The Effects of Practice Modality on Pragmatic Development in L2 Chinese

Shuai Li  
*Georgia State University*, sli12@gsu.edu

Naoko Taguchi  
*Carnegie Mellon University*, taguchi@andrew.cmu.edu

Follow this and additional works at: [https://scholarworks.gsu.edu/mcl_facpub](https://scholarworks.gsu.edu/mcl_facpub)  
Part of the Other Languages, Societies, and Cultures Commons

Recommended Citation

Li, Shuai and Taguchi, Naoko, "The Effects of Practice Modality on Pragmatic Development in L2 Chinese" (2014). *World Languages and Cultures Faculty Publications*. 67.  
[https://scholarworks.gsu.edu/mcl_facpub/67](https://scholarworks.gsu.edu/mcl_facpub/67)

This Article is brought to you for free and open access by the Department of World Languages and Cultures at ScholarWorks @ Georgia State University. It has been accepted for inclusion in World Languages and Cultures Faculty Publications by an authorized administrator of ScholarWorks @ Georgia State University. For more information, please contact scholarworks@gsu.edu.
The Effects of Practice Modality on Pragmatic Development in L2 Chinese

ABSTRACT

This study investigated the effects of input-based and output-based practice on the development of accuracy and speed in recognizing and producing request-making forms in L2 Chinese. Fifty American learners of Chinese with intermediate-level proficiency were randomly assigned to an input-based training group, an output-based training group, and a control group. The input and output groups practiced the target forms over four consecutive days. The control group did not practice the forms. The effects of practice were measured by a Listening Judgment Test (LJT) and an Oral Discourse Completion Test (ODCT). The results showed that the effects of input-based and output-based practice were shared across task modalities on measures of performance accuracy (i.e., accuracy in the LJT and ODCT) but not on measures of performance speed (i.e., LJT response times, ODCT planning times and speech rates).

Key words: interlanguage pragmatics, L2 instruction, Chinese, requests, speech acts, input-based and output-based practice

An important topic in instructional ILP (interlanguage pragmatics) research is the role of various types of instruction in enhancing L2 pragmatic development. Research in this area has centered on comparing the effects of explicit and implicit instructional conditions (hereafter explicit and implicit conditions). However, findings have been inconsistent: Although some studies demonstrated an advantage of explicit over implicit conditions, others showed that both were equally effective or ineffective (see Takahashi, 2010a, 2010b, for recent reviews). A review of the literature indicates that modality of instruction (i.e., input-based, vs. output-based) may influence the effects of explicit and implicit conditions; yet such modality effects have not been empirically examined. A closely related but under-researched topic is how to assess instructional effects. Researchers have focused on gains in performance accuracy, which is an indicator of underlying *pragmatic knowledge*. However, gains in performance speed, which can serve as an indicator of the ability to process pragmatic knowledge (i.e., *processing ability*), has largely been left unexamined. Because *knowledge* and *processing* are key components of pragmatic competence (Kasper, 2001; Taguchi, 2012), both should be examined to better understand instructed L2 pragmatic development. This study is an effort to address the above issues by examining the effects of input-based and output-based practices on the development of accuracy and speed in request-making in L2 Chinese.

**BACKGROUND**

*Knowledge and Processing in Pragmatic Competence*

   Pragmatic competence has been theorized to incorporate *knowledge* and *processing* components (e.g., Bialystok, 1993; Kasper, 2001; Taguchi, 2012). The development of
knowledge involves expanding pragmalinguistic (i.e., linguistic forms for conveying pragmatic functions) and sociopragmatic (i.e., sociocultural conventions governing language use) repertoires (Leech, 1983; Thomas, 1983). It also involves making connections of the two so as to form a refined understanding of form-function-context mappings, that is, the knowledge of which form(s) to use for conveying intended meaning(s) in context(s) (Bialystok, 1993; Faerch & Kasper, 1984). With refined pragmatic knowledge, accuracy of performance can improve. Various measures of performance accuracy (e.g., appropriateness ratings for DCT, accuracy scores in multiple-choice questionnaire) can thus serve as indicators of pragmatic knowledge.

On the other hand, processing refers to the executive abilities to efficiently access, integrate and demonstrate relevant pragmatic knowledge in real-time communication (Taguchi, 2012). The development of processing ability involves acquiring “control strategies to attend to the intended interpretations in contexts and to select the forms from the range of possibilities that satisfy the social and contextual needs of the communicative situation” (Bialystok, 1993, p.54). With repeated comprehension and production of pragmatic meanings, the execution of such strategies can gradually become automatized to allow fluent performance. In interlanguage pragmatics, indicators of processing ability include measures of performance speed, such as response times for interpreting implied meanings (e.g., Taguchi, 2005, 2007a, 2012), and planning times and speech rates in producing speech acts (e.g., Li, 2012; Taguchi, 2008). In the following sections, we will review instructional studies based on knowledge and processing dimensions of pragmatic competence.

Effects of Instruction on the Development of Pragmatic Knowledge

Concerning the development of pragmatic knowledge, research comparing the effects of explicit versus implicit instructional conditions has dominated the field for over two decades. Because instructors provide metapragmatic information in an explicit condition and withhold such information in an implicit condition (Rose, 2005), the explicit condition is more effective than the implicit condition in drawing learners’ attention to target features. By referring to Schmidt’s (1990, 1993, 2001) *Noticing Hypothesis*, which posits that noticing (of linguistic features) is a necessary condition for acquisition, researchers generally consider the explicit condition to be more conducive to pragmatic development than the implicit condition (e.g., Jeon & Kaya, 2006; Kasper & Roever, 2005).

Empirical results, however, do not fully support the above view. The inconsistency in research findings becomes prominent when modality of instruction (i.e., input-based, output-based) is considered in evaluating instructional effectiveness. This is shown in the following review of 15 studies published between 1990 and 2012 that compared the effects of explicit and implicit conditions. These studies were marked with an * in the list of references, and the review chart is available in Appendix A. The 15 studies were classified according to instructional modality: input-based studies (e.g., watching video clips, reading and analyzing dialogues), output-based studies (e.g., role play), and dual-modality studies.

The output-based studies unanimously showed an advantage for the explicit condition over the implicit condition in terms of instructional effectiveness (Félix-Brasdefer, 2008; Ghobadi & Fahim, 2009; House, 1996). For example, in House’s (1996) study, the explicit-instruction group (hereafter explicit group) received teacher-fronted metapragmatic instruction and did role-play activities. On the other hand, the implicit-instruction group

(hereafter implicit group) was given handouts listing situationally appropriate utterances and did role-play activities. After a 14-week instructional period, the explicit group outperformed the implicit group on all measures of a role play task. In this study, the explicit metapragmatic instruction facilitated the learners’ noticing of the target features, and the subsequent role play activities enabled them to strengthen the explicitly taught pragmatic knowledge. Although the implicit group also practiced the target features through role plays, the lack of metapragmatic instruction made them less likely to notice the features than the explicit group, and this difference could explain their moderate gain when compared with the explicit group. Like House’s study, the other two output-based studies showed similar findings. However, because none of the output-based studies included a control group that did not receive instruction and only one study adopted a delayed posttest, the comparative edge for the explicit condition and its durability need to be confirmed in future research.

Turning to the input-based studies, we found no clear advantage of explicit over implicit conditions. In some cases, the effects of both conditions were negligible (Pearson, 2006; Tateyama, 2001). In other cases, the two conditions were equally effective (Takimoto, 2006a, 2006b, 2008, 2009). Still, some other studies reported mixed findings, with the explicit condition leading to more gains than the implicit condition in one measure but not in the other (Rose & Ng, 2001). Clearly, the difference between explicit and implicit conditions cannot explain these varied findings. Rather, the issue is whether input-based instruction can push learners to process target features beyond the level of noticing. If this criterion is met, both conditions can be equally effective; otherwise, neither can be effective. This observation is supported by comparing Takimoto’s study (2009) with those by Pearson (2006) and

Tateyama (2001).

