Title: The effects of explicit meta-pragmatic instruction on EFL learners' performance of constructive criticisms in an academic setting

Author(s): Nguyen Thi Thuy Minh, Pham Minh Tam and Cao Thuy Hong

Source: T. Greer, D. Tatsuki & C. Roever (Eds.), Pragmatics and language learning volume 13 (pp. 213–244)

Published by: National Foreign Language Resource Center, University of Hawai'i

This document was published under a Creative Commons Attribution license (CCBYNCSA).

Citation: Nguyễn, T. T. M., Pham, M. T., & Cao, T. H. (2013). The effects of explicit meta-pragmatic instruction on EFL learners' performance of constructive criticisms in an academic setting. In T. Greer, D. Tatsuki & C. Roever (Eds.), Pragmatics and language learning volume 13 (pp. 213–244). Honolulu, HI: University of Hawai'i, National Foreign Language Resource Center.
The Effects of Explicit Metapragmatic Instruction on EFL Learners’ Performance of Constructive Criticism in an Academic Setting

Nguyen Thi Thuy Minh
Nanyang Technological University, Singapore

Pham Minh Tam
Vietnam National University, Hanoi, Vietnam

Cao Thuy Hong
Vietnam National University, Hanoi, Vietnam

Previous research has shown that second language learners can experience considerable difficulty when giving constructive criticism to peers in instructional settings (see Nguyen, 2005). The current study examines whether metapragmatic instruction facilitates the learning of this speech act set. The study is conducted in an attempt to address the need for further research on L2 pragmatic instruction (see Rose, 2005) and expand the range of speech acts under investigation, which is currently restricted to a small set of well-defined acts and excludes more complex speech act sets such as constructive criticism. Thirty Vietnamese high-intermediate-level learners of English as a foreign language were recruited for the comparison and treatment conditions. Over a 15 week semester, the treatment group received explicit instruction in constructive criticism in English while the comparison group did not receive any equivalent instruction. As pre- and post-tests, learners participated in oral peer-feedback sessions where they provided constructive criticism on peers’ actual written assignments. The treatment group also completed a post-test discourse completion task for the purpose of triangulation and wrote a one-page
essay to reflect on what they had learned. Findings show positive instructional effects on learners’ pragmatic performance and raise issues about task effects on measuring L2 pragmatic performance.

Introduction

This exploratory study sets out to investigate the effects of explicit metapragmatic instruction on a group of EFL learners’ performance of constructive criticism in institutional settings. More specifically, it examines whether explicit teaching helps to improve learners’ use of pragmatic strategies for delivering constructive criticism and linguistic devices to compensate for the potential face-threat of the criticism. This study is part of a larger ongoing research project which aims to compare the relative effectiveness of implementing explicit and implicit pedagogical interventions in the pragmatic realm. It is conducted to address the need to further research on L2 pragmatics instruction (Jeon & Kaya, 2006; Kasper & Rose, 2002; Rose, 2005; Rose & Kasper, 2001), especially regarding under-represented speech act sets such as giving constructive criticism to peers in the classroom context.

Previous L2 pragmatics research has indicated that non-native speakers can experience considerable difficulty when performing speech acts in the target language (TL) due to their incomplete L2 pragmatic knowledge, limited proficiency in the TL, and also because of the influence of their first language (L1) knowledge of pragmatics (see Ellis, 2008; Kasper & Rose, 2002 for reviews). Uninstructed learners in foreign language (FL) contexts might encounter even greater difficulty compared to those in second language (SL) contexts who are exposed to the TL use to a considerably greater extent (Kasper & Rose, 2002; Rose & Kasper, 2001). Learner difficulty with the pragmatics of another language can cause more problems than their difficulty with grammar when they participate in intercultural communication. This is because pragmatic errors are often perceived by native speakers (NS) as rudeness rather than a lack of competence in the L2 (Boxer, 1993; Thomas, 1983). Instruction at the pragmatic level is therefore important in helping L2 learners to communicate more appropriately in the TL.

Recent years have witnessed an increase in research interest in the instruction of different speech acts or speech act sets, including Olshtain and Cohen (1990) on apologies, Billmyer (1990) and Morrow (1995) on refusals and complaints, Kondo (2001) on refusals, Rose and Ng (2001) on compliments and compliment responses, Fukuya and Clark (2001), Safont (2003), Salazar (2003) and Takahashi (2001, 2005) on requests, and Martínez-Flor and Fukuya (2005) on suggestions. These studies have sought to address three important topics: the teachability of a particular speech act, the benefits of instruction versus exposure, and the relative effectiveness of different teaching approaches (Jeon & Kaya, 2006; Rose, 2005).
Overall, findings regarding the first and second topics have suggested that although certain L2 pragmatic areas remain difficult for learners, L2 pragmatics can be taught and instruction has a beneficial effect on pragmatic development. For example, Olshtain and Cohen (1990) found an increased use of intensifiers and a wider variety of apology strategies for instructed learners. Safont (2003) reported a marked increase in the use of request modifications after the participants received instruction in this area. Billmyer (1990) found that instructed learners generally enjoyed an advantage over their uninstructed peers in using compliments and compliment responses, especially with regard to frequency, spontaneity, appropriateness, and adjectival repertoire. A meta-analysis by Jeon and Kaya (2006) that examined the efficacy of 13 instructed L2 pragmatics studies, many of which involved the teaching of speech acts, reported a medium effect size when comparing the performance of instructed and uninstructed learners, and a large effect size when comparing the performance of the instructed learners before and after interventions, also attesting to the benefits of L2 pragmatics instruction in general.

However, findings regarding the third topic have been more controversial. Although there seems to be evidence to suggest that explicit teaching has more effect than implicit teaching (Rose & Ng, 2001; Takahashi, 2001), Jeon and Kaya (2006) have warned against drawing premature conclusions because of the limited data available and a variety of methodological issues, including unequal treatment lengths for explicit and implicit groups and variation in the data collection methods. To achieve more conclusive research outcomes, this line of research should be continued and methodological issues should be addressed in future studies.

Another limitation of the existing body of research on instructed L2 pragmatics lies in its narrow scope of investigation. Despite the growing interest in this realm, much of the current research focuses on a rather restricted range of relatively ‘well-defined’ speech acts (see above). Thus, it remains to be seen whether instruction works for more complex speech act sets such as constructive criticism, which may cause even more problems for L2 learners and therefore require at least equally careful pedagogical attention (Nguyen, 2005, 2008a, 2008b).

Constructive criticism in general refers to negative assessment of an individual’s work-in-progress with the aim of improving current or future performance (Nguyen, 2005). In this sense, it includes giving critical feedback on task performance in the language classroom. In an empirical study of constructive criticism given to peers by Vietnamese EFL learners and NSs of Australian English in writing feedback sessions, Nguyen (2005) found that criticism is made up of multiple components, none of which is the head act (Blum-Kulka & Olshtain, 1984) and therefore may be better described as a speech act set rather than a single speech act. For example, criticism can be
realized by means of any combination of the following: expression of disapproval, expression of negative evaluation, statement of problem, and advice for change. Because it involves a negative evaluation of the recipient’s (R) work with the view of influencing R’s future action, criticism may pose a threat to both the recipient’s positive face (the desire to be approved or accepted by others) and negative face (the desire to be free from imposition from others) (Brown & Levinson, 1987). Therefore, this speech act set is often mitigated by means of various external and internal modifiers, such as compliments (i.e., positive remarks to preface the criticism), grounders (justifications for the criticism), hedges, and downgraders. (see Appendix 1 for definitions and examples of these categories).

