their learning (Gilbert, 1989, as cited in Callele, 2008), they take responsibility for the quality of it as well. The classroom becomes learner-oriented and the student understands which problems to solve and in what order; as a result, the learners will understand clearly why the information and skills should be learned (Driscoll, 2000, as cited in Callele, 2008). In learning writing, “the solving of one problem often reveals another, and the writing process continues” until the learners become satisfied with their learning. Although the teachers have some specific learning objectives in their minds, they should work as a tour-guide, as suggested by Driscoll (2000), “just pointing out the sights that must be seen” and facilitating the learning process (as cited in Callele, 2008, p. 10).

Therefore, it is believed that reflective learning is a crucial factor in learner academic development; thus, there has been a growing interest in this issue in the field of English language teaching (ELT) (Brockbank & McGill, 1998; Cohen, Ritter, & Haynes, 2007; Moon, 1999). Reflective learning enables learners to explore inwards not only within themselves but also within the courses they have been offered to “access information and inspiration about their efforts in the classroom” (Murphy, 2001, p. 499). It also helps them “move from a level where they may be guided largely by impulse, intuition or routine, to a level where their actions are guided by reflection and critical thinking” (Farrell, 1998, p. 10). Therefore, “reflective learning enables EFL learners to think critically about their successes and failures in a realistic context.” (El-Gharmy, 2015). Without reflection, a check, or review, the same mistakes are often repeated in their later compositions (Howrey & Tanner, 2008).

According to Zubizarreta (2009), learners need to learn from their previous writing experiences through reflection and development of critical writing skills. On the other hand, as Zubizarreta argued, “Reflection does not come naturally or easily to many students. … Students must have opportunities to practice reflection in class and must have generous encouragement on their work in progress.” (p. 172). Consequently, the effect of reflection on writing performance has been explored by some researchers using some techniques such as reflective learning portfolios (see e.g., Hemmati & Soltanpour, 2012; Sabooni & Salehi, 2015) as well as journal writing (see e.g., Alijani & Barjesteh, 2016; Duppenthaler, 2004; Lew & Schmidt, 2011; Minjong, 1997; Razak & Asmawi, 2004; Salem, 2007; Yoshihara, 2008).

In this study, in light of the above-mentioned theories and issues, the researchers decided to explore the effect of keeping reflective notes (RNs) in addition to the effect of
the feedback provided by MS Word. As a result, the following questions were addressed.

1. Does the feedback provided by the MS Word improve the students’ grammatical accuracy in writing?
2. Does the feedback provided by the MS Word improve the Mechanics factor in students’ essays?
3. Is there any significant difference between the ‘CALL with RNs’ and ‘CALL without RNs’ in improving the students’ grammatical accuracy in writing?
4. Is there any significant difference between the ‘CALL with RNs’ and ‘CALL without RNs’ in improving the Mechanics factor in students’ essays?
5. What are the attitudes of the participants in each group towards the treatments they received?

2. Method

2.1. Participants
The participants were 50 (37 women and 13 men) Iranian B.A. English Translation majors who took Academic Essay Writing (Advanced Writing) course in the semester this study was conducted. There were 19 women and six men in the “CALL with RNs” group, and 18 women and seven men in the “CALL without RNs” group. They all had already passed Grammar courses (8 units) and Paragraph Development (2 units).

2.2. Sampling Procedures
The study was a pretest-posttest as well as a comparison-group one. Two intact classes were used, but the classes were randomly assigned to each treatment. Thus, there were two independent variables called ‘CALL with RNs’ and ‘CALL without RNs’ as well as two dependent variables named grammatical accuracy and accuracy in Mechanics in writing.

2.3. Instrumentation
To ensure that the individual participants and the groups were homogeneous, the Straightforward Quick Placement & Diagnostic Test was used. The test had 50 questions, each worth one point. The first 40 were grammar questions and the final 10 were vocabulary questions. To decide on the students’ level, a conversion chart was provided with the bandings as a guide. Most of the students’ scores in the researcher’s two classes ranged from 33 to 39 (i.e., the intermediate level, based on the guide). The very few students whose scores were below 33 or above 39 were in the class during the
term and received the treatments, but their performances were not considered in this study as the participants. In addition to the Straightforward Quick Placement & Diagnostic Test, the participants’ pretest scores in both groups were also compared with each other to ensure that the two groups were homogeneous.