Takimoto (2009) examined the effects of three kinds of input-based tasks on Japanese EFL learners’ acquisition of request downgraders: explicit metapragmatic instruction followed by structured input activities (EI+SI), implicit structured input activities only (SI), and implicit problem-solving activities only (PS). The outcome measures included a timed listening judgment task, an appropriateness rating task, a DCT, and a role play task. These three input-based tasks were equally effective in facilitating pragmatic acquisition, with the only exception that the EI+SI group did not maintain the gain for the listening judgment task at the delayed posttest while the other two groups did. Here, although the SI and the PS activities were implicit by definition, the tasks forced the learners to process the target features beyond mere noticing. For example, the PS task asked learners to assess contextual variables of request scenarios and to mark down the differences in request forms in the accompanying dialogues. The learners also rated the appropriateness level of the request forms according to scenarios. This PS task thus pushed the learners not only to pay attention to the target features but also to make connections between pragmalinguistic forms (request utterances) and sociopragmatic variables (contextual variables). This kind of processing corresponds to what Schmidt (1993, 2001) termed *understanding* (i.e., recognizing underlying pragmatic rules), which represents a deeper level of processing than *noticing*. Consequently, the learners in the PS group likely developed the target form–function–context mappings even without explicit metapragmatic instruction. Clearly, it was the level of processing afforded by instructional activities, rather than the distinction between explicit and implicit conditions, that could explain Takimoto’s findings.
Different from Takimoto’s study, the explicit and implicit conditions yielded negligible effects in two other input-based studies (Pearson, 2006; Tateyama, 2001). In both cases, the explicit and implicit groups watched videos that contained target features. Before video watching, the explicit group received metapragmatic information but the implicit group did not. Because the video-watching activity merely exposed the target features to the learners without pushing for a deeper level of processing, this activity did not enable the explicit group to reinforce their explicitly learnt metapragmatic knowledge. Consequently, the explicit group, along with the implicit group, showed little gain after instruction. Together, the above three studies (Pearson, 2006; Takimoto, 2009; Tateyama, 2001) suggest that effective input-based instruction needs to push for a deeper level of processing beyond mere noticing.

Like the input-based studies, the dual-modality studies also showed mixed findings regarding the effects of explicit versus implicit conditions. While Nguyen, Pham & Pham (2012) found a clear advantage of explicit over implicit conditions, Martínez-Flor & Fukuya (2005) reported comparable effects of the two conditions. Still, the remaining two studies (Alcón–Soler, 2005; Martínez–Flor & Alcón–Soler, 2007) showed an advantage for the explicit condition only in one, but not in the other, outcome measure. In light of the above discussion, these findings are difficult to explain because the output-based studies demonstrated an advantage of explicit over implicit conditions while the input-based studies showed comparable effects of the two conditions (either effective or ineffective). Therefore, until we know the role of input-based and output-based instruction in pragmatic development, it would be difficult to interpret these seemingly contradictory results.

Our review thus suggests a need to empirically investigate the effects of instructional

modality on L2 pragmatic development. To date, no study has directly explored this topic, though Takahashi’s study (2001), which focused on Japanese EFL learners’ acquisition of request-making forms, is relevant. It included four instructional conditions: an explicit instruction (EI) condition, an implicit form-comparison (FC) condition, an implicit form-search (FS) condition, and an implicit meaning-focused (MF) condition. The explicit EI condition was output-based while the remaining three implicit conditions were input-based. After the instruction, the EI group outperformed the other three groups in a DCT. Although Takahashi’s original goal was to compare the effects of different levels of explicitness, her findings could also be interpreted as showing a possible advantage of output-based instructions over input-based instructions. However, because modality (input-based, output-based) and explicitness (explicit, implicit) were confounded in the design, and because DCT as a production-based outcome measure might bias the results for the output-based modality, the findings need further confirmation.

Like Takahashi’s study, all output-based studies and three of the four dual-modality studies adopted only production task(s) as outcome measure(s); in contrast, most input-based studies used both production and comprehension tasks as outcome measures. Hence, an interesting question is whether differences in outcome measures contributed to the mixed findings among the input-based and output-based studies reviewed above. Because this is an uninvestigated issue in interlanguage pragmatics, one can be informed by studies comparing the effects of input-based and output-based instructions on L2 grammar and vocabulary learning (e.g., Erlam, 2003; Morgan–Short & Bowden, 2006; Qin, 2008; Shintani, 2011; Toth, 2006). In this line of research, the effects of different modalities of instruction were assessed

by production and comprehension tasks. Learners’ performance on both tasks was taken as evidences for or against claims regarding the advantage of one instructional modality over the other. Researchers in L2 pragmatics instruction can learn from this research methodology and incorporate outcome measures tapping both production and comprehension of pragmatic features in order to better understand the effects of instructional modality on pragmatic development.

*Effects of Instruction on the Development of Pragmatic Knowledge and Processing Ability*

Few ILP studies have examined the development of knowledge and processing in instructed environment. Besides House’s (1996) study, Li’s (2012) work is probably the only empirical effort to date. Li’s study was informed by skill acquisition theory (Anderson, 1993; DeKeyser, 2001, 2007c, 2009) and its application to research on L2 grammar instruction (e.g., Byun, 2009; DeKeyser, 1996, 1997; DeKeyser & Sokalsky, 1996). The theory posits that the initial stage of complex cognitive skill development involves the learning of declarative knowledge (knowledge *that*). Because using declarative knowledge requires conscious information retrieval from memory, the resulting performance is typically slow and erroneous. The next stage involves developing procedural knowledge (knowledge *how*) through repeated applications of declarative knowledge to target skills (e.g., comprehension, production). The final stage is an automatization process, in which procedural knowledge is gradually automatized through a large amount of practice to enable automatic processing, which, in turn, enables stable and fluent performance.

According to skill acquisition theory, declarative and procedural knowledge differ in important ways. Declarative knowledge is factual knowledge (e.g., the suffix –*ing* denotes

progressive tense). It is skill-neutral and can be accessed in performing different skills. Practice in one skill domain (e.g., comprehension) can thus contribute to the development of declarative knowledge shared by a different skill domain (e.g., production). On the other hand, procedural knowledge encodes behaviors: It consists of condition-action pairs that specify the actions to be taken once certain conditions are satisfied (e.g., attaching the suffix –ing to verbs when describing continuous actions). The condition-action pairs can only be processed from condition to action and cannot be reversed. This commitment to directionality can promote efficient skill execution, yet it also makes procedural knowledge skill-specific and hard to transfer across different skill domains (DeKeyser, 2007a, 2007b, 2007c).

Therefore, skill-specific practice is needed to develop procedural knowledge associated with different skill domains. To illustrate, linguistic comprehension and production involve different sets of condition-action pairs (i.e., procedural knowledge): whereas comprehension requires analyzing linguistic input in order to interpret the encoded communicative intention, production asks the language user to convert communicative intention into linguistic output. Therefore, for example, practice in comprehension can develop the procedural knowledge associated with comprehension but not with production.

Informed by skill acquisition theory, Li (2012) aimed to understand the effects of differential amount of input-based practice on the development of request-making in L2 Chinese. After a metapragmatic instruction session, an intensive training (IT) group and a regular training (RT) group engaged in input-based practice over two consecutive days. The IT group practiced twice as much as the RT group. Meanwhile, a control group did not practice the forms. The learners’ judgment and production of the target forms were assessed.

by a listening judgment test and an oral production test for accuracy and speed. The results were complex. No group gained significantly in accurate judgment of the target forms. As for the speed (i.e., response times) of judgment, only the IT group made significant improvement over time, yet it did not outperform the other two groups. Concerning the oral production test, the IT and RT groups both made significant gains in accurate production after practice, but only the IT group outperformed the control group. No improvement was found in the speed (i.e., speech rates, planning times) of producing the forms. These results showed that a larger amount of practice generally led to more gains in pragmatic knowledge (as indicated by the accuracy measures) and processing ability (as indicated by the speed measures), and that the development of processing ability necessitates more practice than the development of knowledge.

Li’s study suggests several areas for future investigation. To start with, his findings indicated a possible effect of practice modality on processing ability. This was shown in the IT group’s performance: after input-based practice, the IT group showed a trend towards gain in judgment speed but not in production speed. However, due to its exclusive focus on input-based practice, Li’s study provided limited evidence to support an argument for a modality effect on pragmatic development. To better investigate the issue, the present study simultaneously examined the effects of different modalities of practice (i.e., input-based, output-based). Moreover, Li’s study was limited in the scope of instruction. For one thing, the instruction and practice focused on forms for making request head acts only and did not cover other elements such as internal modifications. In addition, the input-based practice activities emphasized sociopragmatics (e.g., appropriateness judgment according to contexts) and did

not provide opportunities oriented towards pragmalinguistics (e.g., helping learners to better understand the linguistic structures of the target forms). As Li mentioned, this lack of pragmalinguistic training likely led to the negligible effect of practice on gains in judgment accuracy. The present study aimed to address the above issues by incorporating internal modifications into instruction and practice, and by offering practice activities for pragmalinguistic development.

**Summary and Research Question**

Both knowledge and processing are key components of pragmatic competence underlying performance. However, the field of L2 pragmatics instruction has centered on knowledge and overlooked processing ability for the most part. Our review also suggests the need to examine the effects of instructional modality (input-based, output-based) to better understand whether and how instruction can facilitate the development of both knowledge and processing ability. This study addresses these issues and asks the following research question:

**RQ1.** How do different modalities of practice (input-based, output-based) influence the development of accuracy and speed in recognizing and producing request-making forms in L2 Chinese?

