

ScholarWorks@GSU

The Effects of Input-based Practice on Pragmatic Development of Requests in L2 Chinese

Authors	Li, Shuai
Citation	Li, Shuai. 2012. "The Effects of Input-Based Practice on Pragmatic Development of Requests in L2 Chinese." <i>Language Learning</i> 62 (2): 403–38. https://doi.org/10.1111/j.1467-9922.2011.00629.x .
DOI	https://doi.org/10.1111/j.1467-9922.2011.00629.x
Download date	2026-06-06 23:21:10
Link to Item	https://hdl.handle.net/20.500.14694/10648

Running head: PRAGMATICS INSTRUCTION IN L2 CHINESE

The Effects of Input-based Practice on Pragmatic Development of Requests in L2 Chinese

Shuai Li

Carnegie Mellon University

Cite as:

Li, S. (2012). The effects of input-based practice on pragmatic development of requests in L2 Chinese. *Language Learning*, 62, 403–438.

Abstract

This study examined the effects of input-based practice on developing accurate and speedy requests in L2 Chinese. Thirty learners from intermediate-level Chinese classes were assigned to an intensive training group (IT), a regular training group (RT), and a control group. The IT and the RT groups practiced using four Chinese request-making forms via computerized structured input activities over two consecutive days. During this time, the IT group practiced using the request-making forms twice as much as the RT group. The control group did not practice. The results show that the input-based practice was effective in promoting accuracy in an Oral Discourse Completion Task and in enhancing speed in a Pragmatic Listening Judgment Task. No other effects of practice were observed.

Keywords: Pragmatics, interlanguage pragmatics, L2 teaching, Chinese, requests, speech acts, input processing instruction

The Effects of Input-based Practice on Pragmatic Development of Requests in L2 Chinese

Background

Previous research in interlanguage pragmatics (ILP) has shown that, although second language (L2) pragmatic competence does develop among instructed learners in the absence of formal ILP instruction (e.g., Taguchi, 2007a, 2008), learners with high grammatical proficiency can still be pragmatically inappropriate when compared with native speaker norms (Bardovi-Harlig, 2001). Pragmatics instruction is therefore necessary, and the role it plays in promoting L2 pragmatic development is worth empirical investigation. In this regard, previous instructional ILP studies, with a majority targeting L2 English, have shown that a wide range of pragmatic features are teachable. These pragmatic features include various speech acts (e.g., Codina-Espurz, 2008; Fukuya & Clark, 2001; Martínez-Flor, 2008; Martínez-Flor & Fukuya, 2005; Pearson, 2006; Tateyama, 2009), discourse markers (Ishida, 2007; Kakegawa, 2009); routines (e.g., House, 1996); hedging devices (e.g., Wishnoff, 2000), modal particles (e.g., Belz & Vyatkina, 2005; Vyatkina & Belz, 2006), overall discourse characteristics (e.g., Liddicoat & Crozet, 2001; Lyster, 1994), and pragmatic comprehension skills (e.g., Bouton, 1994; Kubota, 1995). In addition, instructed learners generally outperformed uninstructed counterparts (e.g., Billymer, 1990; Da Silver, 2002; Lyster, 1994; Yoshimi, 2001; for review, see Kasper, 1997, 2001; Kasper & Roever, 2005; Kasper & Rose, 2002).

Since L2 pragmatics is teachable and instruction does make a difference, researchers have strived for more effective ways to implement L2 pragmatics instruction. A variety of instructional approaches and techniques have been examined, including Focus on Form (e.g.,

Fukuya & Zhang, 2002; Martínez-Flor & Fukuya, 2005), consciousness-raising (e.g., Kondo, 2008; Takahashi, 2001; Tateyama, 2009), processing instruction (e.g., Takimoto, 2009), explicit feedback (Takimoto, 2006), and inductive vs. deductive instruction (e.g., Rose & Ng, 2001; Martínez-Flor, 2008; Takimoto, 2008), among others. Much effort has been made to compare the differential effects of explicit versus implicit instruction. In explicit instruction, meta-pragmatic information is taught to learners, whereas such information is withheld from learners in implicit instruction (Rose, 2005). Referring to Schmidt's (1990) Noticing Hypothesis, researchers argue that explicit instruction is more effective than implicit instruction because, by providing meta-pragmatic information, it can better draw learners' attention to target pragmatic features, thus increasing the chances for internalizing the features.