Moreover, two samples of the Task 2 of the General Writing Module of IELTS were used for both pretest and posttest. The task types in both tests were opinion-led, i.e., they required the participants to indicate whether they agree or disagree with the mentioned opinion; they also needed to present their own opinions and provide reasons and examples. It should be noted that the participants had passed Grammar and Paragraph Writing courses in the previous semesters; the participating students had not passed the Academic Essay Writing course and they had just taken it, so they were not expected to know the correct format of such essay; however, the main focus of the study was just on the grammatical accuracy as well as the accuracy in the mechanics factor, and thus, the other factors in writing an essay were not considered in determining the results of this study.

To evaluate the learners’ essays, a detailed analytic scoring rubric for foreign language writing which was provided by East (2009) was used. It consisted of five main items:

1) Cohesion, coherence and rhetorical organization,
2) Knowledge of lexis, idiomatic expressions,
3) Grammatical competence: syntax, sentence-grammar semantics,
4) ‘Mechanics’: spelling and punctuation, and
5) Knowledge of register and varieties of language; knowledge of cultural references (where appropriate).

The scores in the rubric ranged from 0–7. As it was already mentioned, in this study, the participants’ performance and achievements in the two items of Grammatical Competence and Mechanics were considered.

Furthermore, to collect qualitative data, the participating students were asked to take part in a semi-structured interview. Then, their opinions were qualitatively analysed.

2.4. Procedure
At the outset, the teacher-researcher explained this study’s purpose, requirements and procedures to the students in each class. In addition, the students were assured that their anonymity would be maintained. The pretest lasted about 45 minutes of a session and the rest of the time, the teacher-researcher, using a laptop and projector, explained and indicated what the students had to do at home after every session of the class and
during the semester. The teacher indicated how the students needed to use the word processor tips and help to check and correct their grammatical, spelling and punctuation errors. Treatments lasted 10 sessions and then the posttest was done.

Two points should be noted: (1) No student took the tests twice, and (2) The topics of essays were different in pretests and posttests. However, the topics were the same for both experimental groups. Intact classes were used and the classes were arbitrarily assigned to one treatment or another. This semi-randomization procedure was mentioned by Mackey and Gass (2005) as a possibility for dealing with non-randomization of individuals when intact classes are used.

In order to prevent the possibility of the researcher’s bias and considering the rater reliability, each essay was scored independently by both researchers. The final score consisted of the average score of the two raters.

Every session, the teacher taught a part the book, entitled, Longman Academic Writing Series 4: Essay (5th edition) written by Oshima and Hogue (2014). This book was used as the course source during the semester. Both groups were taught by one of the researchers of the study. The students were first required to press the F7 key to disable the program’s checking option and then type their essays relevant to the lesson subject. After the essays were typed, the students released the key and thus had the opportunity to get feedback from program. They were required to notice the underlined sections in green and red color, check the correct form offered by the word processor and check the Help, provided by the software.

The participants in the ‘CALL with RNs’ group were also required to keep a notebook and take notes on their problems and the corrected forms indicated by the word processor as well as any necessary further detailed explanation relevant to their errors. Every time they obtained feedback provided by the software, they needed to review their notes to check whether or not they had already made the same errors.

Every class session, besides the instructor’s teaching a part of the book, the students in both groups had sufficient time to consult the teacher-researcher about the problems they faced during the week while working on MS Word and the feedback they received from it. The participants in “CALL with RNs” group were required to add teacher’s feedback to their notes, too. In order to ensure that the participants had been following the required procedures, the individuals in both groups had to take snapshot of their working page on MS Word and showed it to the teacher-researcher.

It should be noted that the teacher-researcher also carefully corrected the students’ essays, but she provided feedback only on Cohesion, coherence and rhetorical organization as well as vocabulary and idiomatic expressions as just a form of focused feedback. Consequently, according to Ellis (2009), the students’ burden of attending to
all types of their errors simultaneously would be reduced and thus, they would be able to reflect much on each of their errors and as a result the feedback might be more effective.

Another important point, which was considered, was that the “CALL with RNs” group spent more time reflecting and taking notes on their errors; this extra time, which was spent, could affect the reliability of the findings; therefore, to compensate for it, the teacher-researcher suggested that the participants in the “CALL without RNs” refer to Part 3 of the book, covering the essential grammatical structures and relevant exercises, so the students could access metalinguistic clues for extra metalinguistic feedback.