**METHODOLOGY**

**Participants**

The participants were 50 American learners of Chinese (39 females, 11 males, mean age $= 20.56$ years, $SD = 1.76$). They were all native English speakers. There were four African Americans, 12 Chinese Americans, 25 Caucasians, three Japanese, and six Koreans. The

Participants were recruited from intermediate Chinese classes in six study abroad programs in China. These programs all emphasized grammar and vocabulary instruction and did not teach the target features. The participants received 15 to 19 hours of formal Chinese instruction each week. Before studying abroad, they had two to four semesters of formal Chinese study.

The participants were randomly assigned to three groups: an input-based practice group (hereafter “input group”), an output-based practice group (hereafter “output group”), and a control group. A standardized Chinese proficiency test adapted from the C. Test (HSK Center, of BLCU, 2009) was administered to check comparability in general Chinese proficiency among the groups, and no difference was found, $F (2, 47) = 0.362, p > .05$. Due to equipment failure, one participant in the control group was excluded from data analysis. Hence, 49 participants remained, with 17 in the input group, 17 in the output group, and 15 in the control group.

**Target Pragmatic Features**

The target features were four Chinese request-making forms (Table 1). They were selected based on a survey given to 20 Chinese university students. The survey results showed that, when making minor requests to good friends, the students tended to employ direct request strategies by using either Form 1 or Form 2 (Table 1). In contrast, when making major requests to professors, they tended to adopt indirect request strategies by using either Form 3 or Form 4. Five lexical downgraders serving as internal modifications were underlined in Table 1.

_______________

Insert Table 1 here

_______________

*Instruction and Practice*

*Metapragmatic Instruction.* A 40-minute computerized metapragmatic instruction session introduced the target features to the participants on Day One. The participants first completed a DCT for assessing their initial knowledge of the target features. The DCT had two situations involving minor requests to friends (i.e., FM situation) and two situations involving major requests to professors (i.e., PM situation). The participants wrote down in Chinese characters or in *Pinyin* (a Chinese transliteration system) what they would say in these situations. Afterwards, the computerized metapragmatic instruction introduced the structure of a request sequence (e.g., head act and internal modification), the concept of direct and indirect requests, the key contextual factors (i.e., power, social distance, and imposition) that influence the choice of request strategies and the target request forms. Finally, the participants completed another DCT (a parallel version of the previous DCT) to confirm their understanding of the targeted features. A comparison of the DCT results showed increased use of the target request head act forms (i.e., from 21.42% to 92.35%) and the target lexical downgraders (i.e., from 9.18% to 90.31%). From Day Two to Day Five, the participants engaged in their respective activities (detailed below).

*Input-based Practice.* The input group engaged in four computerized input-based practice sessions (20-25 minutes each) over four consecutive days. The practice activities of each session were organized as two FM (friend – minor request) and two PM (professor – major

(request) scenarios. Within each request scenario, there was a grammaticality judgment task followed by a dialogue-reading task (see Appendix B for sample activities). During the grammaticality judgment task, the participants first read a request scenario in English and judged the grammaticality of two written request-making forms in Chinese. The participants made their judgments by clicking on the “Yes” or “No” button on the screen. Following their choice(s), explicit feedback on the target forms appeared on the computer screens. During the dialogue-reading task, the participants read the same request scenario and judged the level of imposition of the target request by clicking the “small favor” or the “big favor” button. Feedback regarding the correctness to their choices then popped up. Afterwards, they moved on to the next screen showing a request-making dialogue. In that dialogue, there were two underlined parts where the participants chose the best request utterance out of three options: (1) a pragmatically appropriate and grammatically accurate utterance, (2) a pragmatically appropriate and grammatically inaccurate utterance, and (3) a pragmatically inappropriate and grammatically accurate utterance. The order of these three options was randomized across items. Following the participants’ choice, explicit metapragmatic feedback popped up on the screens. The participants had to make correct choices to move on to the next screen showing the whole dialogue with appropriate and accurate request utterances (underlined and in bold font). Finally, the participants listened to the dialogue twice.

*Output-based Practice*. The output group engaged in four computerized output-based practice sessions (20–35 minutes each) over four consecutive days. Each session contained the same request scenarios as the input-based practice session. Within each scenario, a sentence translation task was followed by a dialogue completion task (see Appendix B for
sample activities). During the sentence translation task, the participants read a request scenario in English and translated two English request sentences into Chinese by using the target forms. They typed their answers in *Pinyin* (the computer program did not recognize Chinese characters as input). The participants then moved on to the next screen showing their own sentences and correct request sentences. During the dialogue completion task, the participants read the same scenario and judged the nature of the request (i.e., a minor or a major request) by clicking the corresponding buttons on the screen. Explicit feedback appeared following their choices, and the participants had to provide the correct answer to move on. The participants then read a dialogue based on the scenario. The dialogue was the same as the one used in the input-based practice, except that there were two blanks instead of two underlined parts. The participants typed a request sentence (in *Pinyin*) into each blank using the target forms. Finally, they read their own responses and the target answers on the next screen.

*The Control Group.* The control group completed four sessions of Chinese reading comprehension exercises (20–30 minutes each). They read short texts and answered comprehension questions.

*Outcome Measures*

A computerized Listening Judgment Test (LJT) and an Oral Discourse Completion Test (ODCT) were used to assess the effects of practice. The LJT measured speedy and accurate recognition of the target request-making forms in contexts. The LJT had 32 items: two practice items, 24 target items, and six distractor items. The 24 target items were evenly divided between FM and PM scenarios. Half of the scenarios came from the practice

materials, while the other half was new. In each LJT item, after a brief vocabulary session, the participants read along while listening to a request scenario in English. They then heard a Chinese request utterance. Immediately afterwards, a beep introduced three options: (a) pragmatically appropriate and grammatically accurate, (b) pragmatically appropriate and grammatically inaccurate, and (c) pragmatically inappropriate and grammatically accurate. The participants chose the one that best described the request utterance they had just heard. The computers recorded their answers and response times. The 24 target request utterances were counterbalanced across the three options. The LJT had three comparable versions used at pre-, post-, and delayed posttest. The following is a sample LJT item.

Request scenario (visual and aural input)

*Li Xiaochen and Professor Chen are attending an academic conference in another city. Li Xiaochen is going to present tomorrow. Unfortunately, Li Xiaochen’s computer broke down. Li Xiaochen knows that Professor Chen brought a computer and would like to borrow it for tomorrow. Li Xiaochen explains the situation and says:*

Request utterance (aural input only)

*Chén lǎo shī, nín kàn wǒ néng yòng nín de diàn nǎo yī xià ma?*

‘Professor Chen, do you think I can use your computer a little bit?’

Options (visual input only)

a. Pragmatically appropriate and grammatically accurate.

b. Pragmatically inappropriate and grammatically accurate.

c. Pragmatically appropriate and grammatically inaccurate.

The Oral Discourse Completion Test (ODCT) measured the ability to produce the target
request-making forms within context. The ODCT had items, including two practice items, 16 target items, and four distractor items. Like the LJT, the target ODCT items were evenly divided between FM and PM scenarios. Half of the scenarios came from the practice materials while the other half was new. Each item started with the same vocabulary session as in the LJT. The participants then read along while hearing a request scenario in English, which ended with a beep. Upon hearing the beep, the participants started to speak into the microphones their requests for that scenario. They were instructed to say the request head acts only. The computers recorded their oral responses. The ODCT had three comparable versions. The following is a sample ODCT item:

Request scenario (visual and aural input)

*It’s a bit hot in the classroom. Li Xiaochen wants to ask Wang Ning, who is sitting close to the window, to open the window. Li Xiaochen explains the situation and says:*

______________.

*Procedures*

All training and assessment sessions were held individually in a lab on campus and were monitored throughout. During Week One, all participants attended the metapragmatic instruction session and took a pretest on Day One. The pretest was administered after the metapragmatic instruction session in order to show the effects of practice only. From Day Two through Day Five, the participants engaged in their respective activities. On Day Five, after the practice, all participants took an immediate posttest. Two weeks after the immediate posttest, they took a delayed posttest. The ODCT was always administered before the LJT.

*Analysis of Data*

The LJT and ODCT data were analyzed along five measures, following Taguchi (2007b) and Li (2012). First, recognition accuracy was operationalized as LJT accuracy scores. Each correct choice received one point and the maximum score was 24 for the test. Second, recognition speed was operationalized as LJT response times, which was calculated by averaging the number of seconds taken to make correct answers. Third, production accuracy was operationalized as ODCT accuracy scores. This score consisted of three sub-scores for request head act frames, lexical downgraders, and overall grammaticality. Regarding request head act frames, two points were awarded for using the target head act frame(s), one point for using non-target yet acceptable head act frames, and zero point for non-target and unacceptable head act frames. As for lexical downgraders, two points were awarded for using the target downgraders, one point for non-target (but appropriate) downgrader, and zero point for no use of downgrader. Concerning overall grammaticality, a grammatical request utterance received one point and an ungrammatical request utterance received zero point. The maximum ODCT accuracy score was 80 for the test (i.e., 5 points per scenario x 16 scenarios). The first author and another Chinese native speaker independently rated 832 (35.37%) utterances, and the ratings were highly correlated (Pearson’s $r = .94$). The remaining 1,520 utterances were rated by the first author. Finally, production speed was operationalized as ODCT planning times (i.e., averaged number of seconds taken to prepare for ODCT responses) and ODCT speech rates (i.e., averaged number of Chinese syllables spoken per minute when producing pragmatically appropriate request utterances, excluding false starts, repetitions, partial repetitions, and repairs).

