Investigating the Effect of Metalinguistic Feedback in L2 Pragmatic Instruction

Behzad Barekat
Department of English, University of Guilan, Guilan, Iran
E-mail: behzadbarekat@yahoo.com

Masoumeh Mehri (Corresponding author)
Department of English, University of Guilan, Guilan, Iran
E-mail: Masomeh_mehri200x@yahoo.com

Received: January 11, 2013   Accepted: February 5, 2013   Published: April 25, 2013
doi:10.5296/ijl.v5i2.3032   URL: http://dx.doi.org/10.5296/ijl.v5i2.3032

Abstract
The present study investigates the relative effectiveness of two types of instructions, consciousness-raising instruction (C-R) and consciousness-raising instruction with feedback (C-R F) for teaching English requestive downgrades forms. The study was conducted by 3 intact classes randomly assigned as control group, C-R treatment group, and C-R F treatment group. For this purpose forty five homogenous students took part in this study. They were at intermediate level, and all were male. The result of participants’ performance on pre-test, post-test, and follow-up test indicated that two treatment groups outperformed control group, and C-R F group performed better than C-R group. The results confirmed the benefits of L2 pragmatic instruction in EFL context.

Keywords: Consciousness-raising task, Metalinguistic feedback, Pragmatic competence
1. Introduction

Teaching methodology had shifted toward communicative language teaching in the mid-1970s. A shift of direction paid attention to meaning in use rather than meaning in abstract. Bachman (1990, p. 81) suggests “the ability to use language communicatively involves both knowledge of competence in the language, and the capacity for implementing, or using this competence.”

The main idea underlying above statement refers to the fact that in order to make learners communicatively competent, not only their grammatical competence but also their pragmatic competence need to be fostered.

“Among the different pragmatic aspects, the most dominant area of pragmatics in second language studies is the study of speech acts” (Bardovi-Harling, 2002a; 232). In fact, learners for communication in second language require using different kinds of speech acts such as request, apology, compliment, refusal etc. (Cohen, 2005; Usó-Juan and Martínez-Flor, 2010)

However, According to Kasper and Rose (2001) “in addition to cross-cultural differences, the indexical meaning of speech acts and strategies varies inter and intra culturally” (p, 8). Lack of adequate knowledge in this respect may lead to a breakdown in communication known as “pragmatic failure” which may be considered as impolite and impudent utterance by native speakers.

Research conducted on speech acts in EFL context have shown that both classroom input and pedagogical material are inadequate to result in pragmatics development. (Alcón and Safont, 2001; Usó-Juan, 2007).

Iran as an EFL context is not an exception in this regard. Regarding the speech act of request, the findings of the research by Eslami-Rasekh et al. (2006) confirmed the previous studies on second language speech acts which have demonstrated that even Advanced Iranian language learners used nonnative speech act features on written DCT.

Bardovi-Harlig (2001) indicated that without the benefit of instruction, many L2 pragmatic aspects are not acquired or they are learned more slowly. This suggests that instruction can facilitate second language pragmatic acquisition. Rose (2005, p. 13) argued that, “…pedagogical intervention has at least an important facilitative role, which is especially good news for learners in foreign language contexts”. According to Kasper and Rose (2002), despite a growing literature on L2 pragmatic acquisition ability, the issue of how to teach pragmatic competence still remains.

In this regard, a number of ILP studies have examined the efficacy of certain instructional methods in different contexts. (Takahashi, S., 2001; Tateyama, Y., 1997; Fukuya and Zhang, 2002; Alcón, 2005; Dastjerdi and Farshid, 2011; Takimoto, 2012).

Therefore, the present study was carried out to broaden the scope of studies have been done so far in the area of speech act instruction. This study tried to examine the effects of explicit feedback which rarely has been investigated in the area of L2 pragmatics development.
2. Theoretical & Research Background

According to Schmidt (1990), noticing is necessary for the acquisition of target language feature. He argued that “input does not become intake for language learning unless it is noticed, that is, consciously registered” (Schmidt, 2010, p.1).

Schmidt (1993b) extended the role of consciousness in the learning of pragmatic rules. Based on the studies of L1 and L2 acquisition of pragmatic, Schmidt (1993a) argued that L2 pragmatic learning requires “attention to linguistic forms, functional meanings, and relevant contextual features” (p. 35).