Finally, in the session after finishing the treatment period, the semi-interview was held. To encourage the participants to express themselves fully and very well, they were told they could speak in Persian or English as they preferred. No student spoke in English. Thus, their statements were translated into English by the researchers. The participants were individually interviewed and the interviews were audio-recorded. First, they were asked a question about their attitudes about the effectiveness of the technique used as feedback. Depending on their responses, they were then asked to elaborate on their reasons. According to Berg’s (2007) guidelines, the interviews were then transcribed and content analysed in detail.

3. Data Analyses and Results

3.1. The Normality Tests
In the ‘CALL without RNs’ group, the following tests did not enjoy normal distribution: pretests of both grammatical accuracy and accuracy in the Mechanics factor. In the ‘CALL with RNs’ group, only the pretest of accuracy in the Mechanics factor was the only non-normal data. Their outcomes were not within the ranges of +/-1.96 (Field, 2013). Consequently, the tests that enjoyed normal distribution were analyzed through the parametric independent t-test, but the other tests which did not meet the assumption of normality were analyzed through the non-parametric test of Mann-Whitney U test (Field, 2013; Pallant, 2013).

3.2. Placement Tests
An independent t-test was run to compare the mean scores of both groups on the placement test in order to find out whether or not the two groups were homogeneous.
Table 1: Descriptive statistics placement test by groups

<table>
<thead>
<tr>
<th>Group</th>
<th>N</th>
<th>Mean</th>
<th>SD</th>
<th>SEM</th>
</tr>
</thead>
<tbody>
<tr>
<td>CALL with RNs</td>
<td>25</td>
<td>35.64</td>
<td>1.95</td>
<td>.39</td>
</tr>
<tr>
<td>CALL without RNs</td>
<td>25</td>
<td>35.60</td>
<td>1.97</td>
<td>.39</td>
</tr>
</tbody>
</table>

Table 2: Independent t-test placement test by groups

<table>
<thead>
<tr>
<th>Levene’s Test for t-test for Equality of Means</th>
<th>Equal variances assumed</th>
<th>Equal variances not assumed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Equality of Variances</td>
<td>F</td>
<td>.00</td>
</tr>
<tr>
<td>Mean Difference</td>
<td>Std. Error Difference</td>
<td>Std. Error Difference</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Equal variances assumed</td>
<td></td>
<td></td>
</tr>
<tr>
<td>F</td>
<td>.00</td>
<td>Sig.</td>
</tr>
<tr>
<td>df</td>
<td>48</td>
<td>Sig. (2-tailed)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mean Difference</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Std. Error Difference</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Confidence Interval</td>
</tr>
<tr>
<td>Equal variances not assumed</td>
<td></td>
<td></td>
</tr>
<tr>
<td>F</td>
<td></td>
<td></td>
</tr>
<tr>
<td>df</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>.0747.99</td>
<td></td>
</tr>
</tbody>
</table>

As Table 2 indicates, the assumption of homogeneity of variances was not met (Levene’s F=.00, p=.00<.05). That is why the second row of Table 2 “Equal variances not assumed” is reported (Pallant, 2013). Based on the results displayed in Tables 1 and 2, it could be concluded that the difference between the two groups was not significant (t(47.99)=.07, p<.05); the p-value for this t was .94 (Sig(2-tailed)=.94>.05); however, it represented a very small-sized effect (r=.01) based on Cohen (1988, as cited in Pallent, 2013).

3.3. Pretest of Grammatical Accuracy

A Mann-Whitney U test was run to compare the CALL with RNs and CALL without RNs groups on the pretest of grammatical accuracy in writing in order to prove that both groups were homogeneous regarding the level of grammatical accuracy prior to the administration of the treatments. The test revealed no significant difference in the grammatical accuracy of CALL with RNs group (Md=30.00) and CALL without RNs (Md=30.50), U=283.50, z=−.62, p=.53, r=−.087.

3.4. Pretest of Accuracy in the Mechanics Factor

A Mann-Whitney U test was run to compare the CALL with RNs and CALL without RNs groups on the pretest of accuracy in the Mechanics factor in order to prove that both groups were homogeneous with regard to the accuracy in the mechanics factor of the participants’ written texts prior to the administration of the treatments. The test indicated that there was no significant difference between the two groups. (Md for both groups=3.50), U=298.50, z=−.29, p=.77, r=−.041.
3.5. Inter-rater Reliability Indices

The Cronbach alpha indices were calculated as inter-rater reliability coefficients.