Due to the small sample size, we adopted non-parametric statistic procedures. To

To investigate changes within each group over time, we conducted separate Friedman tests for each measure mentioned above. Following significant results of the Friedman tests, Wilcoxon tests with the adjusted alpha level of .016 (for three paired comparisons) were performed. To examine group difference at pretest, immediate posttest, and delayed posttest, we performed separate Kruskal-Wallis tests for each measure. Following significant results of the Kruskal-Wallis tests, Mann-Whitney *U* tests were conducted with the adjusted alpha level of .016 (for three paired comparisons) to locate any significant group difference.

**RESULTS**

Table 2 shows the means and standard deviations of the five performance measures: LJT accuracy scores, LJT response times, ODCT accuracy scores, ODCT planning times, and ODCT speech rates.

<table>
<thead>
<tr>
<th>LJT Accuracy Scores</th>
</tr>
</thead>
</table>

In terms of within group comparison over time, the input group improved significantly, $\chi^2 (2, n = 17) = 26.01, p < .001, \eta^2 = .52$. There was a significant increase in accuracy scores from pretest to immediate posttest ($Z = -3.52, p < .001, \eta^2 = .77$), and the gain was maintained from immediate to delayed posttest ($Z = -1.62, p = .12$). No significant improvement was observed for the output group, $\chi^2 (2, n = 17) = 5.12, p = .078$, and for the control group, $\chi^2 (2, n = 15) = 4.54, p = .107$. Concerning between-group comparisons, there was no significant difference at pretest, $\chi^2 (2, N = 49) = 0.29, p = .864$. However, the three
groups differed significantly at immediate posttest, $\chi^2(2, N = 49) = 22.98, p < .001, \eta^2 = .48$, with the input group outperforming the output group ($Z = -3.42, p = .001, \eta^2 = .35$) and the control group ($Z = -4.29, p < .001, \eta^2 = .59$). There was no difference between the output group and the control group ($Z = -2.13, p = .033$). At delayed posttest, the groups again showed significant difference, $\chi^2(2, N = 49) = 19.76, p < .001, \eta^2 = .41$. That is, the input group scored higher than the output group ($Z = -3.54, p < .001, \eta^2 = .38$) as well as the control group ($Z = -4.02, p < .001, \eta^2 = .52$). The difference between the output and control groups was not significant ($Z = -0.86, p = .403$).

Because the input group outperformed the output group at immediate and delayed posttests, a post hoc error analysis was performed to see in what way the input group was better than the output group. Recall that the participants in both groups had to judge three types of request utterances in the LJT. Table 3 presents the mean scores and accuracy rates of judgment. One noticeable difference between the input and output groups was in their ability to recognize Option B type request utterances (i.e., pragmatically appropriate and grammatically inaccurate): While the input group showed considerable improvement from pretest to immediate posttest (i.e., the accuracy rates increased from 40.44% to 69.85%), the output group only made negligible progress (i.e., the accuracy rates increased from 40.44% to 44.11%). Another notable group difference was in their ability to recognize Option C type request utterances (i.e., pragmatically inappropriate and grammatically accurate). Both groups started out with identical accuracy rates, yet at immediate posttest the input group scored higher (86.03%) than the output group (75.74%). The gap between the two groups further enlarged at delayed posttest (90.44% vs. 71.32%). Regarding the ability to recognize
Option A type request utterances, however, the degree of improvement was comparable between the two groups.

**LJT Response Times**

In terms of within group comparisons over time, the input group significantly reduced response times, $\chi^2 (2, n = 17) = 7.18, p = .028, \eta^2 = .14$. Specifically, a significant difference was found between pretest and immediate posttest ($Z = -2.63, p = .007, \eta^2 = .43$), but not between immediate and delayed posttests ($Z = -0.40, p = .701$). The output group did not show significant reduction of response times, $\chi^2 (2, n = 17) = 5.76, p = .063$, nor did the control group, $\chi^2 (2, n = 15) = 4.13, p = .135$. On the other hand, there was no group difference at pretest ($\chi^2 (2, N = 49) = 0.03, p = .987$), immediate posttest ($\chi^2 (2, N = 49) = 1.05, p = .590$), and delayed posttest ($\chi^2 (2, N = 49) = 0.89, p = .640$).

**ODCT Accuracy Scores**

Concerning within group comparisons over time, the input group gained significantly, $\chi^2 (2, n = 17) = 18.58, p < .001, \eta^2 = .37$. A significant increase in accuracy scores was found between pretest and immediate posttest ($Z = -3.62, p < .001, \eta^2 = .82$), and the gain was maintained from immediate to delayed posttests ($Z = -1.64, p = .107$). The output group also gained over time, $\chi^2 (2, n = 17) = 18.12, p < .001, \eta^2 = .36$, with a significant difference found between pretest and immediate posttest ($Z = -3.48, p < .001, \eta^2 = .76$), but not between immediate and delayed posttests ($Z = -0.97, p = .349$). The control group did not improve, $\chi^2 (2, n = 15) = 0.74, p = .711$. Regarding between-group comparisons, there was no significant difference at pretest, $\chi^2 (2, N = 49) = 0.66, p = .718$. However, the three groups showed significant difference at immediate posttest, $\chi^2 (2, N = 49) = 20.76, p < .001, \eta^2 = .43$. That is,

the input group outperformed the control group ($Z = -3.66, p < .001, \eta^2 = .43$), and the output group also outperformed the control group ($Z = -4.10, p < .001, \eta^2 = .54$). There was no significant difference between the input and output groups ($Z = -1.14, p = .261$). At delayed posttest, there was again a significant difference between the groups, $\chi^2 (2, N = 49) = 12.80, p = .002, \eta^2 = .27$. This time, the only significant difference was found between the output and control groups ($Z = -3.14, p = .001, \eta^2 = .32$). There was no significant difference between the input and output groups ($Z = -2.29, p = .022$).

**ODCT Planning Times**

Regarding within group comparisons, the input group did not show significant reduction of planning times ($\chi^2 (2, n = 17) = 4.59, p = .105$), nor did the control group ($\chi^2 (2, n = 15) = 4.93, p = .096$). However, the output group significantly shortened their planning times, $\chi^2 (2, n = 17) = 23.06, p < .001, \eta^2 = .46$, with a significant difference found between pretest and immediate posttest ($Z = -3.53, p < .001, \eta^2 = .78$), but not between immediate and delayed posttests ($Z = -1.59, p = .117$). Comparisons between the three groups showed no significant difference at pretest ($\chi^2 (2, N = 49) = 2.46, p = .293$), immediate posttest ($\chi^2 (2, N = 49) = 2.73, p = .255$), and delayed posttest ($\chi^2 (2, N = 49) = 0.25, p = .885$).

**ODCT Speech Rates**

In terms of within group comparisons, the input group made significant improvement overtime, $\chi^2 (2, n = 17) = 13.06, p = .001, \eta^2 = .26$. However, follow-up comparisons did not show significant difference between pretest and immediate posttest ($Z = -1.68, p = .098$), or between immediate and delayed posttests ($Z = -1.92, p = .057$). The only significant

difference was found between pretest and delayed posttest ($Z = -2.96, p = .002, \eta^2 = .55$).

The output group also demonstrated significant improvement, $\chi^2 (2, n = 17) = 19.88, p < .001, \eta^2 = .39$. Their performance on immediate posttest was better than that on the pretest ($Z = -2.68, p = .006, \eta^2 = .45$), and additional gain was observed from immediate to delayed posttests ($Z = -2.86, p = .003, \eta^2 = .51$). Finally, the control group also showed significant gains over time, $\chi^2 (2, n = 15) = 12.93, p = .001, \eta^2 = .29$, with a significant improvement observed from pretest to immediate posttest ($Z = -3.18, p < .001, \eta^2 = .72$), but not from immediate to delayed posttests ($Z = -1.53, p = .135$). On the other hand, there was no significant difference between the three groups at pretest ($\chi^2 (2, N = 49) = 0.52, p = .773$), immediate posttest ($\chi^2 (2, N = 49) = 0.60, p = .739$), and delayed posttest ($\chi^2 (2, N = 49) = 0.58, p = .750$).

DISCUSSION

To discuss the effects of practice modality on development in accuracy (an indicator of pragmatic knowledge) and speed (an indicator of processing ability) of pragmatic performance, we examined the results of the LJT and the ODCT separately before reviewing the findings together.