In practice, the bulk of studies comparing the effectiveness of explicit and implicit instruction have generally shown the advantage of explicit over implicit instruction (e.g., Alcón-Soler, 2005; House, 1996; Koike & Pearson, 2005; Rose & Ng, 2001; Takahashi, 2001; for review, see Rose, 2005; for a meta-analysis, see Jeon & Kaya, 2006). For example, Alcón-Soler (2005) compared the effects of explicit and implicit instruction on how Spanish speakers learn English request-making forms. The explicit instruction group received meta-pragmatic information of English request-making forms. Then they completed Discourse Completion Task (DCT) exercises to practice using the target request strategies according to different situational scenarios. Feedback on their DCT performance was provided by the instructor. The learners in the implicit instruction group did not receive meta-pragmatic information and instead watched some videos that highlighted the target

request-making forms and contextual factors. They also searched for appropriate forms to complete a number of request utterances extracted from the videos. The same DCT exercises and a range of possible answers were provided for the implicit instruction learners as well. Both groups were found to improve from pretest to posttest (as measured by a role play task and a meta-pragmatic questionnaire) and to outperform a control group that did not receive any instruction. The explicit instruction group did, however, exhibit more gains than the implicit instruction group. Similar results were reported in Jeon & Kaya's (2006) meta-analysis. Based on 13 original instructional ILP studies, the authors found that explicit instruction yielded larger mean effect sizes than implicit instruction. All these findings point to the edge that explicit instruction has over implicit instruction (see Takahashi, 2010 for a recent review on this issue).

Like Alcón-Soler (2005), many researchers have compared learners' pragmatic performance before and after instruction to examine its effects on L2 pragmatic development. Pragmatic performance is typically defined as *pragmatic performance accuracy*, that is, the ability to produce meaning in a socially appropriate manner and to interpret meaning based on contexts (Thomas, 1995). However, *pragmatic performance speed*, that is, the efficiency of carrying out pragmatic tasks, has been largely neglected. In fact, the predominant use of paper-and-pencil type of outcome measures (e.g., written DCT) makes it difficult to assess pragmatic performance speed. Even in cases where oral production data have been collected (e.g., via oral DCT tasks), there has usually been no discussion about pragmatic performance speed.

The lopsided focus on pragmatic performance accuracy rather than pragmatic

performance speed in the literature seems to have overlooked empirical findings documenting L2 pragmatic development among instructed learners without formal ILP instruction.

Specifically, speed and accuracy of L2 pragmatic performance have been identified as distinct components that not only follow different developmental routes but also interact with different social and cognitive factors (Taguchi, 2007a, 2008). For example, Taguchi (2007a) examined the development of 92 Japanese EFL learners' ability to comprehend indirect refusals and indirect opinions over seven weeks in the absence of ILP-specific instruction. A computerized listening task was used to assess the learners' speedy and accurate pragmatic comprehension. The results revealed that, despite significant improvement for both speed and accuracy of pragmatic comprehension, the magnitude of gains for accuracy was larger than for speed over time. In addition, significant correlations were found between pragmatic comprehension accuracy and English proficiency, as well as between pragmatic comprehension speed and lexical access speed. Comprehension accuracy and speed were not correlated. These findings suggest that accuracy and speed are distinct components of L2 pragmatic performance and do not necessarily develop hand in hand. Since pragmatic performance speed develops slowly without targeted formal instruction, as shown in Taguchi's study, it is desirable to explore pedagogical means to facilitate its development. While previous studies have documented instructional effects on the development of pragmatic performance accuracy, it remains unclear whether instruction plays a role in developing pragmatic performance speed.