Providing constructive criticism in institutional settings is normatively expected of teachers. However, delivering criticism from one peer to another is often tricky, not only because learners generally lack the knowledge required to give fair criticism, but also because they lack the pragmatic competence to express their criticism in an appropriate manner. Research has shown that while students from some countries may find giving constructive criticism that can improve a colleague’s work to be a positive exercise, students from other cultures (particularly Asian cultures) are often uncomfortable expressing criticism of another student’s work (Nelson & Carson, 1998). Other studies have indicated that learners of English may give constructive criticism very differently from native speakers, which may adversely affect their communication with NS peers. For example, Nguyen (2005, 2008a, 2008b) found that when commenting on peers’ essays, Vietnamese EFL university students tended to soften their criticism less frequently and provided direct criticism more often than their fellow Australian L1 students. The L2 speakers also employed modal verbs such as must, should, and have to inappropriately and thus needed pedagogical assistance to improve their pragmatic competence.

Although to date a great deal of pedagogical effort has been devoted to orienting L2 learners to the content of peer assessment and the structure of peer assessment sessions (Liu & Hansen, 2002; Mendonca & Johnson, 1994; Rollinson, 2005), little attention has been focused on the language used to provide negative assessment. The present study aims to address some of the language problems that L2 speakers can have with constructive criticism and thus constitutes an attempt to expand the range of learning targets under investigation. It focuses on a group of TEFL (Teaching English as a Foreign Language) teacher-students who have had quite limited exposure to English use outside the classroom. It is argued that these students need training in how to give constructive criticism appropriately so that they can successfully transfer the acquired knowledge and skills to their future professional practices. Further, since their only regular source of TL pragmatic input seems to come from their formal English classes, some pragmatics-focused instruction is necessary to help them achieve a higher level of pragmatic competence in the TL.
The study defines explicit teaching as a pedagogical approach that combines metapragmatic generalizations (DeKeyser, 2003), output practice and explicit correction of forms and meanings. Multiple instructional strategies have been employed because it is believed this can produce more positive learning effects than the adoption of a single teaching strategy (Izumi, 2002; Martínez-Flor & Fukuya, 2005). In conceptualizing its pedagogical interventions, the present study draws on Schmidt’s (1993, 1995) noticing hypothesis, Swain’s (1985, 1995, 2005) output hypothesis, and Long’s (1983, 1996) interaction hypothesis, which specify conditions for language learning as opportunities for input noticing, corrective feedback and output. In other words, it is believed that learners can benefit from a type of instruction that allows them to attend to linguistic forms and see the relationship between forms and pragmatic meanings, receive negative evidence about their output and modify their output accordingly.

In designing its teaching materials the present study draws on a resource for teaching constructive criticism developed by Nguyen and Basturkmen (2010). This resource included activities originally devised for a group of young Vietnamese EFL learners in an English for Academic Purposes program. The activities aimed to increase the learners’ sensitivity to cultural issues involved in giving constructive criticism in peer assessment sessions, and to help them find socially appropriate language for performing this speech act set in academic contexts. Observations of the learners’ constructive criticism after the implementation of these activities showed evidence of improvement. Details of the instructional implementation will be described in the section below. To evaluate the effectiveness of this instructional implementation, the present study aims to answer the following two research questions:

1. Do learners benefit from explicit instruction in constructive criticism in English?
2. Do some aspects of giving constructive criticism benefit more from explicit instruction than others?

Methodology

Study design and participants

The present study adopted a quasi-experimental, pretest-posttest design with a comparison group. Two intact EFL classes consisting of a total of thirty students were recruited for the study. At the time of the data collection, the students (1 male and 29 female) were pre-service EFL teachers majoring in their third year of English at a teacher training institution in Hanoi, Vietnam. In order to be admitted to this four-year teacher training program, the students had to sit for a nationwide university entrance exam on English grammar and reading comprehension (alongside math and literature). After the students passed the entrance exam, they took another English placement test which included all four language skills. The top fifteen students were placed on a fast-track program and the remaining
cohort was randomly assigned to different classes in a mainstream program. The students stayed in the same class from Year 1 through to Year 4.

Two classes were randomly selected from among the mainstream classes and assigned to the treatment \((N=12)\) and comparison \((N=18)\) conditions. Although there was no official information on the students’ English proficiency levels, according to the institution, the mainstream third year class was considered to be at the high-intermediate level. There were no notable differences between the two groups in terms of age range, lengths of English study and exposure to English outside the classroom. In both groups, the learners’ ages ranged between 20 and 23 (average 21). Their lengths of English study ranged between 6 and 9 years. None of them had ever visited an English-speaking country. They all had limited exposure to English in their daily lives and little chance to use English for communication outside the classroom.

Treatment was incorporated into a writing syllabus where the students were taught how to write paragraphs and different types of academic essays in English. The writing class met for three class hours every week. Instruction on constructive criticism was implemented for the treatment group for one class hour (i.e., 45 minutes) every week over a 15 week semester between February and May, 2010. The second author taught the comparison group and the third author taught the treatment group. Both instructors were trained EFL teachers and had obtained master’s degrees in English language education from English-medium institutions outside Vietnam. The two classes followed the same writing syllabus and schedule. The only difference was the additional instruction of language for giving constructive criticism for the treatment group. Other than the writing course, the students also took part in oral skills and reading courses as well as courses in English linguistics, English language teaching methodology, and the history and geography of English speaking countries.

**Instructional procedures**

Over the 15 week period, the treatment group received explicit instruction on the target forms, participated in productive activities (e.g., role-play, oral peer-feedback), reflected on their own output, and received explicit correction throughout the process. The comparison group, on the other hand, did not receive any equivalent instruction.

The target forms included two major strategies for delivering constructive criticism: (1) statement of problem and (2) giving advice for improvement, as well as two types of criticism modifiers: (1) external modifiers (compliments, grounders and disarmers) and (2) internal modifiers (past tense, modal verbs, modal adverbs, expression of uncertainty, hedges and downgraders) (see Appendix 1). These strategies and modifiers were selected as instructional foci because they occurred most frequently in NS criticism. They also constituted areas of difficulty for many learners of English (Nguyen, 2005; 2008a). The
teaching activities for the treatment group were adapted from Nguyen and Basturkmen (2010) and are summarized in Table 1.