Regarding LJT accuracy scores, the input group made significant gains from pretest to immediate posttest, maintained the gains at delayed posttest, and outperformed the output group and the control group at immediate and delayed posttests. The output group, however, did not show significant improvement. The input group thus demonstrated a stronger effect of practice than the output group in recognition accuracy. These findings were expected because the input group had opportunities to practice judging request utterances belonging to different...

Levels of appropriateness (e.g., appropriate, inappropriate) and grammaticality (grammatical, ungrammatical). In contrast, the output group did not have many opportunities to encounter inappropriate or ungrammatical utterances during output-based practice. The output group was thus less prepared for judging different types of request utterances in the LJT, especially for Option B type (i.e. pragmatically appropriate and grammatically inaccurate) and Option C type (i.e. pragmatically inappropriate and grammatical accurate). The results of the error analysis (Table 3) supported this interpretation: compared with the input group, the output group made only modest improvement from pretest to immediate posttest in recognizing Option C type utterances and negligible improvement in recognizing Option B type utterances. Regarding the long-term effects of practice on recognition accuracy, the input group retained their gains from immediate to delayed posttest. It also outperformed the output group at delayed posttest. Hence, there was an overall edge that input-based practice held over output-based practice in promoting recognition accuracy.

The input group’s gain in recognition accuracy contrasted with the IT group’s non-improvement reported in Li’s (2012) study (cited earlier). Because the outcome measure (i.e., the listening judgment task), the target features (i.e., Chinese request-making forms), and the length of instruction/practice were comparable across these two studies, the difference in recognition accuracy was attributable to the opportunities for practicing pragmalinguistic forms (e.g., the grammaticality judgment task, the dialogue reading task) during the input-based practice in this study. Even though the practice activities were input-based in both studies, the results differed because of the nuances in what the learners actually did.
The output group in this study did not show overall gain in recognition accuracy. However, they improved notably in recognizing Option A type utterances (pragmatically appropriate and grammatically accurate) from pretest to immediate posttest (Table 3). In fact, another post hoc analysis showed that they improved significantly in this respect, $\chi^2(2, n = 17) = 10.51, p = .004$. There was a significant difference between pretest and immediate posttest ($Z = -2.68, p = .007$), but not between immediate and delayed posttests ($Z = -0.54, p = .781$). The output-based practice, by asking the participants to repeatedly produce request utterances, likely refined the target form–function–context mappings to some degree. Consequently, the output group was able to do better in recognizing request utterances that were appropriate and accurate. However, this practice did not allow better recognition of the other two types of request utterances (i.e., Option B and C types).

Compared with the gains in recognition accuracy, the gains in recognition speed (i.e., LJT response times) were very limited. The input group made significant gains from pretest to immediate posttest and maintained the gains at delayed posttest. While this result pointed to an effect of input-based practice on promoting recognition speed, the effect was rather weak because the input group did not outperform the control group. The output group, however, did not show significant gain in recognition speed. This result indicated that output-based practice was not effective for enhancing recognition speed. Collectively, these findings showed a trend suggesting that input-based practice was more effective than output-based practice in promoting recognition speed.

Turning to the ODCT results, the input and output groups both made significant improvement in ODCT accuracy from pretest to immediate posttest and maintained the gains

at delayed posttest. Both groups also outperformed the control group at immediate posttest. However, while the output group outperformed the control group at delayed posttest, the input group did not. These results showed that the effects of input-based and output-based practice were comparable in terms of immediate gains, but for the retention of gains, output-based practice was more beneficial than input-based practice.

The above findings can be explained by the difference between the input-based and output-based practice activities. In producing request utterances, the output group must engage in exact word-for-word analysis of the target forms. The input-based practice did not require such precise linguistic analysis. For example, during the dialogue-reading task, the participants could rely on a few key phrases to help make their choices. To illustrate, the phrase nín kàn (‘you see’, with a respectful second person pronoun) could help determine that the entire request utterance was appropriate for a “professor – major request” scenario. Hence, the output-based practice likely forced the learners to engage in a deeper level of analyzing the target forms but the input-based practice did not.

In addition, the feedback given to the two groups during practice might also lead to the difference in the depth of linguistic analysis. The output group received correct request utterances next to their self-generated ones on the computer screen. This arrangement probably helped the learners to “notice the gap”, which might lead to further metapragmatic reflections on the target forms (e.g., Swain, 1985, 1995). In contrast, the input group was not likely to perform such metapragmatic analyses because the input-based practice did not ask them to produce request utterances. Hence, the feedback during the output-based practice might enable a deeper level of analysis of the target forms than the feedback during the
input-based practice. This could explain why the output group retained their gains at delayed posttest but the input group did not. Interestingly, however, the input and output groups did not show significant difference at any point in time. Since the time interval between the practice phase and both posttests was relatively short (i.e., two weeks), the effects of processing depth on production accuracy was probably not strong for the output group (relative to the input group). Future research is needed to further investigate this effect of processing depth on retention of gains in pragmatic knowledge.

The difference between the input and output groups was smaller in production speed (i.e., ODCT planning times, speech rates). For both speed measures, there was no significant difference between the three groups at any time point. However, the results showed that the output group improved significantly over time while the input group did not. These findings pointed to a trend showing that the output-based practice was more effective than the input-based practice in developing processing ability associated with the ODCT.

Somewhat unexpectedly, the control group also made significant gains in speech rates from pretest to immediate posttest. To understand this result, it is helpful to consider the measure of ODCT accuracy along with the speech rates measure. For the output group, the increase in speech rates from pretest to immediate posttest was associated with greater production accuracy. This meant that they were in the process of incorporating new pragmatic knowledge into their interlanguage system while becoming more efficient in processing this knowledge in oral production. In contrast, the control group improved in speech rates but not in production accuracy. This increase in speech rates was likely a result of repeating similar production tasks over time (i.e., the ODCT). Hence, the control group’s gain in speech rates
should not undermine the effectiveness of output-based practice in enhancing processing ability. Rather, it is important to consider the nature of increased production speed before taking it as an indicator for the development of processing ability. As Bialystok (1993) theorized, although knowledge and processing are distinct cognitive processes underlying L2 performance, they are connected and prioritized from a developmental perspective, that is, the development of processing ability presupposes the development of knowledge. This is because processing ability, understood as the efficient control of attention, becomes a vacuous construct without relevant knowledge, i.e., the target linguistic and non-linguistic information to which one’s attentional resources are allocated. With this understanding, the gain in speech rates of the output group was a valid indication of development in processing ability.

Two observations emerged after comparing the input and output groups across the accuracy and speed measures. First, for both groups, the effects of practice were always more prominent for the development of knowledge (as indicated by the accuracy measures) than for the development of processing ability (as indicated by the speed measures). For example, the input group demonstrated a strong effect of practice on measures of recognition accuracy but a rather weak effect on recognition speed; the group also made significant gains in production accuracy but not in production speed. Likewise, the output group showed a strong effect of practice on production accuracy but merely a weak effect on production speed; meanwhile, the group made significant gains in recognition accuracy (for Option A type request utterances) but not in recognition speed. Second, the development of pragmatic knowledge benefited from practice regardless of modality type, yet the development of

processing ability seemed to benefit only from modality-specific practice. To illustrate, the input and output groups both showed significant improvement in recognition and production accuracy, and there was no significant difference between the two groups. However, the two groups showed different patterns of development in performance speed. The input group demonstrated a trend towards an effect on recognition speed but not on production speed. Similarly, the output group showed a trend towards an effect on production speed but not on recognition speed. Together, these two observations suggest that: (1) pragmatic knowledge is more amenable to practice than processing ability, and (2) pragmatic knowledge can be developed regardless of practice modality, but the development of processing ability requires modality-specific practice.

The above two observations can be explained by skill acquisition theory (DeKeyser, 2007a, 2007b, 2007c, 2009). In this study, pragmatic knowledge, as indicated by the accuracy measures, can be considered as a kind of declarative knowledge because it deals with the rules governing the relationship between target forms, functions, and contexts. On the other hand, processing ability, as indicated by the speed measures, can be considered as a kind of procedural knowledge because it involves the efficiency of access to pragmatic knowledge in completing various tasks. In this study, the input-based and output-based practice activities provided opportunities for developing procedural knowledge associated with the recognition and production tasks. For instance, the output-based practice repeatedly asked the learners first to assess contextual factors and then to produce request utterances based on the target forms. Because these procedures resembled those involved in responding to the production task (i.e., ODCT), the procedural knowledge developed through output-based practice can be 30

transferred to performing the production task. However, because the amount of practice was small in both practice conditions (i.e., eight instances for practicing each mapping), the procedural knowledge was probably not well developed enough to enable automatic processing. Hence, it was likely that the effect of practice on performance speed was rather limited. On the other hand, the learners’ declarative knowledge might have been refined and strengthened through repeated practice in input-/output-based activities, thereby enhancing performance accuracy. This explains the first observation that has already been mentioned: The gain in pragmatic knowledge was more prominent than the gain in processing ability. We can perhaps further argue that pragmatic knowledge can be refined to a significantly higher level with a relatively small amount of practice while processing ability requires a larger amount of practice to develop. This finding echoes the results reported in Li’s (2012) study discussed earlier. The present study further suggests that, irrespective of practice modality, pragmatic knowledge is more amenable to practice than the ability for processing this knowledge.