The development of speed and accuracy of L2 performance has been discussed in cognitive theories of skill development. According to Anderson's (1993) ACT-R (Adaptive

Control of Thought – Rational) model, complex cognitive skill development involves three progressive stages: a cognitive stage, an associative stage, and an autonomous stage. The initial stage of L2 development involves the conscious learning of declarative knowledge (knowledge *that*, i.e., knowing specific rules, such as when should the English past tense marker *-ed* be applied). Performance at this stage is slow and erroneous because learners generally rely on conscious rule application. Over time, through repeated practice, declarative knowledge can be proceduralized. With procedural knowledge (knowledge *how*, i.e., procedural rules such as “if a regular past tense verb is used to talk about past events, then the suffix *-ed* should be applied”), performance becomes faster and more accurate but may still be beneath expert levels. Finally, after extensive practice, rule application becomes automatic, leading to even faster and more accurate performance. Therefore, practice is the driving force in promoting speedy and accurate performance.

The ACT-R model’s potential in guiding and explaining L2 teaching and learning has only recently begun to be explored (DeKeyser, 1996, 1997, for L2 learning; DeKeyser, 1998, 2001, 2007a, 2007b; Ranta & Lyster, 2007, for L2 instruction), and empirical studies suggest that L2 grammar learning can be explained by the skill acquisition model. For instance, DeKeyser (1997) found that the effects of practice on learning morpho-syntactic rules of an artificial language were skill-specific (i.e., rules practiced in comprehension tasks were not readily transferable to production tasks), and that performance in both comprehension and production tasks had less interference from a parallel task. The drop of reaction times and error rates for individual morpho-syntactic rules was sharp during initial practice sessions, and the learning curve became more smooth and flat with increasing amounts of practice.

Thus, the overall developmental patterns fit the power law of practice. According to DeKeyser (1997, 2007c), the initial sharp decrease of reaction times and error rates reflected a qualitative change of knowledge representation (i.e., proceduralization), whereas the learning curve of later practice sessions reflected the process of automatization. These findings demonstrate that practice, as predicted by the ACT-R model, can be effective in promoting speedy and accurate L2 performance.

In interlanguage pragmatics, no study has examined the role of practice (in the sense of skill acquisition theory) in promoting speedy and accurate L2 pragmatic performance. Only two instructional studies touched upon similar issues (Takimoto, 2008, 2009). In these studies, Takimoto examined the effects of three types of input-based instruction on the learning of English request downgraders by Japanese EFL learners with intermediate proficiency. The three types of input-based instruction were: structured input with explicit meta-pragmatic information (EI+SI), structured input only (SI), and problem solving (PS) ¹. The intervention lasted for four sessions (40 minutes each) over two weeks. Four outcome measures were used: a DCT task, a role play task, a written appropriateness judgment task, and a timed listening judgment task (TLJT). The TLJT asked the learners to rate the appropriateness level of the aurally delivered request utterances on five-point scales within two seconds. All other outcome measures were untimed. The results showed that all instructed learners improved significantly on all measures from pretest to posttest, and all outperformed a control group who did not receive instruction on target pragmatic features. On the delayed posttest administered two weeks after the posttest, all three experimental groups retained their gains on all measures. The only exception was that the EI+SI group failed to maintain the gains on

the TLJT. These findings suggest that input-based instruction can be effective in promoting accuracy in pragmatic production (as reflected by the role play and the DCT scores), and that it can also enhance speedy and accurate L2 pragmatic judgment (as reflected by the TLJT results).

The structured input activities in Takimoto's study originated from VanPatten's (2004) Input Processing (IP) model. According to VanPatten (2007), language acquisition starts with the fundamental process of making appropriate form-meaning mappings during the act of comprehending input. The form-meaning connections that are processed become intake, which may subsequently be incorporated into learners' developing interlanguage system and be accessed by the learners as output or production. Based on the IP model, Processing Instruction (PI) is focused on helping learners to first derive and later reinforce target form-meaning connections from input through structured input practice (Wong, 2004). In practice, PI starts with teaching target form-meaning connections, followed by a series of structured input activities. From a skill acquisition perspective, the explicitly taught form-meaning connections can serve as declarative knowledge, and implemented PI "clearly aims at building the procedural knowledge needed for the use of grammar rules in comprehension after the declarative knowledge of these rules has been taught explicitly" (DeKeyser, 2007a, p. 7). Since declarative knowledge is shared across different skill domains and procedural knowledge requires skill-specific practice (DeKeyser, 2007b), it is expected that input-based instruction can improve performance accuracy in both comprehension and production tasks. However, it would be effective in promoting performance speed only in comprehension tasks (rather than in production tasks). These predictions were generally

supported by Takimoto's findings cited above. However, because Takimoto did not adopt any measures of production speed, it remains to be investigated whether input-based instruction can contribute to the development of procedural pragmatic knowledge required in production tasks.