**Table 1.** Instructional procedure for the treatment group

<table>
<thead>
<tr>
<th>week</th>
<th>focus and activities</th>
<th>description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>pre-test reflection on experience</td>
<td>Learners worked in groups to reflect on and share their experience of giving and receiving constructive criticism in equal-status academic encounters. They also reflected on the similarities and differences between their native language and culture and English in giving constructive criticism.</td>
</tr>
<tr>
<td>2</td>
<td>strategies for giving criticism: awareness-raising activity metapragmatic instruction on criticism strategies</td>
<td>Learners were guided to identify criticizing strategies from authentic NS samples of speech. Explicit instruction on two major strategies for delivering constructive criticism: (1) statement of problem and (2) giving advice for improvement. Distribution of explanatory handouts.</td>
</tr>
<tr>
<td>3</td>
<td>modifying criticism: awareness-raising activity metapragmatic instruction on criticism modifiers</td>
<td>Learners were guided to identify types of modifiers from authentic NS samples of speech. Explicit instruction on two major types of criticism modifiers: (1) external (e.g., compliments and disarmers) and (2) internal (e.g., hedges, downgraders, etc.) Distribution of explanatory handouts.</td>
</tr>
<tr>
<td>4</td>
<td>practice: recognizing degree of softening</td>
<td>Learners identified modifiers from samples and judged the relative degrees of softening of the given criticism, using their knowledge of criticism modifiers.</td>
</tr>
<tr>
<td>5</td>
<td>practice: softening criticism using appropriate modifiers</td>
<td>Learners added modifiers to samples to make given criticism more sociopragmatically appropriate.</td>
</tr>
<tr>
<td>6</td>
<td>practice: role-play</td>
<td>Learners paired up to write and act out on structured role-plays involving giving constructive criticism to peers in classroom situations.</td>
</tr>
<tr>
<td>7–15</td>
<td>practice: oral peer-feedback on actual written assignments reflection on own output post-tests</td>
<td>Learners paired up to give oral constructive criticism on peers’ actual written assignments, then reflected on and evaluated their own output for sociopragmatic appropriateness and pragmalinguistic accuracy.</td>
</tr>
</tbody>
</table>

**Data collection**

The learners’ use of constructive criticism before and after the treatment was measured by means of an oral peer-feedback task and a discourse completion
task (DCT) (see Appendix 2). These two instruments were developed and validated by Nguyen (2005).

To elicit data on constructive criticism, Nguyen (2005) first asked the participants to write a 250-word argumentative essay on a controversial topic. They were then arranged into pairs to give oral constructive critical feedback on each other’s essays. The feedback was based on three main assessment criteria, namely the organizational structure of the essay, the quality of argumentation, and grammar and vocabulary. To triangulate this set of data, Nguyen devised a DCT consisting of four criticizing situations. These situations were constructed based on peer-feedback data taken from a pilot study with four dyads of learners and three dyads of NSs one month prior to the main study. The purpose was to make the situations as comparable to the peer-feedback task as possible.

Similar procedures for conducting the oral peer-feedback task were followed in the current study and the learners’ feedback conversations were audio-taped and transcribed for later analysis. For the DCT task, learners were asked to give responses to four hypothetical situations involving giving critical feedback on a friend’s writing assignment. These situations were taken from Nguyen (2005).

Pre-test data were collected in Week 1, at the onset of the study, and consisted of learners’ oral peer-feedback conversations. Post-test data were collected in Week 15, at the end of the treatment period, and consisted of oral peer-feedback conversations for both groups plus DCT responses for the treatment group. In addition, learners from the treatment group were also required to write a reflection essay in English after the procedure, in which they commented on their learning experience in the course (see Appendix 2). Data from this source were then analyzed for instances of input noticing.

Data analysis

Data from one learner from the treatment group and two learners from the comparison group were excluded from statistical analysis because they did not participate in all pre- and post-tests. Therefore, out of the 30 learners, only 27 were considered for analysis. This included 11 learners in the treatment group and 16 in the comparison group. The data were coded independently into different types of criticizing strategies and modifiers, and then carefully cross-checked by two of the researchers, adapting the categorization scheme devised and validated by Nguyen (2008b) (summarized in Appendix 1). This procedure achieved an inter-rater agreement rate of 95%.

An analytical assessment was then conducted to assign scores to each learner for his or her use of constructive criticism in the tests, using a 10 point scale adapted from Martínez-Flor and Fukuya (2005). This scale consisted of two parts, allowing the researchers to assess both sociopragmatic appropriateness and pragmalinguistic accuracy in learners’ constructive criticism (see Appendix 2).
Each part was rated from 0 to 5, making a total score ranging from 0 to 10. Sociopragmatic appropriateness was assessed in terms of knowledge of what to say to a particular interlocutor in a particular context or situation and determined by the right choice of realization strategies and politeness devices. Pragmalinguistic accuracy was assessed in terms of the learners’ knowledge of various expressions for conveying intentions and determined by the correct use of relevant linguistic structures.

Note that although the learners were assessed for both sociopragmatic appropriateness and pragmalinguistic accuracy, they were to be awarded scores in the latter area only when they were awarded scores in the former area. In other words, pragmalinguistic accuracy was scored only when sociopragmatic appropriateness had been achieved. Note also that scores were awarded only when the learners made use of one of the target forms which had been taught to them during the treatment (see Appendix 1). A learner’s final score on a task was obtained by averaging the sum of the sub-scores that he or she achieved for each criticism made while performing the task. The scoring procedures were conducted independently and cross-checked carefully by two of the researchers, reaching an agreement rate of 98%.

Analysis of the data was based on both between-group (comparison vs. treatment) and within-group (pre-test vs. post-tests) comparisons. Statistical procedures for the former case included ANCOVA, which was to test the difference in means between groups while controlling pre-existing differences due to non-random selection of participants, and Chi-square tests for relatedness, which was to test the differences in frequencies. Statistical procedures for the latter case included one-way repeated measures ANOVA, which was to test differences in means among three test conditions (in the case of the treatment group) and Friedman, which was similar to repeated measure ANOVA but employed in the absence of a normal distribution of data. Where a significant difference was found, paired samples t and Wilcoxon were utilized respectively for post-hoc comparisons. These two procedures were also used to compare differences in means between two test conditions (in the case of the comparison group). The former was used when data were normally distributed and the latter was used when data were not. A modified Bonferroni procedure was conducted when multiple tests on the same set of data were run in order to reduce the chance of Type 1 error.

Results

This section reports the results for the treatment group with respect to their overall pragmatic scores, frequency of use of target forms, and use of modifiers after instruction. Reference to the comparison group will also be made in order to examine whether improvements (if any) observed for the treatment group resulted from instruction.
Table 2 summarizes the two groups’ pragmatic performance as measured by their overall scores gained in the pre- and post-tests. First, comparisons were made between the treatment group’s scores on the two post-tests with the comparison group’s oral post-test scores. Results of ANCOVA tests showed that the treatment group gained higher scores when measured by both oral and DCT post-tests ($M=4.6$ and $5.9$ respectively for the treatment group vs. $M=3.0$ for the comparison group), but significantly differed from the comparison group only when they were measured by the DCT post-test $[F(2, 24)=9.794, p=.005, d=1.37$, suggesting a very large effect size (see Cohen, 1988)]. The difference between their oral post-test and the comparison group’s oral post-test was not found to be significant $[F(2, 24)=2.854, n.s., at p=.10]$, although the effect size calculated for this comparison was considered relatively large ($d=.86$) following Cohen’s (1988) recommendation, suggesting large-magnitude superiority of the treatment group over the comparison group.

Table 2. Means and standard deviations of the pre and post-test overall scores for the comparison and treatment groups

<table>
<thead>
<tr>
<th></th>
<th>comparison ($n=18$)</th>
<th>treatment ($n=12$)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$M$</td>
<td>$SD$</td>
</tr>
<tr>
<td>pre-test</td>
<td>3.3</td>
<td>1.5</td>
</tr>
<tr>
<td>oral post-test</td>
<td>3.0</td>
<td>2.0</td>
</tr>
<tr>
<td>DCT post-test</td>
<td>n.a</td>
<td></td>
</tr>
</tbody>
</table>

Indeed, this superiority became more evident when it came to within-group comparisons, i.e., comparisons of the performance by the same group on different test conditions. Results of a paired samples $t$ suggested that the comparison group did not score higher in their post-test $[t(15)=.682, n.s., at p=.51]$. However, the treatment group’s scores on both their oral and DCT post-tests increased considerably after interventions (Table 2).