Raising consciousness in the pragmatics has been emphasized by other scholars (Judd, 1999; Kasper and Schmidt, 1996; Rose 1997; Rose and Ng, 2001). Kasper and Schmidt (1996) highlighted the need for pedagogical intervention which draws learners’ attention to both contextual features and its linguistics properties.

Drawing on Schmidt Noticing Hypothesis, Ellis (1994) argued that noticing accounts for the way input are integrated into the learner’s developing system. According to Ellis (2003, p. 163), “a consciousness-raising task consists of: (1) data containing exemplars of the target feature, and (2) instructions requiring the learners to operate on the data in some way”. Therefore Features of consciousness raising task such as isolating specific pragmalinguistic features, providing learners with an explicit rule, describing the feature, and engaging learners’ intellectual efforts, are of great help to make the pragmalinguistic forms and sociopragmatic rules salient enough for ESL and EFL learners.

A number of studies have been conducted to investigate claims about the role of awareness and attention in L2 pragmatic learning. Among them, some recent studies on request speech act have employed input-based approaches in teaching pragmatics.

Takahashi (2001) conducted one of the most complex designs of input enhancement. He investigated four input enhancement conditions in acquisition of biclausal request forms. These four conditions differed from each other in terms of the degree of input enhancement. The results of study showed that explicit group performed better than the other three groups in using request strategies.

Alcón (2005) investigated the benefits of explicit and implicit teaching of L2 request strategies with EFL learners in Spain. The participants of this study were 132 students randomly assigned to three groups, explicit group with direct consciousness raising task activities and written metalinguistic feedback, implicit group with implicit consciousness-raising activities, and control group. The effectiveness of both kinds of instruction was approved in the post-test. However, the explicit instruction gained better results over the implicit one.

Takimoto (2009) assessed the effects of three types of input-based instruction for teaching English request forms to Japanese learners of English: comprehension-based instruction, structured input instruction, and consciousness-raising instruction. Three experimental group and one control group were considered in the study. The experimental groups received four
40-minute sessions of instruction. Pre-tests, post-tests, and follow-up tests consisting of a discourse completion test, a role-play test, a listening test, and an acceptability judgment test revealed that the all three treatment groups performed considerably better than the control group. However, the effect of treatment did not sustain for comprehension-based instruction between the post-test and follow-up test in the listening test.

Takimoto (2012) investigated the effects of metapragmatic discussion on learners’ recognizing and producing English request downgraders. Three groups participated in this study, two treatment group and one control group. The two experimental groups received the following instructional treatments: problem-solving tasks with metapragmatic discussion (PTW) and problem-solving tasks without metapragmatic discussion (PTO). The results demonstrated that the PTW and PTO groups performed significantly better than the control group on the unplanned written-production and the unplanned written-judgment tests. There were no statistically significant differences between the two experimental groups on the unplanned written-judgment test; but a significant difference in the two treatment groups in the unplanned written-production test.

Thus, the present study was carried out to broaden the scope of studies carried out so far in the area of interlanguage pragmatics. In this study, we tried to investigate the relative effectiveness of consciousness-raising task as an input enhancement method and consciousness-raising task along with explicit feedback on Iranian EFL learners’ production of requestive downgrades in English.

Research questions

This study was carried out to answer the following questions:

1) Does consciousness-raising task instruction enhance Iranian EFL learners’ pragmatic awareness?

2) Does consciousness-raising task along with metalinguistic feedback enhance Iranian EFL learners’ pragmatic awareness to a greater degree than consciousness-raising task instruction alone?

3. Method

3.1 Participants

The population of this study is comprised of 45 students from a private language institute in Ghaemshahr called Marefat Language Institution. Based on the institute rules, the students were in intermediate level when we were conducting this study. The age of the learners ranged from 16 to 20, and all were male. Their first language was Persian and none of them has been abroad. The reason for choosing this level was that some part of book content at this level is allocated to different kinds of speech acts such as request, apology, compliment and etc. Therefore students were familiar with the different kinds of speech act in general.

According to Hatch and Lazaraton (1991), we cannot always have randomization of individuals rather we are more dependent on the contexts that already exist such as intact
classes for our research. Due to Institutional constraints, the study had been conducted with 3 intact classes which were selected randomly and the choice of which class to be C-R treatment group, C-R F treatment group or control group were also random.