Another under-researched area in L2 pragmatics instruction is the amount of practice needed to promote speedy and accurate pragmatic performance. The findings of Jeon and Kaya's (2006) meta-analysis are relevant here. Based on thirteen original instructional studies in L2 pragmatics, the authors set an arbitrary five-hour cut-off point to compare the effects of instruction that lasted for over 5 hours (i.e., long treatment group) with those that lasted for less than 5 hours (i.e., short treatment group). It was found that long treatments yielded larger mean effect sizes in pragmatic gains than short treatments. However, one should be cautious in interpreting the findings because length of instruction and type of instruction were confounded (i.e., explicit instruction constituted the majority of long treatment studies). Additionally, the complexity of instructional target(s) may further complicate the issue. Certain target pragmatic features may require longer periods of instruction than others due to higher levels of complexity (e.g., teaching the whole speech acts of English compliments and compliment responses versus teaching just English request downgraders). Hence, examining the effect of instructional length *per se* without considering the complexity of instructional targets may oversimplify the issue.

Jeon and Kaya's findings suggest that length of instruction does play a role in L2 pragmatic development, yet it is helpful to go beyond the results and ask why longer treatment can lead to more gains. According to Schmidt (1993), pragmatics acquisition entails

processing target pragmatic features by connecting linguistic forms, illocutionary functions, and applicable contexts (i.e., form-function-context mappings). Instruction offers opportunities for learners to process target form-function-context mappings in different pedagogical activities (e.g., role play). From a skill acquisition perspective, processing target mappings in different tasks constitutes practice, and more practice can lead to a higher level of proceduralization of declarative knowledge. As a result, pragmatic performance speed and accuracy can be enhanced. In this view, a longer treatment is more beneficial than a shorter one because it can provide more instances for processing target pragmatic features. Clearly, empirical evidence is needed to test this prediction. One way to do this is by manipulating the frequency for processing target pragmatic features, that is, the amount of practice.

Finally, since speed and accuracy of pragmatic performance follow different developmental patterns in the absence of targeted ILP instruction (Taguchi, 2007a, 2008), it is reasonable to ask whether similar developmental trajectories can be observed when specific pragmatic instruction is provided. Since almost all ILP instructional studies to date have focused on the gains in pragmatic performance accuracy, little can be said about the relationship between length of instruction and gains in pragmatic performance speed. Hence, studies are needed to investigate whether and how different amounts of practice (i.e., frequency of processing target pragmatic features) affect the gains in both speed and accuracy of L2 pragmatic performance.

In summary, very few instructional ILP studies have examined the effects of formal instruction on the development of both speed and accuracy in L2 pragmatic performance. To this end, skill acquisition theory that recognizes the role of practice in promoting speedy and

accurate performance can shed light on this issue. The limited available research findings show that practice in general can be effective in facilitating L2 pragmatic development, yet more fine-tuned research is needed to show what kind of practice can promote which aspects of L2 pragmatic development. Input-based practice, which has been found effective in promoting the development of L2 grammar and pragmatics, seems promising. Particularly, the relationship between amount of input-based practice and L2 pragmatic development deserves empirical investigation. Finally, investigations focusing on English have dominated the field of L2 pragmatics instruction, and no study to this author's knowledge has ever investigated instructional effects on pragmatic development in L2 Chinese. To fill these gaps in the literature, this study asked: Does the amount of input-based practice affect the development of speedy and accurate performance in making requests in L2 Chinese?