Results of a one-way repeated measures ANOVA showed a significant difference when comparing three sets of scores from the treatment group, pre-test, oral post-test and DCT post-test, suggesting a significant improvement for this group after 15 weeks of instruction $[F(2, 18)=7.864$ at $p=.004]$. Results of post hoc paired samples $t$ with Bonferroni correction found that this improvement was evident when the treatment group’s performance was measured by both the oral ($p=.025$) and DCT post-tests ($p=.002$) (Table 3). Effect sizes calculated for both of these analyses were considered very large (pre-test vs. oral post-test: $d=1.49$; pre-test vs. DCT post test: $d=1.94$).
Effects of instruction on the treatment group’s pragmatic competence were also evident when examining their post-instructional improvement in those pragmatic features that caused them problems before the experiment. Table 4 shows the frequencies and percentages of “giving advice” categorized as target forms as used by the comparison and treatment groups in their pre-and post-tests. This strategy was selected for the analysis because previous research has shown learners of English often employed the modal verbs *must, should,* and *have to* inappropriately and rarely hedged their advice when giving constructive criticism to peers (Nguyen, 2005, 2008b). Initial screening of pre-test data also showed that “giving advice” seemed more problematic to learners than “statement of problems.” Thus, it would be worthwhile knowing to what extent this area of difficulty responded to instruction. For the analysis below, an instance of advice-giving was categorized as “target-like” only when it corresponded to the pragmalinguistic features that had been taught to the learners in the experiment and contained the taught modifiers.

### Table 3. Results of post hoc paired samples t with Bonferroni correction for the treatment groups’ pre- and post-test scores

<table>
<thead>
<tr>
<th>absolute value of t</th>
<th>P value</th>
<th>critical alpha</th>
</tr>
</thead>
<tbody>
<tr>
<td>pre-test vs. DCT post-test</td>
<td>4.218</td>
<td>.002</td>
</tr>
<tr>
<td>pre-test vs. oral post-test</td>
<td>2.642</td>
<td>.025</td>
</tr>
</tbody>
</table>

As can be seen, the comparison group consistently made use of a high percentage of nontarget-like advice in both of the tests (71% and 75%, respectively; Table 4). Outcomes of a Chi square test revealed no significant difference for this analysis $[\chi^2(1, N=75)=.015, n.s., at p=.46]$. The treatment group, on the other hand, displayed a significant improvement in their use of target forms after instruction. Their rate of use of target-like advice increased from 12% in the pre-test to 50% in the oral post-test and 72% in the DCT post-test. Their use of non-target forms reduced accordingly from 88% in the pre-test to 50% in the oral post-test and 28% in the DCT post-test (Table 4). These differences were found significant by Chi-square test results (pre-test vs. oral post-test and DCT post-test).
oral post-test: $\chi^2(1, N=87)=13.070$ at $p<.001$; pre-test vs. DCT post-test: $\chi^2(1, N=106)=37.830$ at $p<.001$.

Below are some illustrative examples of how a learner in the treatment group employed the strategy of “giving advice” while carrying out constructive criticism with her peer in three areas (essay organization, ideas and grammar) as observed in her pre-test as well as in the oral and DCT post-tests. As can be seen, before receiving instruction on constructive criticism, she tended to draw heavily on the structure “You should” to advise her classmate how to make changes in the three areas. However, after the 15 week course she used more target-like structures such as may and the conditional clause.

**Pre-test**
Organization: I think you should change the way you arrange your writing.
Ideas: You should give a clear opinion about this [the topic].
Grammar: You make some mistakes. It makes the readers confused.

**Oral post-test**
Grammar: You use (read aloud) but I think it could be (read aloud).

**DCT post-test**
Organization: You may think of rearrange [rearranging] some ideas to make your writing easier to follow.
Ideas: You may need to focus on the topic and omit some irrelevant ideas.
Grammar: Your writing would be much better if you pay some attention to linking words to make the essay cohesive.

Similarly, another student learned to give advice in a more target-like manner by making use of modal structures such as “might” and “could” and conditional structure, which she did not employ before receiving instruction on constructive criticism.

**Pre-test**
Ideas: You need to improve your ideas to make your writing more effective.
Grammar: You need to improve the vocabulary.

**Oral post-test**
Organization: Your writing would be better if you can balance the length between different points.
Ideas: It’s better to rewrite the [thesis] statement or illustrate it in another way, not repeat it.

**DCT post-test**
Organization: You might rearrange the paragraphs to make them clearer.
The effects of explicit metapragmatic instruction

In order to more closely examine whether instruction affects the treatment group’s scores for sociopragmatic appropriateness or pragmalinguistic accuracy, their pre-test and post-test scores in these two aspects were also compared. Table 5 presents the mean scores for appropriateness and accuracy, standard deviations, and the results of the one-way repeated measures ANOVA run for these statistics. As can be seen, learners scored significantly higher for both aspects in the two post-tests. Their scores for appropriateness increased considerably from 1.5 in the pre-test to 3.2 in the oral post-test and 3.5 in the DCT post-test \(F(2, 18)=9.415, p=.002\). There was a corresponding increase in their accuracy scores from 0.6 in the pre-test to 1.4 in the oral post-test and 2.4 in the DCT post-test \(F(2, 18)=5.083, p=.018\) (Table 5).

**Table 5.** Results of one-way repeated measures ANOVA for mean scores of appropriateness and accuracy

<table>
<thead>
<tr>
<th></th>
<th>M</th>
<th>SD</th>
<th>F</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>appropriateness</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pre-test</td>
<td>1.5</td>
<td>1.1</td>
<td>9.415</td>
<td>.002</td>
</tr>
<tr>
<td>oral post-test</td>
<td>3.2</td>
<td>0.9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DCT post-test</td>
<td>3.5</td>
<td>1.3</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>accuracy</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pre-test</td>
<td>0.6</td>
<td>1.0</td>
<td>5.083</td>
<td>.018</td>
</tr>
<tr>
<td>oral post-test</td>
<td>1.4</td>
<td>2.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DCT post-test</td>
<td>2.4</td>
<td>1.0</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Results of post hoc paired samples t with Bonferroni correction indicated that the learners significantly improved their sociopragmatic appropriateness on both the oral and DCT post-tests (pre-test vs. oral post-test: \(p=.005, d=1.69\); pre-test vs. DCT post-test: \(p=.002; d=1.66\)). However, significant improvement in the area of pragmalinguistic accuracy was evident only in the DCT (\(p=.003, d=1.79\)) and not in their oral post-test scores (\(p=.22, n.s., d=.48\)) (see Table 6), although the effect size calculated for the latter analysis was considered medium (see Cohen, 1988).
Table 6. Results of post hoc paired samples $t$ with Bonferroni correction for the treatment group’s scores of appropriateness and accuracy

<table>
<thead>
<tr>
<th></th>
<th>absolute value of $t$</th>
<th>$P$ value</th>
<th>critical alpha</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>appropriateness</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pre-test vs. DCT post-test</td>
<td>4.238</td>
<td>.002</td>
<td>.05/3=.017</td>
</tr>
<tr>
<td>pre-test vs. oral post-test</td>
<td>3.541</td>
<td>.005</td>
<td>.05/2=.025</td>
</tr>
<tr>
<td><strong>accuracy</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pre-test vs. DCT post-test</td>
<td>3.840</td>
<td>.003</td>
<td>.05/3=.017</td>
</tr>
<tr>
<td>pre-test vs. oral post-test</td>
<td>1.304</td>
<td>.2</td>
<td>.05/2=.025</td>
</tr>
</tbody>
</table>