3.2 Material

The C-R F experimental group received same consciousness raising tasks as C-R group. The only difference was that this group received metalinguistic feedback while performing consciousness-raising task (e.g., do you think the appropriate score should be 4? Or no, you should choose 4 instead of 5 because the speaker used modification devices in her request in order to be much more polite). It is worth mentioning that the metalinguistic feedback was given only if the learners answered questions incorrectly. In addition to providing C-R tasks aims, the metalinguistic feedback reinforces the pragmalinguistic and sociolinguistic connection during the metalinguistic discussions of the target features of speech act of request. For conducting this kind of instruction, the researcher allotted 20 to 25 minutes of each session to it.

3.3 Instruments

Written Discourse Completion Test (WDCT) consists of 20 requesting situation items adopted from Takimoto (2009) was given to the students to measure participants’ requestive speech act performance. He examined the construct validity of the WDCT through factor analysis procedure and the result revealed high construct validity, and the reliability estimates for the WDCT was 9.33.

The participants read 20 short scenarios and responded in English. The entire situations in WDCT were based on three sociolinguistic variables (Power, degree of Imposition, and Distance). These three sociolinguistic variables were selected because they were identified as three sensitive and dependent variables in cross cultural studies which play a crucial role in requestive speech act behavior (Hudson et al., 1995).

According to Takimoto (2009), situations with the high degree of imposition, combined with power and distance, required participants to use downgrades. Therefore, the focus of the present study is on high level of imposition combined with power and distance. The WDCT of the present study included ten high-imposed situation (HI) items combined with power and distance. However, ten low-imposed situation items combined with power and distance are added as distracters. In order to minimize test learning effect or test order effect, three versions (A, B, and C) of WDCT were used. In fact, the order of presentation of the same situations was counterbalance across the pre-test, the post-test, and the delayed posttest.

3.4 Procedure

To evaluate the effects of the treatments on the learners’ realizations of speech act of request, WDCT, as pre-test and post-test and follow-up test, was employed. The Pre-test was given to three groups during the learners’ class period over two class days prior to administration of treatment. After the participants wrote their responses, the pre-tests were collected and transcribed for analysis. A week after the pretests, the instructional treatment started. All
participants in the experimental groups, C-R and C-R F group received the treatment during regular class periods. However, the control group received no instruction on speech act of request but had instructor-led lessons from the textbooks. Both treatments lasted for 6 weeks. The post-test was conducted nine to ten days after the instruction. Finally, the follow-up test was administered four weeks after instruction.

3.5 Data Analysis

The participants’ responses to WDCT tests and the explanation of how to use the rating scale were sent to two English native speakers. The raters could ask their questions about rating Web Chat. They rated the appropriateness of participants’ response to HI request forms on a 5-point scale ranging from 1 to 5.

Those responses that reflected mastery of downgrades received 5 points; on the other hand those responses that were considered completely inappropriate received 1 point. Since there were 10 HI items on the test, the maximum score was 50.

In this research, a 5-point Scale is unlikely to be equidistant and considered to be ordinal since there is no way to ensure that raters view the difference between “1” and “2” the same as they might view the difference between “2” and “3”. Inter-rater correlations from the study yielded an acceptable level of agreement for inter-rater reliability \( r = 0.93 \).

4. Results

In order to identify that whether the differences between groups on WDC pre-tests are statistically significant or not we need to compare the mean scores of 3 groups. One-way between-groups ANOVA was used for this purpose.

Table 1. One-way between-groups ANOVA on the WDCT pretest scores

<table>
<thead>
<tr>
<th>Groups</th>
<th>Sum of Squares</th>
<th>df</th>
<th>Mean Square</th>
<th>F</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between Groups</td>
<td>1.440</td>
<td>2</td>
<td>.720</td>
<td>.096</td>
<td>.908</td>
</tr>
<tr>
<td>Within Groups</td>
<td>284.121</td>
<td>38</td>
<td>7.477</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>285.561</td>
<td>40</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

As it is obvious from the table 1, all three groups had similar and approximately equal mean on WDC pre-tests. As table 1 shows the significant value is .90, which is greater than .05, indicating no statistically significant differences between the 3 groups for the pre-test scores. In other words, the three groups did not differ in their pre-test scores on the speech act of request test prior to the treatment.