Methodology

Target Pragmatic Features

The instructional target was two pairs of form-function-context mappings for producing Chinese request head acts (Table 1). A request head act is the minimum unit of a request sequence that can realize request intention independent of other elements (Blum-Kulka, House & Kasper, 1989). The mappings were selected based on an empirical study carried out under the framework of Brown and Levinson's (1987) politeness model (Li, 2007). In that study, the data collected from 20 Chinese undergraduates showed that, when two Chinese interlocutors share equal power status (=P) and know each other well (-D), and when the favor being asked is less imposing (-R), it is most appropriate to use direct request strategies, which are realized through the imperatives (e.g., *Jiè wǒ yīxià nǐ de qiānbǐ*. "Lend me your

pencil a little bit”) and the *bǎ* structure² (e.g., *Bǎ chuānghu kāikāi*. “Open the window”). On the other hand, when the speaker occupies lower power status than the hearer (+P) and the two interlocutors know each other well (-D), and when the favor being asked is more imposing (+R), it is most appropriate to use conventionally indirect request strategies, which are realized through the following two structures: “*Néngbùnéng...?*” (e.g., *Nín néngbùnéng ràng wǒ wǎn yī tiān kǎoshì?* “Can you allow me to take the exam one day later?”) or “*Néng...ma?*” (e.g., *Wǒ néng tuīchí yīdiǎnr jiāo wǒ de lùnwén ma?* “Can I submit my paper a little later?”).

Insert Table 1 about here

Participants

Thirty six learners of Chinese were recruited from intermediate level classes in three Chinese programs for this study. All three Chinese programs focused on teaching grammar and vocabulary. Pragmatic features (e.g., address terms) were mentioned only sporadically as they appeared in the lessons. Although different textbooks were used in the three programs, none of them included sections directly related to the target of instruction for this study. The instructors were also interviewed to make sure that they had not taught or discussed anything related to the instructional target of the study.

Among the 36 participants, 13 came from a Chinese program at a US institution (US program). By the time of the study, the 13 participants had received Chinese instruction for about three and a half semesters (four hours per week, and 15 weeks per semester). Another

10 participants were from a study abroad program in Shanghai, China (Shanghai program). The Shanghai program was an eight-week intensive Chinese program with approximately 17.5 hours of instruction per week. Before going abroad, the 10 students had finished two semesters of Chinese language courses at the same US institution as the first group. Finally, 13 participants were recruited from another study abroad program in Beijing, China (Beijing program). The Beijing program was also an eight-week intensive Chinese program with approximately 21 hours of instruction per week. Before going abroad, the participants from the Beijing program had taken two to three semesters of Chinese language courses at their respective institutions in the US. Data collection started in the fourth week of the Shanghai program and the Beijing program, and by then the participants had received comparable amounts of Chinese instruction as their counterparts from the US program.

Among the 36 participants, three were excluded either because they failed to attend all instructional sessions or because of equipment failure. Another three participants were identified as outliers (i.e., $Z > |\pm 3.29|$, $p < .001$) for at least one of the measures of this study³. The remaining 30 participants included 15 Caucasian-Americans, seven Koreans, two Indians, one Japanese, one Japanese-American, one Filipino, one Thai, one Frenchman, and one Spaniard. All participants were between 18 and 23 years of age. Among the 30 participants, 13 were from the US program, nine came from the Shanghai program, and the remaining eight were from the Beijing program. The participants attended the present study in three locations, that is, the US, Beijing, and Shanghai. A Chinese listening comprehension test adapted from a standardized Chinese proficiency test – *Test of Practical Chinese* (HSK Center, 2006) – was administered to see if there was any pre-existing difference among the

three participant groups. The test had 20 multiple choice items, and the results of the *Kruskal-Wallis* test showed no significant difference, $\chi^2(2, N = 30) = 2.65, p > .05$. The 30 participants were randomly divided into three groups: an intensive training group (IT, $n=10$), a regular training group (RT, $n=10$), and a control group ($n=10$). The three groups were comparable in terms of program background, gender distribution, and the number of non-native speakers of English (Table 2). All participants received seven US dollars for each hour of participation.

Insert Table 2 about here

Instrument Development

This study aimed to teach two pairs of form-function-context mappings for making Chinese requests (Table 1). These mappings show that the appropriateness of a request is determined by the context in which the request is made. Hence, it was crucial to develop situations that require the use of the target request-making forms. This study focused on two types of situations (i.e., making small requests to friends, and making big requests to professors). In the process of developing the two types of situations, efforts were made to ensure cross-cultural comparability and authenticity (detailed below).