<table>
<thead>
<tr>
<th>Tests</th>
<th>Indices</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pretest of Grammatical Accuracy</td>
<td>.83</td>
</tr>
<tr>
<td>Pretest of Accuracy in the Mechanics Factor</td>
<td>.84</td>
</tr>
<tr>
<td>Posttest of Grammatical Accuracy</td>
<td>.95</td>
</tr>
<tr>
<td>Posttest of Accuracy in the Mechanics Factor</td>
<td>.98</td>
</tr>
<tr>
<td>Delayed Posttest of Grammatical Accuracy</td>
<td>.94</td>
</tr>
<tr>
<td>Delayed Posttest of Accuracy in the Mechanics Factor</td>
<td>.98</td>
</tr>
</tbody>
</table>

3.6. The first Research Question

A paired-sample t-test was conducted to evaluate the impact of CALL without RNs on students’ pretest and posttest scores on grammatical accuracy of their written texts. There was a statistically significant increase in the above-mentioned scores from the pretest ($M=3.44$, $SD=.46$) to posttest ($M=4.22$, $SD=.84$), $t(24)=-8.11$, $p=.000<.05$ (two-tailed). The mean increase was .78 with a 95% confidence interval ranging from -.97 to -.58. The eta squared statistic (.73) indicated a large effect size as proposed by Cohen (1988, cited from Pallant, 2013).

Another paired-sample t-test was conducted to evaluate the impact of CALL without RNs on students’ pretest and delayed posttest scores on grammatical accuracy of their written texts. There was a statistically significant increase in the above-mentioned scores from the pretest ($M=3.44$, $SD=.46$) to delayed posttest ($M=4.30$, $SD=.80$), $t(24)=-8.10$, $p=.000<.05$ (two-tailed). The mean increase was .86 with a 95% confidence interval ranging from -1.07 to -.64. The eta squared statistic (.73) indicated a large effect size as proposed by Cohen (1988, cited from Pallant, 2013).

3.7. The Second Research Question

A paired-sample t-test was conducted to evaluate the impact of CALL without RNs on students’ pretest and posttest scores on the accuracy in the Mechanics factor of their written texts. There was a statistically significant increase in the above-mentioned scores from the pretest ($M=3.46$, $SD=.45$) to posttest ($M=4.52$, $SD=1.07$), $t(24)=-7.29$, $p=.000<.05$ (two-tailed). The mean increase was 1.06 with a 95% confidence interval ranging from -1.35 to -.76. The eta squared statistic (.68) indicated a large effect size as proposed by Cohen (1988, cited from Pallant, 2013).

Another paired-sample t-test was conducted to evaluate the impact of CALL without RNs on students’ pretest and posttest scores on the accuracy in the Mechanics factor of their written texts.
factor of their written texts. There was a statistically significant increase in the above-
mentioned scores from the pretest (M=3.46, SD=.45) to delayed posttest (M=4.64,
SD=1.06), t(24)=−8.19, p=.000<.05 (two-tailed). The mean increase was 1.18 with a 95%
confidence interval ranging from -1.47 to -.88. The eta squared statistic (.73) indicated a
large effect size as proposed by Cohen (1988, cited from Pallant, 2013).

3.8. The Third Research Question
An independent t-test was run to compare the mean scores of the CALL with RNS and
CALL without RNSs groups on the posttest of grammatical accuracy in order to probe
the effect of the two types of treatments on the improvement of the grammatical
accuracy of the students’ written texts after the administration of the treatments.

Table 4: Descriptive statistics posttest of grammatical accuracy by groups

<table>
<thead>
<tr>
<th>Group</th>
<th>N</th>
<th>Mean</th>
<th>SD</th>
<th>SEM</th>
</tr>
</thead>
<tbody>
<tr>
<td>CALL with RNS</td>
<td>25</td>
<td>5.32</td>
<td>.55</td>
<td>.11</td>
</tr>
<tr>
<td>CALL without RNS</td>
<td>25</td>
<td>4.22</td>
<td>.84</td>
<td>.16</td>
</tr>
</tbody>
</table>