Instruction also seemed to affect the treatment group’s frequency of use of modifiers to soften their constructive criticism. Table 7 presents the results of Friedman tests run for medians of modifiers produced per criticism by this group in three test conditions. Overall, the learners employed a notably greater number of modifiers in the two post-tests than in the pre-test (pre-test: $Mdn$=.20, oral post-test: $Mdn$=1.0, DCT post-test: $Mdn$=1.0, $\chi^2$=11.870 at $p$=.003). However, post hoc Wilcoxon analyses showed that although there were significant improvements for both the oral and DCT post-tests (Table 7), these improvements were regarding learners’ use of internal modifiers (pre-test vs. oral post-test: $p$=.020; $d$=1.24; pre-test vs. DCT post-test: $p$=.006, $d$=1.63) rather than their use of external modifiers, although the effect size calculated for post-instructional improvement in external modifiers when measured in the oral post-test was rather large (pre-test vs. oral post-test: $p$=.019, n.s., $d$=1.07; pre-test vs. DCT post-test: $p$=.20, n.s., $d$=.20) (Table 8).

Table 7. Results of Friedman for mean scores of criticism modifiers

<table>
<thead>
<tr>
<th></th>
<th>$mdn$</th>
<th>$\chi^2$</th>
<th>$P$ value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>external modifiers</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pre-test</td>
<td>.20</td>
<td>10.255</td>
<td>.006</td>
</tr>
<tr>
<td>oral post-test</td>
<td>.95</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DCT post-test</td>
<td>.45</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>internal modifiers</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pre-test</td>
<td>.00</td>
<td>9.814</td>
<td>.007</td>
</tr>
<tr>
<td>oral post-test</td>
<td>.20</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DCT post-test</td>
<td>.50</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>total mean</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pre-test</td>
<td>.20</td>
<td>11.870</td>
<td>.003</td>
</tr>
<tr>
<td>oral post-test</td>
<td>1.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DCT post-test</td>
<td>1.0</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 8. Results of post hoc Wilcoxon for the treatment group’s means of modifiers

<table>
<thead>
<tr>
<th>absolute value of z</th>
<th>P value</th>
<th>critical alpha</th>
</tr>
</thead>
<tbody>
<tr>
<td>external</td>
<td></td>
<td></td>
</tr>
<tr>
<td>pre-test vs. oral post-test</td>
<td>2.355</td>
<td>.019</td>
</tr>
<tr>
<td>pre-test vs. DCT post-test</td>
<td>1/247</td>
<td>.20</td>
</tr>
<tr>
<td>internal</td>
<td></td>
<td></td>
</tr>
<tr>
<td>pre-test vs. DCT post-test</td>
<td>2.752</td>
<td>.006</td>
</tr>
<tr>
<td>pre-test vs. oral post-test</td>
<td>2.319</td>
<td>.020</td>
</tr>
<tr>
<td>total</td>
<td></td>
<td></td>
</tr>
<tr>
<td>pre-test vs. DCT post-test</td>
<td>2.714</td>
<td>.007</td>
</tr>
<tr>
<td>pre-test vs. oral post-test</td>
<td>2.361</td>
<td>.018</td>
</tr>
</tbody>
</table>

Discussion

Research question 1
The present study has sought to address two research questions: (1) Do learners benefit from explicit instruction in constructive criticism in English? and (2) Do some aspects of giving constructive criticism benefit more from explicit instruction than others? With regard to the first question, a positive answer was supported by the findings that the instructed students generally outperformed their uninstructed peers and improved their pragmatic performance significantly compared to their pre-test results. They also employed more target-like advice and internal modifiers, which constituted an area of difficulty for them before receiving pedagogical intervention.

To interpret the relative effectiveness of the instructional approach employed in this study as compared to instructional approaches used in previous studies, effect sizes were calculated for both the between-group (i.e., treatment vs. comparison) and within-group (i.e., pre-test vs. post-test) analyses. These were compared with the corresponding figures reported by Jeon and Kaya (2006) in their meta-analysis of instructed L2 pragmatics studies, and Norris and Ortega (2000) in their meta-analysis of instructed L2 grammar studies. In the present study, the mean treatment-versus-comparison effect size ($d=1.12$), calculated by averaging the effect sizes for oral post-test contrast ($d=.86$) and DCT post-test contrast ($d=1.37$), was considered relatively large, following Cohen’s (1988) recommendations. This was larger than both the mean effect size of .59 reported for pragmatic instruction by Jeon and Kaya (2006) and the mean effect size of .96 reported for grammar instruction by Norris and Ortega (2000) (see Table 9).
Similarly, the mean pretest-to-posttest effect size in the present study ($d=1.71$), calculated by averaging the effect sizes for oral post-test contrast ($d=1.49$) and DCT post-test contrast ($d=1.94$), was found to be larger than the corresponding figures reported in Jeon and Kaya (2006) and Norris and Ortega (2000) (see Table 9) and considered relatively large, following Cohen (1988). These results seem to suggest a greater magnitude of effects for the instructional approach employed in the current study.

Positive instructional effects were also evident in the learners’ comments about their learning experience in end-of-course reflective essays. These effects, reported by all twelve learners from the treatment group, included heightened awareness of the politeness aspect of giving constructive criticism, enhanced metapragmatic knowledge, as well as knowledge of the linguistic devices used for expressing and softening their constructive criticism. For example, one learner wrote:

“Since the first year I’ve checked and given feedback to my friends. However, I did not have much knowledge about how to give feedback. I just focused on my friends’ mistakes and gave feedback directly to their mistake without caring much about how my friends felt when receiving my feedback. Recently, I learn something about giving feedback. I suppose that the most important thing I learned is to give feedback that doesn’t hurt my friends but still points out their mistakes by using mitigation.”

The overall benefits of explicit metapragmatic instruction reported in the present study are consistent with the findings of previous studies (Billmyer, 1990; Bouton, 1994; Lyster, 1994; Martínez-Flor & Fukuya, 2005; Olshtain & Cohen, 1990; Safont, 2003; Yoshimi, 2001) and can be attributed to various factors. First, instructional activities such as guided discovery and discussion of metapragmatic rules served to draw the learners’ attention to the targeted features and caused them to notice the relationship between form, meaning and use, thus aiding their acquisition of these features (Schmidt, 1993, 1995). Indeed, instances of input-noticing resulting from this learning process were reported in the learners’ end-of-course reflective essays, adding further evidence about the noticing function of the above-mentioned instructional activities. For example, one learner wrote:

Table 9. A comparison of mean effect sizes for treatment-versus-comparison and pretest-to-posttest contrasts among three studies

<table>
<thead>
<tr>
<th>Instructional Foci</th>
<th>Treatment vs. Comparison</th>
<th>Pre-test vs. Post-test</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Present Study</strong></td>
<td>L2 Pragmatics</td>
<td>1.12</td>
</tr>
<tr>
<td>Jeon &amp; Kaya (2006)</td>
<td>L2 Pragmatics</td>
<td>.59</td>
</tr>
<tr>
<td>Norris &amp; Ortega (2000)</td>
<td>L2 Grammar</td>
<td>.96</td>
</tr>
</tbody>
</table>
“One important thing about giving critical peer feedback I have learned in the last session is that I know how to soften criticism by using the modal verbs like *might, would*, etc. or uncertain structures like *I'm not sure or It seems*.”