The second observation (i.e., the development of pragmatic knowledge can benefit from practice across modalities while the development of processing ability requires modality-specific practice) can be accounted for by the characteristics of declarative and procedural knowledge and how these two types of knowledge are developed. By referring to the skill acquisition theory (DeKeyser, 2007a, 2007b), our literature review section has explained that procedural knowledge is committed to a specific skill domain (e.g., comprehension, production) because it is developed through practice in that skill, and that declarative knowledge is shared across skills and thus can be developed through practice in a
different skill domain. In other words, the effects of practice are skill-specific at the level of procedural knowledge, but at the level of declarative knowledge, the effects of practice can be transferred across skill domains. In this study, the procedural knowledge associated with the recognition and production tasks was developed separately through input-based and output-based practice activities. And this can explain the lack of effect of input-based practice on the development of processing ability associated with the production task (i.e., ODCT), as well as the lack of effect of output-based practice on the development of processing ability associated with the recognition task (i.e., LJT). On the other hand, the input-based and output-based practice both contributed to the development of the declarative knowledge that can be accessed during recognition and production tasks, and this can account for the development of pragmatic knowledge regardless of practice modality.

In summary, the present study showed that, in an explicit instructional condition, input-based and output-based practice differentially facilitated the gains in pragmatic knowledge and processing ability. While instructional ILP research has centered on comparing the effects of explicit versus implicit conditions, this study indicates that modality of instruction merits independent empirical attention. Specifically, our findings suggest that the development of processing ability likely necessitates skill-specific practice, but the development of pragmatic knowledge can benefit from practice across modalities.

REFLECTIONS AND FUTURE RESEARCH

This study aimed to understand the effects of different practice modalities (i.e., input-based, output-based) on the development of pragmatic knowledge and processing ability. Relying on computer technology, we opted for a laboratory-based approach (rather

than a classroom-based approach) in order to better control extraneous variables. Yet as one reviewer insightfully pointed out, the variables of input-based and output-based practice modalities were probably not as univalent as we conceived them to be. In fact, as our discussion shows, the effects of a particular type of practice depended on what the learners actually did and how they were assessed. For example, although the input and output groups both practiced pragmalinguistic structures, the specific setups of the two practice conditions (i.e., whether the learners encountered any “negative evidence” in judging request utterances) led to differences in performance accuracy as assessed by the LJT and the ODCT. In this article, we tried to discuss such effects in detail. In so doing, we hoped to encourage more attention to the specifics in instructional treatment and assessment for interpreting existing findings and for designing future instructional ILP studies.

The findings of this study suggest several issues for future research. To start with, due to the small amount of practice administered over a short period of time (i.e., four days), the effects of practice on the development of processing ability were rather weak. A follow-up study can explore the amount of practice needed for the development of processing ability. One option could be to increase the amount of practice and to track learners over an extended period of time. However, one should be cautious with this approach because the amount of practice needed for developing processing ability (and pragmatic knowledge) may differ across pragmatic features. Although a few studies included multiple pragmatic features as instructional targets (e.g., Billmyer, 1990; Eslami & Eslami-Rasekh, 2008; Eslami-Rasekh, Eslami-Rasekh & Fatahi, 2004; House, 1996; Rose & Ng, 2001), very little is known about whether and how different pragmatic features interact with instruction in affecting the

development of pragmatic knowledge, let along processing ability. Future research on this topic is clearly in order.

In addition, the input-based and output-based practice activities in this study both belonged to the explicit instructional condition. Given the mixed findings regarding the effects of explicit and implicit conditions across different instructional modalities, it would be interesting to investigate the modality effect in implicit conditions. Findings of such studies would enable a better understanding of the existing findings concerning explicit vs. implicit conditions.

Finally, this study examined the effects of two types of pragmatics practice, i.e., input-based and output-based. This line of research can be further refined by investigating the effects of task repetition within input-based and/or output-based practice mode. For example, within the context of input-based instruction, Takimoto (2012) recently compared the effects of identical task repetition and task-type repetition on the acquisition of request downgraders. He found that identical task repetition demonstrated a superior effect than task-type repetition as reflected in both recognition and production of the target features. Under Takimoto’s definition, the practice activities adopted in this study belonged to task-type repetition (i.e., the participants encountered different but comparable request scenarios during practice), and it would be interesting to incorporate a condition of identical task repetition into future research.

NOTES

1. We excluded studies that did not contain a pretest and studies that focused on constructs other than pragmatic knowledge (e.g., confidence level in performing pragmatics tasks). Takahashi’s (2001) study does not appear in the review chart available online because her explicit condition consisted of output-based activities while her implicit conditions consisted of input-based activities. Takahashi’s study is reviewed separately in this article.

2. One reviewer suggested using the term “production practice” to label this instructional condition because the term “output” usually implies speaking accompanied by interaction. While we acknowledge the reviewer’s point, the term “output-based practice” was retained as a parallel term of “input-based practice”. Moreover, because the input-based practice condition of this study was a refined version of Li’s (2012), we chose to use this term to ensure consistency in terminology.

3. During the mini vocabulary lesson, the participants listened to a few Chinese words twice and read these words (in characters and in *Pinyin*, with English translations) on computer screens.

4. One example of “non-target yet acceptable head act frames” is using the form “kě yǐ + verb phrase + ma?” (May + verb phrase + particle), instead of the form “néng + verb phrase + ma?” (Can + verb phrase + particle) when making a major request to professors.

5. Another reason for choosing non-parametric statistic procedures was to allow better comparison of results with Li (2012), the study that directly motivated the present one, because that study also adopted the same non-parametric statistic procedures.

REFERENCES


49–84.


38


Taguchi, N. (2008). The role of learning environment in the development of pragmatic


Takimoto, M. (2012). Assessing the effects of identical task repetition and task-type repetition

on learners’ recognition and production of second language request downgraders.