Table 5: Independent t-test posttest of grammatical accuracy by groups

<table>
<thead>
<tr>
<th>Levene's Test for Equality of Variances</th>
<th>t-test for Equality of Means</th>
</tr>
</thead>
<tbody>
<tr>
<td>F</td>
<td>Sig.</td>
</tr>
<tr>
<td>Equal variances assumed</td>
<td>13.13</td>
</tr>
<tr>
<td>Equal variances not assumed</td>
<td>5.4441.60</td>
</tr>
</tbody>
</table>

As Table 5 shows, the assumption of homogeneity of variances was not met (Levene’s
F=13.13, p=.00<.05). That is why the second row of Table 5 “Equal variances not
assumed” is reported (Pallant, 2013). Based on the results displayed in Tables 4 and 5, it
could be concluded that the difference between the two groups was significant
(t(41.60)=5.44, p<.05); the p-value for this t was .000 (Sig(2-tailed)=.000<.05); it
represented an above medium effect (r=.64) based on Cohen (1988, as cited in Pallent,
2013).

Another independent t-test was run to compare the mean scores of the CALL
with RNSs and CALL without RNSs groups on the delayed posttest of grammatical
accuracy.
Table 6: Descriptive statistics delayed posttest of grammatical accuracy by groups

<table>
<thead>
<tr>
<th>Group</th>
<th>N</th>
<th>Mean</th>
<th>SD</th>
<th>SEM</th>
</tr>
</thead>
<tbody>
<tr>
<td>CALL with RNs</td>
<td>25</td>
<td>5.36</td>
<td>.55</td>
<td>.11</td>
</tr>
<tr>
<td>CALL without RNs</td>
<td>25</td>
<td>4.24</td>
<td>.84</td>
<td>.16</td>
</tr>
</tbody>
</table>

Table 7: Independent t-test delayed posttest of grammatical accuracy by groups

<table>
<thead>
<tr>
<th>Levene's Test for Equality of Variances</th>
<th>t-test for Equality of Means</th>
</tr>
</thead>
<tbody>
<tr>
<td>F</td>
<td>Sig.</td>
</tr>
<tr>
<td>----</td>
<td>------</td>
</tr>
<tr>
<td>Equal variances assumed</td>
<td>13.76</td>
</tr>
<tr>
<td>Equal variances not assumed</td>
<td>5.56</td>
</tr>
</tbody>
</table>

As Table 7 shows, the assumption of homogeneity of variances was not met (Levene’s F=13.76, p=.00<.05). That is why the second row of Table 7 “Equal variances not assumed” is reported (Pallant, 2013). Based on the results displayed in Tables 6 and 7, it could be concluded that the difference between the two groups was significant (t(41.29)=5.56, p<.05); the p-value for this t was .000 (Sig(2-tailed)=.000<.05); it represented an above medium effect (r=.65) based on Cohen (1988, as cited in Pallent, 2013).

3.9. The Fourth Research Questions

An independent t-test was run to compare the mean scores of the CALL with RNs and CALL without RNs groups on the posttest of accuracy in the Mechanics factor in order to probe the effect of the two types of treatments on the improvement of the Mechanics factor of the students’ written texts after the administration of the treatments.

Table 8: Descriptive statistics posttest of accuracy in Mechanics by groups

<table>
<thead>
<tr>
<th>Group</th>
<th>N</th>
<th>Mean</th>
<th>SD</th>
<th>SEM</th>
</tr>
</thead>
<tbody>
<tr>
<td>CALL with RNs</td>
<td>25</td>
<td>5.64</td>
<td>.65</td>
<td>.13</td>
</tr>
<tr>
<td>CALL without RNs</td>
<td>25</td>
<td>4.52</td>
<td>1.07</td>
<td>.21</td>
</tr>
</tbody>
</table>
As Table 9 indicates, the assumption of homogeneity of variances was not met (Levene’s F=4.38, p=.04<.05). That is why the second row of Table 9 “Equal variances not assumed” is reported (Pallant, 2013). Based on the results displayed in Tables 8 and 9, it could be concluded that the difference between the two groups was significant (t(39.61)=4.45, p<.05); the p-value for this t was .000 (Sig(2-tailed)=.000<.05); it represented a medium effect (r=.57) based on Cohen (1988, as cited in Pallent, 2013).

Another independent t-test was run to compare the mean scores of the CALL with RNs and CALL without RNs groups on the delayed posttest of accuracy in the Mechanics factor.