Similarly, another learner commented:

“In terms of language, by learning about giving critical peer feedback, I know how to point out the mistakes without depressing or hurting my friends. For example, instead of saying straightly about the weak points, I can use mitigation devices, like ‘I think this point is a little bit far from the topic in the way that....’ to soften the comment.”

Second, the learners also benefit from the explicit correction they received from their teachers, especially when the need to focus on form arose out of meaningful communication (e.g., giving constructive criticism in oral role-plays and oral peer-feedback tasks). Such interational feedback causes learners to attend to problematic forms, notice the gap between their own output and the target form, and modify their output accordingly, which is believed to contribute to learning (Long, 1983, 1996; Swain, 1985, 1995, 2005). Similar learning benefits also result from learners’ reflection on their own output after each peer-feedback session, which also provides opportunities for gap-noticing and modified output. Instances of gap-noticing were found in learners’ reflective journals too. For example, one learner commented:

“I often directly give comments on what my friends do wrongly. However, now I usually give my friends' strong points first and then I comment on his/ her weak point so that my friends are not really disappointed and hurt.”

Similarly, another student wrote:

“When it comes to giving comments, I usually use expressions such as *I think that you should* or *You should not*, which are direct and clear. However, sometimes I found my friends not really satisfying with my feedback due to its straightforwardness. Therefore, I have tried to soften it. In other words, I have attempted to make use of softeners such as *You might*, or *I'm not pretty sure* so that my friend can find it easier to receive the feedback. The result is that my peers take my comments more seriously.”

**Research question 2**

With regard to the second research question, findings suggested that while instruction generally benefited the learners’ performance of constructive criticism, it seemed to affect different aspects differently. For example, the learners' post-instructional improvement was more evident in sociopragmatic appropriateness than in pragmalinguistic accuracy. While the gains made in their sociopragmatic appropriateness scores were significant in both the oral and DCT post-tests, their gains in pragmalinguistic accuracy were significant
only in the DCT post-test, where they were not under pressure to produce online speech and thus could plan their production more carefully.

**Table 10.** Pretest-to-posttest effect sizes for gains in sociopragmatic appropriateness and pragmalinguistic accuracy

<table>
<thead>
<tr>
<th>Effect size for pretest vs. oral post-test contrast</th>
<th>Sociopragmatic appropriateness</th>
<th>Pragmalinguistic accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.69</td>
<td>.48</td>
<td></td>
</tr>
<tr>
<td>Effect size for pre-test vs. DCT post-test contrast</td>
<td>1.66</td>
<td>1.79</td>
</tr>
<tr>
<td>Mean effect size</td>
<td>1.67</td>
<td>1.13</td>
</tr>
</tbody>
</table>

Comparing pretest-to-posttest effect sizes for gains in these two aspects, it was found that the mean effect size for gains in sociopragmatic appropriateness ($d=1.67$) was larger than that in pragmalinguistic accuracy ($d=1.13$). This was because of the larger magnitude of effects on sociopragmatic appropriateness ($d=1.69$) as compared to effects on pragmalinguistic accuracy ($d=.48$) when both aspects were measured in the oral post-test condition but not when they were compared in the DCT post-test condition ($d=1.66$ for sociopragmatic appropriateness; $d=1.79$ for pragmalinguistic accuracy) (Table 10).

These results suggest some possible interplay between instructional effects and processing issues. Instructional effects were more evident in test conditions such as the written DCT, which allowed a greater amount of planning time and imposed relatively less processing demand. In other words, instructional effects presented themselves more obviously when learners were measured for what they could possibly do under controlled, pressure-free conditions than when they were measured for what they actually did under the pressure of real-time interaction.

On the one hand, these findings carry important implications for future studies in designing their outcome measures to capture instructional effects on different aspects of learners’ pragmatic performance. Written DCTs may be effective in measuring what learners know about L2 pragmatics while oral, spontaneous production tasks are effective in measuring whether they can retrieve this knowledge for use in real time.

The findings also suggest pedagogical implications in fostering pragmatic fluency. In particular, the learners’ lower accuracy in the oral post-test suggested that they did not achieve full control over pragmalinguistic forms due to a lack of fluency in the L2, with the result that they did not use these forms accurately when under time constraints, despite their knowledge of the forms. According to Nation (2011), fluency development requires not only plentiful and repeated communicative practice in which learners can work to avoid communication
breakdown but also some pressure on learners to perform at a faster than usual speed. Although the instructional approach employed in the current study met the first condition, it fell short of the second. While it allowed the learners to be regularly engaged with meaning via communicative tasks such as role plays and oral peer-feedback, it did not require them to produce faster speech each time they delivered it. In the future, teachers should take timed practice into consideration when designing tasks to help improve learners’ automaticity in using pragmalinguistic conventions. In addition, perhaps allowing plentiful time for pre-task planning may also compensate for this lack of pragmatic fluency.

Further, the instructional approach employed in the present study also seemed to produce different effects on learners’ use of external and internal modifiers. Findings showed that instruction caused learners to use significantly more internal modifiers but not more external modifiers. The mean effect size calculated for improvements in internal modifiers ($d=1.43$) was twice as large as that calculated for improvements in external modifiers ($d=.63$). These results are worth noting, given the findings of earlier studies which suggest that internal modifiers tend to pose more challenges to L2 learning than do external modifiers because of their less noticeable pragmatic meaning and the possibility of adding to the structural complexity of the speech (Hassall, 2001; Nguyen, 2008a). Indeed, the findings of the present study indicate that learners generally preferred external modifiers to internal modifiers before pedagogical interventions (c.f.: $M=.3$ for external modifiers vs. $M=.02$ for internal modifiers). After receiving instruction, however, they preferred external modifiers only when they participated in the spontaneous, oral peer-feedback task ($M=1.5$ for external modifiers vs. $M=.2$ for internal modifiers). When they were given more processing time in the pressure-free, written DCT condition, they employed more internal modifiers ($M=.4$ for external modifiers vs. $M=.6$ for internal modifiers).

The above findings might be explained by the observation that unlike external modifiers, which learners already employed extensively before treatment, internal modifiers represented ‘newer’ knowledge to learners, i.e., something they were less likely to notice otherwise, and therefore attract more of their attention. Unfortunately, without conducting retrospective interviews with the learners, no definite answers can be provided at this point and obviously further investigation into the issue is required. All in all, the learners’ considerable improvement in the area of internal modifiers is worth applauding and suggests the advantage of explicit instruction in raising the learners’ awareness of forms and meanings which would have been far less salient to them without instruction.

**Conclusion**

Overall, like many previous studies (Jeon & Kaya, 2006; Rose, 2005), the findings of the present study indicate that as in the case of L2 grammar (Norris & Ortega, 2000), L2 pragmatic development can be enhanced through an appropriate
pedagogical approach. More specifically, learning L2 pragmatics can benefit from opportunities for input noticing, meaningful output practice, corrective feedback and modified output. These findings thus lend support to second language acquisition theories which hold that part of learning must require consciousness and that input and output modified via meaningful communication can aid learning (Long, 1983, 1996; Schmidt, 1993, 1995; Swain, 1985, 1995, 2005). These findings also suggest that constructive criticism is teachable, despite the fact that it is a challenging speech act set which may require a relatively high degree of linguistic complexity as well as pragmatic sophistication. The findings thus make a strong case for teaching the language for delivering constructive criticism in peer assessment tasks in L2 classrooms. Such instruction is especially necessitated in EFL contexts where learners have limited exposure to sociopragmatic rules and pragmalinguistic resources in the TL.