**APPENDIX A**

**Review chart of 15 Studies Comparing Explicit and Implicit Instructional Conditions**

<table>
<thead>
<tr>
<th>Study</th>
<th>Group</th>
<th>Target feature</th>
<th>Design</th>
<th>Outcome measure</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Output-based Instructional Studies</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Félix-Brasdefer</td>
<td>Explicit Implicit</td>
<td>Spanish refusal</td>
<td>Pretest</td>
<td>Role play</td>
<td>Explicit &gt; implicit (except for one measure)</td>
</tr>
<tr>
<td>(2008)</td>
<td></td>
<td></td>
<td>Posttest</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Delayed posttest</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ghobadi &amp; Fahim</td>
<td>Explicit Implicit</td>
<td>English Thanking</td>
<td>Pretest</td>
<td>DCT</td>
<td>Explicit &gt; implicit</td>
</tr>
<tr>
<td>(2009)</td>
<td></td>
<td></td>
<td>Posttest</td>
<td></td>
<td>Explicit &gt; implicit</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Role play</td>
<td></td>
</tr>
<tr>
<td>House (1996)</td>
<td>Explicit Implicit</td>
<td>English gambit; discourse strategy;</td>
<td>Pretest</td>
<td>Role play</td>
<td>Explicit &gt; implicit</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Opening &amp; closing phrase</td>
<td>Mid-test</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Posttest</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Input-based Instructional Studies</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Koike &amp; Pearson</td>
<td>1. Explicit instruction + explicit feedback (EI+EF)</td>
<td>Spanish suggestion</td>
<td>Pretest</td>
<td>Multiple-choice questionnaire</td>
<td>Instructed &gt; control (Overall effect of instruction on posttest; effects not maintained on delayed posttest)</td>
</tr>
<tr>
<td>(2005)</td>
<td>2. Explicit instruction + implicit feedback (EI+IF)</td>
<td></td>
<td>Posttest</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>3. Implicit instruction + explicit feedback (II+EF)</td>
<td></td>
<td>Delayed posttest</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>4. Implicit instruction + implicit feedback (II+IF)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>5. Control</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study</td>
<td>Group</td>
<td>Target feature</td>
<td>Design</td>
<td>Outcome measure</td>
<td>Results</td>
</tr>
<tr>
<td>-------</td>
<td>-------</td>
<td>----------------</td>
<td>--------</td>
<td>-----------------</td>
<td>---------</td>
</tr>
<tr>
<td><strong>Pearson (2006)</strong></td>
<td>Explicit</td>
<td>Spanish directive</td>
<td>Pretest, Posttest, Delayed posttest</td>
<td>DCT</td>
<td>Explicit = implicit = control (Qualitative analysis: some effects of both instructional conditions)</td>
</tr>
<tr>
<td></td>
<td>Implicit</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Control</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Rose &amp; Ng (2001)</strong></td>
<td>Explicit</td>
<td>English compliment &amp; Compliment response</td>
<td>Pretest, Posttest, Delayed posttest</td>
<td>DCT</td>
<td>Explicit &gt; implicit</td>
</tr>
<tr>
<td></td>
<td>Implicit</td>
<td></td>
<td></td>
<td></td>
<td>Explicit=implicit (no effect)</td>
</tr>
<tr>
<td></td>
<td>Control</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Takimoto (2006a, 2006b)</strong></td>
<td>1. Structured input (implicit)</td>
<td>English request downgraders</td>
<td>Pretest, Posttest, Delayed posttest</td>
<td>DCT</td>
<td>Explicit = implicit &gt; control</td>
</tr>
<tr>
<td></td>
<td>2. Structured input + explicit feedback (explicit)</td>
<td></td>
<td></td>
<td>Role play</td>
<td>Explicit = implicit &gt; control</td>
</tr>
<tr>
<td></td>
<td>3. Control</td>
<td></td>
<td></td>
<td>Listening judgment</td>
<td>Explicit = implicit &gt; control</td>
</tr>
<tr>
<td></td>
<td>4. Control</td>
<td></td>
<td></td>
<td>Rating task</td>
<td>Explicit = implicit &gt; control</td>
</tr>
<tr>
<td></td>
<td>2. Problem-solving (implicit)</td>
<td></td>
<td></td>
<td>Role play</td>
<td>Explicit = implicit &gt; control</td>
</tr>
<tr>
<td></td>
<td>3. Structured input (implicit)</td>
<td></td>
<td></td>
<td>Timed listening judgment</td>
<td>Explicit = implicit &gt; control</td>
</tr>
<tr>
<td></td>
<td>4. Control</td>
<td></td>
<td></td>
<td>Rating task</td>
<td>(Explicit group did not maintain gain on delayed posttest, yet still outperformed control group)</td>
</tr>
<tr>
<td></td>
<td>4. Control</td>
<td></td>
<td></td>
<td></td>
<td>Explicit = implicit &gt; control</td>
</tr>
<tr>
<td><strong>Tatemaya (2001)</strong></td>
<td>Explicit</td>
<td>Japanese routine</td>
<td>Pretest, Posttest</td>
<td>Role play</td>
<td>Explicit = implicit (no effect)</td>
</tr>
<tr>
<td></td>
<td>Implicit</td>
<td></td>
<td></td>
<td>Multiple-choice questionnaire</td>
<td>Explicit = implicit (no effect) (Qualitative analysis: Explicit &gt; implicit)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Study</th>
<th>Group</th>
<th>Target feature</th>
<th>Design</th>
<th>Outcome measure</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Dual-modality Instructional Studies</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Alcón–Soler (2005)</td>
<td>Explicit</td>
<td>English requests</td>
<td>Pretest</td>
<td>Metapragmatic questionnaire, Written role play</td>
<td>Explicit = implicit &gt; control</td>
</tr>
<tr>
<td></td>
<td>Implicit</td>
<td></td>
<td>Posttest</td>
<td></td>
<td>Explicit &gt; implicit &gt; control</td>
</tr>
<tr>
<td>Martínez–Flor &amp; Fukuya (2005)</td>
<td>Explicit</td>
<td>English suggestion</td>
<td>Pretest</td>
<td>E-mail task, Telephone task</td>
<td>Explicit = implicit &gt; control</td>
</tr>
<tr>
<td></td>
<td>Implicit</td>
<td></td>
<td>Posttest</td>
<td></td>
<td>Explicit = implicit &gt; control</td>
</tr>
<tr>
<td></td>
<td>Implicit</td>
<td></td>
<td>Posttest</td>
<td></td>
<td>Explicit &gt; implicit &gt; control</td>
</tr>
<tr>
<td>Nguyen, Pham, Pham (2012)</td>
<td>Explicit</td>
<td>English criticism</td>
<td>Pretest</td>
<td>DCT, Role play, Oral feedback</td>
<td>Explicit &gt; implicit &gt; control</td>
</tr>
<tr>
<td></td>
<td>Implicit</td>
<td></td>
<td>Posttest</td>
<td></td>
<td>Explicit &gt; implicit &gt; control</td>
</tr>
<tr>
<td></td>
<td>Control</td>
<td></td>
<td>Delayed posttest</td>
<td></td>
<td>Explicit &gt; implicit &gt; control</td>
</tr>
</tbody>
</table>

*Note.* > refers stronger instructional effects; = refers to equal instructional effect.
APPENDIX B

Sample Scenario (for both Input-based and Output-based Activities)

Li Xiaochen and Wang Ning take the same computer course. The professor of the course sent out an assignment via e-mail but Li Xiaochen lost the e-mail. Wang Ning still has the e-mail, so Li Xiaochen wants to ask Wang Ning to send it to him/her.

A. Sample input-based practice activities

Sample grammatical correctness judgment task

(The English translations were not available to the participants)

a. 给我发一下电子邮件吧。(Meaning: Send me the e-mail. Using Form 1 as listed in Table 1, this sentence was grammatically correct.)

b. 把电子邮件给我发。(Meaning: Send me the e-mail. Using Form 2 as listed in Table 1, this sentence was grammatically incorrect.)

Sample dialogue reading task

Lǐ Xiǎochén : Wáng Níng, chén lǎoshī zuó tiān bùzhí le zuòyè, duì ba?
李晓晨：王宁，陈老师昨天布置了作业，对吧？

Wáng Níng : duì , tā gěi wǒmen fā le diànzǐ yóujiàn.
王宁：对，他给我们发了电子邮件。

Lǐ Xiǎochén : Wǒ debú jiàn le. nǐ (a) néng zài gěi wǒ fā nà ge diànzǐ yǒujiàn ma?
李晓晨：我不见了。你(a)能再给我那个电子邮件吗？

(b) bǎ nà ge diànzǐ yǒujiàn zài gěi wǒ fā.

(b) 把那个电子邮件再给我发。

(c) 把那个电子邮件再给我发一下吧。

Wáng Níng:  tā  gěi wǒmen fā le liàng ge diànzǐ yóujiàn . nǐ yào nǎ yī ge?

王宁: 他给我们的发了两个电子邮件。你要哪一个?

Lǐ Xiǎochén : (a) gěi wǒ fā yīxià zuótiān xiàwǔ nà ge ba .

李晓晨: (a) 给我发一下昨天下午那个。

(b) néng bù néng gěi wǒ fā zuótiān xiàwǔ nà ge ?

(b) 能不能给我发昨天下午那个?

(c) gěi wǒ fā zuótiān xiàwǔ nà ge yīxià ba .

(c) 给我发昨天下午那个一下吧。

Wáng Níng:  hǎo , wǒ huí sùshè yǐhòu gěi nǐ fā.

王宁: 好, 我 回 宿舍 以后 给 你发。

English translation of the dialogue (not available to the participants)

Li Xiaochen: Wang Ning, Professor Chen assigned homework yesterday, right?

Wang Ning:  Yes. He sent us an e-mail.

Li Xiaochen: Mine was lost. (a) Can you forward that e-mail to me? (b) Forward that e-mail to me (there was a grammar error in the Chinese version of this sentence). (c) Forward that e-mail to me a bit.

Wang Ning:  He sent us to e-mails. Which one do you want?

Li Xiaochen: (a) Forward to me the one sent to us yesterday afternoon. (b) Can

you forward to me the one sent to us yesterday afternoon?

(c) Forward to me the one sent to us yesterday afternoon (there was a grammar error in the Chinese version of this sentence).

Wang Ning: OK. I will forward (it) to you after I get back to my dorm.

---

**B. Sample output-based practice**

**Sample sentence translation task**

(The participants were instructed to use both Form 1 and Form 2 as listed in Table 1 to translate the following English request utterances).

a. Send me that e-mail.

b. Send that e-mail to me.

**Sample dialogue completion task**

(The participants were instructed to use both Form 1 and Form 2 as listed in Table 1 to fill in the blanks in the following dialogue).

Lǐ Xiǎochén: Wáng Níng, chén lǎoshī zuó tiān bùzhì le zuòyè, duì ba?

李晓晨：王宁，陈老师昨天布置了作业，对吧?

Wáng Níng: dui, tā gěi wǒmen fā le diànzǐ yóujìàn .

王宁：对，他给我们发了电子邮件。

Lǐ Xiǎochén: Wǒ de bú jiàn le. nǐ

李晓晨：我的不见了。你______________________。

Wáng Níng: tā gěi wǒmen fā le liǎng ge diànzǐ yóujìàn . nǐ yào nǎ yī ge?

王宁：他给我们发了两个电子邮件。你要哪一个？

Li Xiaochen:

李晓晨：

Wang Ning: 好，我回宿舍以后给你发。

English translation of the dialogue (not available to the participants)

Li Xiaochen: Wang Ning, Professor Chen assigned homework yesterday, right?