As Table 10 indicates, the assumption of homogeneity of variances was met (Levene’s F=3.68, p=.06>.05). That is why the first row of Table 11 “Equal variances assumed” is reported (Pallant, 2013). Based on the results displayed in Tables 10 and 11, it could be
concluded that the difference between the two groups was significant ($t(48)=4.56$, $p<.05$); the $p$-value for this $t$ was $.000$ (Sig(2-tailed)$=.000<.05$); it represented a medium effect ($r=.55$) based on Cohen (1988, as cited in Pallent, 2013).

3.10. The Fifth Research Question
What follows represent the participating students’ responses to the interview questions. Attempts are made to provide relevant quotes from their responses to convey key themes and what they said about their perspectives and reasons.

Generally, the participants in both groups had positive attitudes towards receiving feedback by MS Word. The following are some samples of the participants’ statements:

- “In my opinion, typing essay on MS Word is helpful. I think my essay writing will improve greatly when it is assisted by word processor.”
- “Now I’m interested in typing my essay on MS Word. I like it when I correct my mistakes and I know I can get a better mark.”
- “I am more relaxed when MS Word helps me find my mistakes.”
- “I feel less stressed because I know I will have fewer mistakes.”
- “I am more motivated to write extra essays at home and check my mistakes using MS Word.”
- “I think writing an essay with the help of MS Word is easier and more interesting than writing it on a paper.”
- “My grammar knowledge was not or I can say is not good, so writing has always been stressful to me because I knew I would have lots of grammatical mistakes and thus I would lose mark; I like the feedback provided by the word processor because I liked it when I had time to correct my mistakes before the teacher corrected it for me. I liked reading the tips when I made a mistake and the word underlined it. I also felt more relaxed.”
- “I love to be tidy, so I’m used to typing my essays and I was familiar with word processor and its underlining the texts. I already used the MS Word checking system, but this time, during the semester, I was more careful. In addition, I didn’t use to read the tips provided by the word-processor; I only corrected my wrong sentence according to the suggested correct form, but I understood I could use the tips and it helped me. When I read them, the grammatical structure was reminded to me. It was helpful.”
- “I have always liked typing my essays but because my essays would be cleaner and tidier; I used to use the Word processor’s different fonts and margins, but I didn’t use to pay attention to capitalization, punctuation, and the grammatical structures given by the
word-processor. I’m going to use the Word program’s feedback forever. My writing skill will improve.”

As it is evident in the participants’ statements, they had positive attitudes towards using the word-processor mainly because they felt less stressed and more relaxed when they received immediate feedback. In addition, they believed that receiving such computer-assisted feedback made them more interested in writing essay and thus helped them improve their writing skill.

In spite of the above-mentioned information, few participants indicated that using the word-processor’s feedback was not sufficient for them and they needed their teacher’s help, too. Two samples are mentioned below.

- “I really liked this experience, but I think I also need my teacher’s help. Sometimes I couldn’t understand why the MS Word underlined a part of my writing.”

- “It was good, but I felt more relaxed when I knew I would have the opportunity to ask my teacher the parts that I hadn’t understood on my computer.”

Very few students said that the assignment was time-consuming for them, especially because of their own weaknesses in typing skill or working on computer; nevertheless, they reported that the experience was beneficial for them.

- “I’m not good at typing the texts quickly and it was time-consuming for me, but it was so good that I could use the Word program’s feedback to check my errors.”

- “At first, I didn’t like this work because I can’t work on computer very well, but little by little, I became interested in it and using word-processor feedback was useful to me.”

- “I had problem with typing the texts because I was so slow and I could use only two fingers to type. It took time longer than writing using pen and paper, but honestly, the feedback was really good and helpful.”

Considering the RNs in the “CALL with RNs” group, the participants also had positive attitude towards it.

- “I loved keeping RNs; they helped me review my past mistakes and reading the explanations.”

- “RNs were helpful to me because I constantly remembered what I needed to notice while writing.”

- “I loved RNs because I love taking notes. I always take notes when I study my lessons and then review my notes during the tests.”

- “Using Word program and RNs, the time I spent at home writing my essays was like a class; I really liked it because I could have two teachers reminding me my mistakes and explaining the important necessary points to me.”
“Before being accepted into university, I used to go to English institutes. I wrote essays there and my teachers always corrected my mistakes and gave me marks, but I had a new interesting experience this term at university by MS Word and RNs. While I was writing an essay at home, I felt I was studying a lesson carefully. Both techniques were really useful. I will continue doing them.”