The findings of this study also suggest some effects of test methods on learners’ pragmatic performance, as pointed out in many previous studies (e.g., Bardovi-Harlig & Hartford, 1993; Beebe & Cummings, 1996; Hartford & Bardovi-Harlig, 1992; Hinkel, 1997; Jeon & Kaya, 2006; Johnston, Kasper, & Ross, 1998; Nguyen, 2005; Rose, 1992; Sasaki, 1997; Wolfson, 1989; Yuan, 2001). In this study, learners’ superior performance in the DCT task, especially regarding their overall pragmatic scores as compared to their uninstructed peers, and gains in scores for pragmalinguistic accuracy as compared to gains in sociopragmatic appropriateness, might have been explained by the written nature of the DCT. More specifically, the task allows them more processing time, and thus is less imposing on their processing capacity, which enables them to attend to those forms which have not yet been automatized (Nguyen, 2005). In this way, the DCT provides useful information about learners’ declarative knowledge of L2 pragmatics, or what learners can do under controlled conditions, in addition to their procedural knowledge, or what they actually do in real-time communication, which might be measured more faithfully by means of spontaneous interaction, e.g., the oral peer-feedback task (see Anderson, 1976, 1980, for further discussion of declarative and procedural knowledge).

These findings raise several issues for future research in designing data collection methods. Employing multiple and different modes of data collection would potentially help capture the subtle changes and differences caused by instruction. Comparing studies that employ elicited data exclusively with those employing both elicited and naturally occurring data, Jeon and Kaya (2006) found a larger effect size for the latter. However, they warn against any premature conclusion because of the limited number of studies available for their meta-analysis. Further systematic research might be conducted to address the relationship between observed instructional effectiveness and types of outcome measure so as to inform the methodology of future instructed L2 pragmatics studies.
The findings of the present study also carry important pedagogical implications. Bialystok (1993) assumes that the acquisition of pragmatic knowledge in the TL requires two separate cognitive processes: that of acquiring the knowledge and that of gaining automaticity in using this knowledge. Previous research shows that high proficiency learners sometimes learn grammatical forms but do not learn all their functions, with the result that they do not always put these forms to their correct pragmatic use (Hassall, 2001; Nguyen, 2005, 2008a). In other cases, difficulty does not arise from deviant knowledge of form-function mapping but from learners’ inability to draw on this knowledge when they need it, as was the case with the treatment group in the present study. Therefore, pragmatic instruction should allow not only focus on attention to form-function mapping but also control over this attention so that learners can produce pragmatically appropriate and linguistically accurate speech acts in real time.

Finally, a word of caution is in order. Despite some interesting insights, the findings in this chapter are still exploratory and should not be generalized without careful consideration of the limitations of the study. These limitations included the employment of a small, nonrandomized, gender biased sample (with all but one of the students being female), which affects the representativeness of the findings. Further, the lack of a delayed post-test also made it impossible to measure the long-term retention of instructional effect on the learners. Therefore, while positive instructional effects were observed at the end of the intervention, one cannot conclude with confidence that these effects will be maintained beyond immediate post-interventional observations. What is more, the fact that the study employed two different instructors for the treatment and comparison groups may have also contributed to the differences in learning outcomes. In particular, although the instructors followed the specified instructional protocols, they might still have differed in their teaching styles, which could have affected the students’ motivation to learn. Future studies need to include classroom observation in order to document how the instructors deliver the teaching and how the students participate in the activities under investigation.

Notes

1. A speech act set is composed of a range of strategies, any combination of which could perform it (Olshtain & Cohen, 1983) and was first used by these researchers to describe apologies.

2. Criticizing refers to the act of finding fault with another person (Tracy, van Dusen, & Robinson, 1987). It may involve criticizing another’s work, ideas, or personal style (Tracy & Eisenberg, 1990). However, it should be noted that criticizing someone’s work might be expected to be constructive and supportive in nature. Thus, the type of criticism under inquiry in the present study may involve a lower level of ‘infraction’
than the more ‘biting’ types of criticism such as criticizing someone’s appearance or behavior.

In many studies on workplace communication or communication for teachers in the context of mentoring, the terms ‘criticism’ and ‘critical feedback’ are used interchangeably (Tracy & Eisenberg, 1990; Wajnryb, 1993, 1995).

Although this concept was first mentioned in a published article by Blum-Kulka and Olshtain (1984), as they point out in that article, it was originally coined and used by Susan Ervin-Tripp and David Gordon in their coding manual for analyzing requests within the framework of their unpublished Social Development and Communication Strategies project.

The students were not told what counted as ‘constructive criticism’ but were instead simply asked to find points in their peer’s essays they felt unsatisfied with and to comment on these points. The instruction also indicated that although the task required them to comment specifically on the points they were unsatisfied with, they might also comment on the good points (if any) in the essay (see Appendix 2).

References


Appendix 1: Target forms included in the study

<table>
<thead>
<tr>
<th>target forms</th>
<th>examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Statement of problems: stating the problem or errors found with R’s choice,</td>
<td>I thought you had two conclusions.</td>
</tr>
<tr>
<td>work, or products</td>
<td>I didn’t see your conclusion.</td>
</tr>
<tr>
<td>Giving advice: proposing a potential solution to the problem or errors</td>
<td>You might want to delete the comma.</td>
</tr>
<tr>
<td>identified</td>
<td>Perhaps you could pay more attention to grammar.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>modifiers</th>
<th>examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. External: additional comments, separate from the problem identification</td>
<td>It was an interesting paper.</td>
</tr>
<tr>
<td>and advice giving</td>
<td>That was a great presentation.</td>
</tr>
<tr>
<td>Compliment: saying something good about the thing you are going to criticize</td>
<td></td>
</tr>
<tr>
<td>Disarmer: showing awareness of the potential offence that your comment may</td>
<td>You had a few spelling mistakes here and there but I think that’s because you’re writing pretty quick, nothing too major.</td>
</tr>
<tr>
<td>cause R</td>
<td></td>
</tr>
<tr>
<td>Grounder: giving reasons to justify your comment</td>
<td>I think “is” would be better than “are” there because traffic is single.</td>
</tr>
<tr>
<td>2. Internal: linguistic softeners</td>
<td></td>
</tr>
<tr>
<td>Question: using questions rather than bald statements or imperatives to</td>
<td>Did you summarize the main idea?</td>
</tr>
<tr>
<td>identify a problem or propose a potential solution</td>
<td>Could this work?</td>
</tr>
<tr>
<td>Past tense: creating a sense of distance between the speaker and the</td>
<td>I thought it would make more sense that way.</td>
</tr>
<tr>
<td>comment</td>
<td>Maybe you could’ve explained it a little bit more.</td>
</tr>
<tr>
<td>Modal verbs (e.g., may, might [want to], could, would)</td>
<td>I’m not sure but maybe you could cut out the second section.</td>
</tr>
<tr>
<td>Modal adverbs</td>
<td>maybe, perhaps, probably Perhaps you might want to check that again.</td>
</tr>
<tr>
<td>Uncertainty phrases: showing hesitation or uncertainty about the criticism</td>
<td>I wasn’t sure that was the best phrase you could’ve used.</td>
</tr>
<tr>
<td></td>
<td>I don’t know that I agree with the point you made.</td>
</tr>
<tr>
<td>Hedges (e.g., kind of, sort of, seem)</td>
<td>This sentence was sort of unclear.</td>
</tr>
<tr>
<td>Downgraders (e.g., a bit, a little [bit], quite, rather)</td>
<td>Your introduction seemed a little too long.</td>
</tr>
</tbody>
</table>