After ten sessions of treatment, mixed between-within subjects ANOVA was performed in order to compare participants’ scores on the WDC test in three groups, Control, C-R and C-R
F group, across the three time junctions (pre-test, post-test, and follow-up).

The results of mixed between-within subjects ANOVA performed on the high imposed situation of WDCT showed a significant between-subjects effect for instructions (the C G, C-R G, and C-R F G), F (2, 28) = 122.93, p < .05. It also revealed a significant within-subjects effect for time (the pre-test, post-test, and follow-up test), F (2, 28) = 328.40, p < .05. Moreover, a significant interaction between time and instruction was found F (4, 56) = 139.76, p < .05. The following figure clearly shows the interaction.

**Figure 1. Interaction plot for WDCT**

Note: CG = Control Group, C-R G = Consciousness-raising Group; C-R F-G = Consciousness-raising instruction with feedback Group.

Bonferroni post hoc test was performed for time comparisons. Table 2 demonstrates the result of post hoc test:

**Table 2. Pairwise Comparisons of time**

<table>
<thead>
<tr>
<th>(I) Time</th>
<th>(J) Time</th>
<th>Mean Difference (I-J)</th>
<th>Std. Error</th>
<th>Sig.</th>
<th>95% Confidence Interval for Difference</th>
<th>95% Confidence Interval for Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>(I-J)</td>
<td></td>
<td></td>
<td>Lower Bound</td>
<td>Upper Bound</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>-11.705</td>
<td>.507</td>
<td>.000</td>
<td>-12.732</td>
<td>-10.678</td>
</tr>
<tr>
<td>1</td>
<td>3</td>
<td>-11.484*</td>
<td>.490</td>
<td>.000</td>
<td>-12.475</td>
<td>-10.492</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>11.705</td>
<td>.507</td>
<td>.000</td>
<td>10.678</td>
<td>12.732</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>.222</td>
<td>.111</td>
<td>.053</td>
<td>-.003</td>
<td>.446</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>11.484</td>
<td>.490</td>
<td>.000</td>
<td>10.492</td>
<td>12.475</td>
</tr>
<tr>
<td>3</td>
<td>2</td>
<td>-.222</td>
<td>.111</td>
<td>.053</td>
<td>-.446</td>
<td>.003</td>
</tr>
</tbody>
</table>

Based on estimated marginal means

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

a. Adjustment for multiple comparisons: Bonferroni.
Although table 2 shows that all pairs of time differ significantly from each other (p < 0.05), the amount of the mean difference vary for each of pairs. According to the table, the mean difference between time1 and time 2 (time2-tim1) is 11.70. It means that participants’ scores increase sharply from pre-test to posttest. However, this mean difference between time2 and time3 is not significant (p>.05) and the mean difference of -.22 shows that participant’ scores on WDC test decrease a little from the posttest to the follow-up test.

However, the result of tests of within-subjects effect indicated that time interacted in some way with the type of the group. It means there is not the same change in mean scores over time for the three different groups. In fact, the interaction shows a relatively large superiority on WDCT scores from time1 to time 2 for the two treatment groups over the control group.

Table 3. group * Time

<table>
<thead>
<tr>
<th>group * Time</th>
<th>Time</th>
<th>Mean</th>
<th>Std. Error</th>
<th>95% Confidence Interval</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Lower Bound</td>
</tr>
<tr>
<td>C G</td>
<td>1</td>
<td>23.929</td>
<td>.731</td>
<td>22.449</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>24.143</td>
<td>.973</td>
<td>22.174</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>24.214</td>
<td>.900</td>
<td>22.393</td>
</tr>
<tr>
<td>C-R G</td>
<td>1</td>
<td>23.846</td>
<td>.758</td>
<td>22.311</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>38.462</td>
<td>1.009</td>
<td>36.418</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>38.154</td>
<td>.934</td>
<td>36.271</td>
</tr>
<tr>
<td>C-R F G</td>
<td>1</td>
<td>23.500</td>
<td>.731</td>
<td>22.021</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>43.786</td>
<td>.973</td>
<td>41.817</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>43.357</td>
<td>.896</td>
<td>41.543</td>
</tr>
</tbody>
</table>