“Word software helped me understand that I had made a mistake, then showed me the tips; then I noted down my mistake and the explanations, and I also reviewed my notes whenever I was writing an essay. All these made me be more careful and they really helped me.”

On the other hand, there were a few students who didn’t like RNs technique, not because they found it useless, but mainly because of some personal preferences which could stem from the individual differences. The following include some samples:

“To tell you the truth, I didn’t like RNs because it was so boring to me. I dislike writing, but I can’t say RNs is not helpful.”

“I think RNs can be useful, but I think I don’t need this technique. I think it was enough for me to read the Word processor tips. I could remember them. Writing about my mistakes in RNs was time-consuming and not necessary.”

“RNs were good and useful, but I really preferred to keep audio reflective notes not a written one. I liked it more if I were allowed to record my voice while explaining the points and saved the tracks as RNs. I think if I had done it, I would have benefitted greatly.”

4. Discussion and Conclusion

In this study, the effect of corrective feedback provided by MS word processor on the writing accuracy of the L2 learners was investigated. The participants in both groups, the one with the technique of keeping reflective notes and the one without it, showed progress in their writing accuracy; however, the students who were required to take reflective notes together with receiving the feedback of the program’s checking system outperformed the other group. What follows is a summary of the probable effective factors contributing to the findings of this study.

A factor to be considered is the writing anxiety. Cheng (2004) categorized the sources of writing anxiety into: instructional practices, personal beliefs about writing and learning to write, low self-confidence, and the threats of interpersonal evaluation. Tsai (2008, as cited in Tsao, Tseng, & Wang, 2017) also argued that writing anxiety is determined by factors such as linguistic ability, socio-psychological phenomena and
cognitive factors, as well as classroom practices. Some scholars (e.g., Cheng, Horwitz, & Schallert, 1999; Faigley, Daly, & Witte, 1981) found a significant negative correlation between language anxiety and writing performance. On the other hand, other researchers (e.g., Pajares & Johnson, 1994), in spite of the acknowledging the existence of a correlation between anxiety and learners’ writing, thought that writing anxiety does not directly affect students’ writing performance; they believed that anxiety itself is not an independent variable and that some other individual factors, such as the confidence or self-efficacy while performing tasks, can eventually affect their outcome. Huwari and Hashima (2011) also showed that L2 learners with high writing anxiety had some characteristics, including lack of self-esteem and confidence in improving their writing skills; they usually write papers of low quality (e.g., lacking enough grammatical knowledge and well-developed ideas), and lacking motivation in writing. In Rankin-Brown’s (2006) research, highly anxious students avoided writing, probably because of a fear of teacher and peer evaluation. Such students might consider teacher or peer feedback as criticism and simply paid no attention to it (Goodman & Cirka, 2009). As a result, Choi (2013) recommended language instructors adopt more effective ways to ease the anxiety that learners might feel when learning and writing English in order to support successful language learning experiences. Furthermore, Andrade and Evans (2013) argued that finding ways to effectively self-regulate the learners’ writing motivation and emotion in face of setbacks is critical for both students and teachers. In this study, as some students stated, the applied techniques eased their stress, so the techniques might have some positive effects on the students’ individual factors, such as the ones reported above although each issue needs careful investigation.

The next important factor which could be considered is autonomy. Different scholars have presented different views about autonomy. Little (1990) believed that autonomy is a process. Nunan (1996) claimed that autonomy is not an absolute concept and there are degrees of autonomy. Thus, according to Shinge (2005, p. 32), “autonomy is a construct that is continually changing”. It is possible to teach language learners to be autonomous, but the degree to which they will succeed in becoming autonomous differs, which is determined by the individual differences, such as the personality traits, language learning goals, cultural context, etc. (Shinge, 2005). Regarding the present study, as Kenning (1996) already stated, MS Word could be intrinsically supportive of cognitive and metacognitive autonomy because they encourage their users to consider their compositions critically and try to make improvements.

Moreover, autonomy is often found to be linked with higher learner motivation (Dickinson, 1995; Little, 2007). Although Spratt, Humphreys, and Chan (2002) found that “motivation is a key factor that influences the extent to which learners are ready to
learn autonomously” (p. 245), Dörnyei (2005, p. 79) stated, autonomy results “in a significantly higher level of L2 intrinsic motivation”. Dörnyei and Csizer (1998) had already advised teachers to motivate the learners by promoting their autonomy.