Wang Ning: Yes. He sent us an e-mail.

Li Xiaochen: Mine was lost. ______________________________.__

Wang Ning: He sent us to e-mails. Which one do you want?

Li Xiaochen: ______________________________.

Wang Ning: OK. I will forward (it) to you after I get back to my dorm.

**TABLE 1**

Target Form-Function-Context Mappings

<table>
<thead>
<tr>
<th>Request-making forms (head act frames)</th>
<th>Function</th>
<th>Context</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. (bāng máng / bāng wǒ) + verb + yī xià + (object)</td>
<td>Direct</td>
<td>Making</td>
</tr>
<tr>
<td>+ ba *</td>
<td>request with minor</td>
<td></td>
</tr>
<tr>
<td>(help / help me) + verb + a little bit + object +</td>
<td>mitigated</td>
<td>requests to</td>
</tr>
<tr>
<td><em>Particle</em></td>
<td>tone</td>
<td>good</td>
</tr>
<tr>
<td>2. (bāng máng / bāng wǒ) + bǎ +object + verb</td>
<td>friends (FM situation)</td>
<td></td>
</tr>
<tr>
<td>+ yī xià ba</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(help/help me) + prep. + object + verb + a little bit +</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>particle</em></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. nǐn kàn + (subject) + néng + verb + yīxià+ object</td>
<td>Indirect</td>
<td>Making</td>
</tr>
<tr>
<td>+ ma?</td>
<td>request with major</td>
<td></td>
</tr>
<tr>
<td>You see + (subject) + can + verb + a little bit +</td>
<td>mitigated</td>
<td>requests to a</td>
</tr>
<tr>
<td><em>object + particle</em>?</td>
<td>tone</td>
<td>professor</td>
</tr>
<tr>
<td>4. nǐn kàn + (subject) + néng bù néng + verb +</td>
<td>that one</td>
<td></td>
</tr>
<tr>
<td>yī xià + object?</td>
<td>knows well</td>
<td></td>
</tr>
<tr>
<td>You see + (subject) + can or cannot + verb + a</td>
<td>(PM situation)</td>
<td></td>
</tr>
<tr>
<td>little bit + object?</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Note. * The components in the parentheses are optional.

**TABLE 2**

Means and Standard Deviations of the Five Measures

<table>
<thead>
<tr>
<th>Measure</th>
<th>Group</th>
<th>Pretest</th>
<th>Immediate posttest</th>
<th>Delayed posttest</th>
</tr>
</thead>
<tbody>
<tr>
<td>LJT Accuracy</td>
<td>Input (n=17)</td>
<td>Mean</td>
<td>14.05</td>
<td>20.11</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SD</td>
<td>4.09</td>
<td>1.99</td>
</tr>
<tr>
<td></td>
<td>Output (n=17)</td>
<td>Mean</td>
<td>14.64</td>
<td>16.94</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SD</td>
<td>4.75</td>
<td>2.77</td>
</tr>
<tr>
<td></td>
<td>Control (n=15)</td>
<td>Mean</td>
<td>14.60</td>
<td>14.46</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SD</td>
<td>3.66</td>
<td>3.48</td>
</tr>
<tr>
<td>LJT Response Times</td>
<td>Input (n=17)</td>
<td>Mean</td>
<td>4.23</td>
<td>3.16</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SD</td>
<td>1.67</td>
<td>1.18</td>
</tr>
<tr>
<td></td>
<td>Output (n=17)</td>
<td>Mean</td>
<td>4.27</td>
<td>3.54</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SD</td>
<td>1.49</td>
<td>1.24</td>
</tr>
<tr>
<td></td>
<td>Control (n=15)</td>
<td>Mean</td>
<td>4.24</td>
<td>3.86</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SD</td>
<td>1.61</td>
<td>1.85</td>
</tr>
<tr>
<td>ODCT Accuracy</td>
<td>Input (n=17)</td>
<td>Mean</td>
<td>61.88</td>
<td>74.00</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SD</td>
<td>9.64</td>
<td>3.74</td>
</tr>
<tr>
<td></td>
<td>Output (n=17)</td>
<td>Mean</td>
<td>62.11</td>
<td>75.11</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SD</td>
<td>13.11</td>
<td>4.04</td>
</tr>
<tr>
<td></td>
<td>Control (n=15)</td>
<td>Mean</td>
<td>58.93</td>
<td>59.80</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SD</td>
<td>13.15</td>
<td>12.89</td>
</tr>
<tr>
<td>ODCT Planning Times</td>
<td>Input (n=17)</td>
<td>Mean</td>
<td>2.86</td>
<td>1.51</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SD</td>
<td>3.44</td>
<td>0.54</td>
</tr>
<tr>
<td></td>
<td>Output (n=17)</td>
<td>Mean</td>
<td>3.11</td>
<td>1.46</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SD</td>
<td>1.87</td>
<td>0.58</td>
</tr>
<tr>
<td></td>
<td>Control (n=15)</td>
<td>Mean</td>
<td>3.49</td>
<td>2.29</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SD</td>
<td>3.14</td>
<td>1.85</td>
</tr>
<tr>
<td>ODCT Speech Rates</td>
<td>Input (n=17)</td>
<td>Mean</td>
<td>109.89</td>
<td>122.18</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SD</td>
<td>36.56</td>
<td>38.56</td>
</tr>
<tr>
<td></td>
<td>Output (n=17)</td>
<td>Mean</td>
<td>107.11</td>
<td>130.65</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SD</td>
<td>28.62</td>
<td>39.48</td>
</tr>
<tr>
<td></td>
<td>Control (n=15)</td>
<td>Mean</td>
<td>113.20</td>
<td>127.60</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SD</td>
<td>36.27</td>
<td>43.58</td>
</tr>
</tbody>
</table>

**TABLE 3**

Means and Accuracy Rates in Recognizing the Three Types of Request Utterances

<table>
<thead>
<tr>
<th>Request type</th>
<th>Test</th>
<th>Input</th>
<th>Output</th>
<th>Control</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Option A:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pragmatically appropriate,</td>
<td>Pretest</td>
<td>Mean</td>
<td>5.65</td>
<td>6.23</td>
</tr>
<tr>
<td>grammatically accurate</td>
<td></td>
<td>%</td>
<td>70.59%</td>
<td>77.94%</td>
</tr>
<tr>
<td>appropriate</td>
<td>Immediate</td>
<td>Mean</td>
<td>7.53</td>
<td>7.35</td>
</tr>
<tr>
<td>Inaccurate</td>
<td>posttest</td>
<td>%</td>
<td>94.10%</td>
<td>91.91%</td>
</tr>
<tr>
<td>Immediate posttest</td>
<td></td>
<td>Mean</td>
<td>7.82</td>
<td>7.47</td>
</tr>
<tr>
<td><strong>(k =8)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pragmatically appropriate,</td>
<td></td>
<td>%</td>
<td>97.79%</td>
<td>93.38%</td>
</tr>
<tr>
<td>grammatically inaccurate</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Option B:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pragmatically appropriate,</td>
<td>Pretest</td>
<td>Mean</td>
<td>3.24</td>
<td>3.23</td>
</tr>
<tr>
<td>grammatically inaccurate</td>
<td></td>
<td>%</td>
<td>40.44%</td>
<td>40.44%</td>
</tr>
<tr>
<td>appropriate</td>
<td>Immediate</td>
<td>Mean</td>
<td>5.59</td>
<td>3.52</td>
</tr>
<tr>
<td>Inaccurate</td>
<td>posttest</td>
<td>%</td>
<td>69.85%</td>
<td>44.11%</td>
</tr>
<tr>
<td>Immediate posttest</td>
<td></td>
<td>Mean</td>
<td>6.06</td>
<td>3.82</td>
</tr>
<tr>
<td><strong>(k =8)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pragmatically appropriate,</td>
<td></td>
<td>%</td>
<td>75.74%</td>
<td>47.79%</td>
</tr>
<tr>
<td>grammatically accurate</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Option C:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pragmatically inappropriate,</td>
<td>Pretest</td>
<td>Mean</td>
<td>5.12</td>
<td>5.17</td>
</tr>
<tr>
<td>grammatically accurate</td>
<td></td>
<td>%</td>
<td>63.97%</td>
<td>64.70%</td>
</tr>
<tr>
<td>Inappropriate</td>
<td>Immediate</td>
<td>Mean</td>
<td>6.88</td>
<td>6.05</td>
</tr>
<tr>
<td>Inaccurate</td>
<td>posttest</td>
<td>%</td>
<td>86.03%</td>
<td>75.74%</td>
</tr>
<tr>
<td>Immediate posttest</td>
<td></td>
<td>Mean</td>
<td>7.24</td>
<td>5.70</td>
</tr>
<tr>
<td><strong>(k =8)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pragmatically appropriate,</td>
<td></td>
<td>%</td>
<td>90.44%</td>
<td>71.32%</td>
</tr>
</tbody>
</table>