The result reveals a significant effect of group, F (2, 28) =122.93, p<.05. It means that there is significant difference in WDCT scores for the three groups. In order to recognize where the differences between groups occur, we conducted post-hoc test.
Table 4. Multiple Comparisons

<table>
<thead>
<tr>
<th>(I) group</th>
<th>(J) group</th>
<th>Mean Difference (I-J)</th>
<th>Std. Error</th>
<th>Sig.</th>
<th>95% Confidence Interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gabriel</td>
<td>C G</td>
<td>-9.392</td>
<td>1.112</td>
<td>.000</td>
<td>-11.643 to -7.141</td>
</tr>
<tr>
<td></td>
<td>C-R G</td>
<td>-12.786*</td>
<td>1.091</td>
<td>.000</td>
<td>-14.995 to -10.576</td>
</tr>
<tr>
<td></td>
<td>C-RF G</td>
<td>-3.394</td>
<td>1.112</td>
<td>.004</td>
<td>-5.645 to -1.142</td>
</tr>
<tr>
<td>C-R G</td>
<td>C G</td>
<td>9.392</td>
<td>1.112</td>
<td>.000</td>
<td>7.141 to 11.643</td>
</tr>
<tr>
<td></td>
<td>C-RF G</td>
<td>3.394</td>
<td>1.112</td>
<td>.004</td>
<td>1.142 to 5.645</td>
</tr>
<tr>
<td>C-R F G</td>
<td>C G</td>
<td>12.786</td>
<td>1.091</td>
<td>.000</td>
<td>10.576 to 14.995</td>
</tr>
<tr>
<td></td>
<td>C-R G</td>
<td>3.394</td>
<td>1.112</td>
<td>.004</td>
<td>1.142 to 5.645</td>
</tr>
</tbody>
</table>

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

The study results reveal a significant improvement of both treatment groups. However, compared with control group, the participants in the experimental C-R F group generated significantly higher scores on the WDCT than C-R group.

5. Conclusion

The aims of this study were to investigate the effects of two types of instructions, C-R treatment and C-R treatment along with feedback, in L2 learners’ pragmatic awareness.

The result of the study demonstrated that both treatment groups outperformed control group on post-tests and follow-up tests. Moreover, the C-R F treatment group outperformed C-R treatment group on WDC posttest; however, the result of paired comparison statistical analysis revealed little mean difference.

The higher effectiveness of C-R F treatment in the pragmatic improvement can be justified by Craik (2002) who argued that claimed that the quality of a memory trace depends on the depth of mental processing where meaning plays a very important role. Meaning, in this case, includes both pragmalinguistic and sociopragmatic meaning. In other words, the participants of C-R F group in this study had the chance of talking meaningfully about the pragmalinguistic and sociopragmatic connections of request forms during C-R activities, allowing greater retention and processing at a deep level by the learners.

The result of the study confirmed the idea of Bardovi-Harlig (2001) and Rose (2005) who believe in the facilitative role of pragmatic intervention in EFL context. In addition, the benefits of both kinds of instructions on learners’ awareness of speech act of request support Schmidt’s noticing hypothesis (1993a, 1995, 2001), in which “noticing” center on initial input recognition and focuses on the significance of attention and consciousness (1993a) in second-language acquisition.

Second, this study provides evidence for Schmidt’s emphasis on raising learners’ consciousness in L2 pragmatic acquisition. Kasper and Schmidt (1996) highlighted the necessity of pedagogical intervention in L2 pragmatic acquisition which requires learners to attend to both contextual features and its linguistic form. Among three way of raising learners’ consciousness which have been offered by them, we employed teaching material (C-R tasks in this study).

To sum up, the present study reveals the applicability of the focused tasks in the realm of
pragmatics. In addition, the result showed that effective learning occurs when the tasks used provide learners with the opportunity for processing both the forms and meanings of the target features. Although this study focused on speech act of request, this teaching technique may be applicable to other speech acts such as apology, suggestion, refusal, etc. Future studies can investigate the effects of different types of instruction along with different types of feedback. The contextual variables used in this study were “power,” “social distance,” and “imposition of the task.” Further studies could be designed to examine other contextual factors, such as age and gender.

References