On the other hand, Amrhein and Nassaji (2010) stated that the effectiveness of written corrective feedback depends on learners’ opinions of the feedback and their preferences regarding it. In addition, previous research (Grombczewska, 2011; Rollinson, 2005; Shawish & Abdelraheem, 2010) also suggested that different sources as well as types of feedback may affect learners’ motivation and anxiety. Considering this study and as it was explained earlier in the Introduction, computer-based task (i.e. using MS Word checking system) might support flow because as Egbert (2003) said, flow can occur by tasks under the following conditions:

a) there is a perceived balance of task challenge and participation skills during the task,

b) the task offers opportunities for intense concentration and the participants’ attention is focused on the pursuit of clear task goals, (3) the participants find the task intrinsically interesting or authentic, and (4) the participants perceive a sense of control over the task process and outcomes. (p. 513)

Furthermore, a number of scholars (e.g., Robinson, 1995; Schmidt, 1990, 1995, 2001; Tomlin and Villa 1994) argued that attention plays a crucial role in promoting linguistic processing in learners’ second language development; in their opinions, noticing is a prerequisite for L2 learning to take place. Moreover, according to Schmidt (2010), some second language learners notice more than others and individual differences in noticing ability correlate with rate of learning. For example, being a motivated learner or an unmotivated one can affect the person’s noticing ability as Gardner (1988) stated motivated learners are successful because they are active and pay more attention to the information. As a result, they notice more and try harder to understand the noticed language and thus they achieve higher levels of awareness and learning.

In addition, according to the constructivist view of language learning, students learn in problem-solving environments that challenge their prior knowledge and encourage them to reflect on the differences between their own knowledge and that of the course (Nelson, 2002). According to constructivists, learning always involves both analyzing and transforming new information (O’ Donoghue & Clarke, 2010). Moreover, as Chau and Cheng (2012) stated, “Reflection has been increasingly recognized as integral to effective learning.” (p. 15). Research also suggests that systematic reflection on experiences can produce a variety of benefits, such as understanding one’s behaviors and motivations, as well as increasing one’s confidence and competence (Thorpe, 2004).
In this study, students were active participants in their learning. They were required to monitor their produced output and correct their erroneous statements. The participants in the CALL with RNs were even required to be consciously aware of their learning processes. Thus, based on (Gilbert, 1989, as cited in Callele, 2008), because they were in charge and control of their learning, they took responsibility for the quality of it as well. In conclusion, the findings of this study were in line with the theories of both fields of CALL and the effect of reflection explained in the Introduction. Furthermore, what is found supports the findings of the previous studies done in this regard as reported in the Introduction. In addition, based on the findings of the qualitative data (i.e., the interview responses) in this study and what has already been suggested by the scholars (e.g., Evans, Hartshorn, McCollum, & Woltersberger, 2010; Kormos, 2012), the EFL instructors had better keep the essential factor of individual differences in mind, so they could provide the learners with more effective techniques. Regarding the implication of this study, it should be mentioned that as this study was an action research carried out in a real classroom context and beneficial results have been achieved, the employed techniques in this study are recommended to L2 writing instructors.

About the Authors
Fatemeh Soltanpour holds an MA in TEFL from Payame Noor University, Tehran, Iran. She’s been teaching EFL at Educational Center of Jahad, Karaj, since 2003, and Kar Institute of Higher Education, Qazvin, Iran since 2013. Her areas of interest include psycholinguistics, individual differences in language learning, corrective feedback, learner autonomy, CALL, teacher education and development
Mohammadreza Valizadeh holds an MA in TEFL from University of Tehran, Iran and is now pursuing his PhD at Gazi University, Ankara, Turkey. He has published several articles in different scholarly journals. His areas of interest include psycholinguistics, language learning strategies, learner and teacher autonomy, task-based language teaching and individual differences in language learning.

References


33. El-Ghamry, M, A, H, 2015. Integrating reflective learning into TEFL at the secondary stage in KSA. *Humanising Language Teaching*, 17(1): (Online Article). [www.hltmag.co.uk/feb15/sart03.htm](http://www.hltmag.co.uk/feb15/sart03.htm)