(adapted from Nguyen, 2005)
Pragma-linguistic conventions for realizing constructive criticism
(adapted from Nguyen, 2005)
1. Statement of problem:
   • NP was ADJ
   • You V (past tense)
   • You had (a/an) (ADJ) NP
2. Giving suggestion:
   • You can
   • You could + V
   • You could have + V (past participle)
   • You may + V
   • You might + V
   • You might want to + V
   • (If I were you) I would
   • It would be better if you + V
   • It would be better + V (infinitive)
   • If you + V, it may
   • NP may be + V (past participle)
   • Why don’t you?

note: Structures to be avoided because they might produce a negative effect on the hearer:
   • Imperatives: “Give more examples”
   • Strong modal structures: should, must, have to, ought to
   • Negative words: “wrong,” “weak,” etc.
Appendix 2: Data collection instruments

Oral Peer Feedback Task

**Instruction:** You have written a 250 word essay. Now show it to your friend and work in pairs. Take as much time as you need to read your friend’s essay. *Try to find something you are unsatisfied with about the essay and comment on it.* Does your friend think the same? *Discuss with him or her some of the things you think are wrong with the essay.* Your friend will also give comments on your essay. Do you agree with his or her ideas?

Do not try to discuss two essays at the same time. Work on one at a time only. You may discuss between yourselves whose essay to work on first.

**Note**
1. It is important that you understand the task completely, so before you start you are encouraged to ask questions if you find any detail you are not sure of.
2. Although the task requires you to comment specifically on the points you are unsatisfied with in your friend’s essay, you can also comment on the good points (if any) in his or her essay.

You may want to ask yourself the following questions when giving feedback on your friend’s essay:

**Organization**
1. Does the essay directly discuss the topic?
2. Is there a clear organizational structure, i.e., does it have three parts: an introduction, a body, and a conclusion?
3. Is the introduction brief and to the point? Does it indicate the main ideas that the writer will discuss in the body?
4. Are there several paragraphs in the body, each making a different specific point?
5. Is there a brief conclusion that summarizes the main points in the argument?
6. Are the ideas properly linked?

**Ideas**
7. Is the writer’s opinion clear or do you think the writer is not quite sure what he or she thinks?
8. Are the ideas relevant and well supported by evidence and examples?
9. Are the ideas presented logically?
10. Are the ideas developed from one paragraph to another or does the
writer just repeat himself or herself?

**Grammar/vocabulary**
11. Is there a variety of sentence structure and vocabulary or is there a lot
of repetition?
12. Are the linking words (i.e., words used to link ideas) helpful or do they
confuse you?
13. Are the sentences grammatically accurate?
(adapted from Nguyen, 2005)

**Discourse Completion Task**

Please read the instruction and the given situations carefully and write your
answers in English in the space provided under each situation. It is important
that you understand the requirements completely, so before you start, you are
encouraged to ask questions if you find something you do not understand.

Thank you for your assistance.

**Instruction**: In reference to an essay that your friend has just written, what
would you say in the following hypothetical situations?

**Example**

*Situation 0:* What would you say to your friend if you thought her essay
exceeded the limit of 250 words?
You: “I think ….”

*Situation 1:* What would you say to your friend if you thought his or her essay
was not very well organized, so it was rather difficult to follow his or
her ideas?
You:

*Situation 2:* What would you say to your friend if you thought in some instances
he or she didn’t support his or her arguments with relevant examples
and evidence, so these arguments were hard to convince readers?
You:

*Situation 3:* What would you say to your friend if you thought his or her essay
lacked a focus, and so it was difficult to follow his or her arguments?
You:

*Situation 4:* What would you say to your friend if you thought he or she didn’t
often make use of linking words, so the essay seemed to lack cohesion?
You:
(adapted from Nguyen, 2005)
Reflection Essay

At home write a one-page reflection on the experience you gained in giving critical peer feedback, using the following guiding questions to help you. Submit this reflection to your instructor in the next class meeting.

Guiding questions
1. Name AT LEAST one important thing about giving critical peer feedback that you have learned in the last sessions.
2. How has your idea about giving critical feedback changed as a result of this gained experience?
3. How helpful is this experience to you?
4. How are you going to make use of this experience in your other classes?
Appendix 3: Grading criteria

**Scoring for pragmatic appropriateness level**

**Appropriateness**: knowing what to say to a particular interlocutor in a particular context of situation; determined by the right choice of realization strategies, that is, giving suggestion rather than demanding changes because of the equal power status between the interlocutors.

<table>
<thead>
<tr>
<th>grammatical level</th>
<th>score</th>
<th>examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>inappropriate Head Act (HA)</td>
<td>0</td>
<td>You <em>must</em> pay attention to grammar.</td>
</tr>
<tr>
<td>appropriate HA (without modifiers or with inappropriate modifiers)</td>
<td>4</td>
<td>You can rewrite your introduction (without modifiers). You can rewrite your introduction <em>because it’s too bad</em> (with inappropriate modifier).</td>
</tr>
<tr>
<td>appropriate HA + appropriate modifiers</td>
<td>5</td>
<td><em>Perhaps you could</em> pay more attention to grammar.</td>
</tr>
</tbody>
</table>

**Scoring for linguistic accuracy**

**Accuracy**: knowing the expressions for conveying intentions; determined by the correct usage of relevant linguistic structures, e.g., saying "*if I were you, I would*" but not "*If I were you, I will.*"

<table>
<thead>
<tr>
<th>grammatical level</th>
<th>score</th>
<th>examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>incorrect pragma-linguistic form</td>
<td>0</td>
<td>Your ideas <em>would</em> be linked more properly (the correct form is “Your ideas <em>could</em> be linked more properly”).</td>
</tr>
<tr>
<td>correct pragma-linguistic form + incorrect connecting part/inaccurate modifiers</td>
<td>4</td>
<td>It would be better if you could revising it (the correct form is “<em>If you could revise it</em>”).</td>
</tr>
<tr>
<td>correct pragma-linguistic form + correct connecting part/accurate modifiers</td>
<td>5</td>
<td>You <em>may want to explain this a little bit</em> more.</td>
</tr>
</tbody>
</table>

**Note**

- Scores to be assigned for the target forms only (see Appendix 1)
- Scores to be assigned for both linguistic realization strategies (i.e., identifying problem and giving advice) and modifiers (i.e., softeners).
• Scores to be assigned for both pragmatic appropriateness and linguistic accuracy but linguistic accuracy is to be scored only when pragmatic appropriateness has been achieved. In other words, pragmatic appropriateness is a necessary condition for linguistic accuracy.

• Modifiers to be awarded scores only when appropriate head act has been used. E.g., if a candidate says “Your writing is good but you must pay attention to grammar,” the compliment “Your writing is good” is not to be awarded a score because the head act “You must” is not appropriate.

• A candidate’s final score will be calculated following this procedure: (1) calculate the total number of criticisms he or she has made; (2) then generate the score for each criticism by adding its score for appropriateness and score for accuracy; (3) calculate the sum for all criticisms made; (4) finally divide this sum by the number of criticisms made.

(adapted from Martínez-Flor & Fukuya, 2005)