Development of request-making situations. A pilot study was carried out to develop the two types of request-making situations that were cross-culturally equivalent in their degrees of imposition and authenticity. A Context Judgment Questionnaire (CJQ) was created for this purpose. The CJQ included 26 candidate friend - low imposition (henceforth FL) situations

and 23 candidate professor - high imposition (henceforth PH) situations. The degree of imposition was measured by assessing speakers' psychological difficulty in putting forward the request on a six-point scale, with one being "least difficult" and six being "most difficult". To measure authenticity, a separate six-point scale was used, with one meaning "least likely to happen in real life" and six meaning "very likely to happen in real life." The English and Chinese versions of CJQ were administered to 15 native speakers of American English in the US and 20 native speakers of Chinese in China, respectively. The native speakers were all university undergraduates with comparable age and gender distributions.

The mean ratings of imposition and authenticity were calculated for each candidate situation. A situation was considered as low-imposition if the average ratings from both native speaker groups were below or equal to 2.0 (out of six). A situation was regarded as high-imposition if the average ratings from both native speaker groups were above or equal to 3.5 (out of six). The criteria were met by 14 low-imposition and 14 high-imposition situations. An independent samples *t*-test confirmed that, for both native speaker groups, the mean rating of high-imposition situations was significantly higher than that of low-imposition situations: $t(26) = -22.02, p < .001$ for Chinese native speakers, and $t(26) = -15.64, p < .001$ for American native speakers. The 28 situations also received satisfactory authenticity ratings. The means of the 14 FL situations were 5.29 ($SD = 0.21$) among Chinese native speakers and 5.12 ($SD = 0.35$) among American native speakers. On the other hand, the means for the 14 PH situations were 3.38 ($SD = 0.37$) and 3.09 ($SD = 0.49$) for Chinese and American native speakers, respectively. The 28 situations were further piloted with two learners from the US program to determine the difficulty levels (i.e., vocabulary and grammar for making requests)

and subsequent revisions were made. The two students did not participate in the main study. The revised 28 situations formed a situation pool for developing outcome measures and practice activities.

Outcome measures. Two computerized instruments were used to assess learning outcomes: a Pragmatic Listening Judgment Task (PLJT) and an Oral Discourse Complete Task (ODCT). The purpose of the PLJT was to assess the participants' speedy and accurate recognition of target form-function-context mappings. The PLJT included 28 items: two practice items, twenty target items, and six distracters⁴. The twenty target items were 10 FL situations and 10 PH situations. Among the 20 items, eight were "old" items (i.e., situations that the participants encountered during the practice sessions) and twelve were "new" items (i.e., situations that the participants did not encounter during the practice sessions). The following is a sample PLJT item (practice item):

Yesterday, Professor Wang gave out some handouts in the class. Ma Yang didn't come to the class due to illness. Ma Yang wants to get a copy of the handout from Professor Wang. Ma Yang explains situation and says: 王老师, 您能不能给我一份讲义? (Professor Wang, can you give me a copy of the handout?)

- A. *Pragmatically appropriate and grammatically accurate.*
- B. *Pragmatically inappropriate and grammatically accurate.*
- C. *Pragmatically appropriate and grammatically inaccurate.*

For each PLJT item, the participants first received a mini vocabulary lesson by listening to a few useful Chinese words (read three times each). Meanwhile, the Chinese words, the pinyin of the words (i.e., the most widely used Chinese Romanization system) and their

English translations were displayed on the screen. The participants then heard a request situation in English, accompanied by a written description of the situation shown on the screen. Two seconds after the aural description, the participants heard a request utterance in Chinese. The heard request utterances belonged to one of the following three categories: (a) pragmatically appropriate and grammatically accurate (eight items), (b) pragmatically appropriate and grammatically inaccurate (six items), or (c) pragmatically inappropriate and grammatically accurate (six items). Right after that, the written description of the request situation disappeared, and three options (shown in the sample PLJT item above) appeared after a beep. The participants needed to click one of the three options to indicate their choice as soon as possible. Their choices and response times were recorded on a computer. The three options were fixed in order. The three types of request utterances were counterbalanced across request-making situations⁵. The PLJT had two equivalent versions, PLJT-1 and PLJT-2⁶.

Another outcome measure was the Oral Discourse Completion Task (ODCT), designed to assess the participants' ability to produce target request-making forms in corresponding contexts. The ODCT had 14 items: two practice items, 10 target items, and two distracters⁷. Like the PLJT, there were both "old" ($k=4$) and "new" ($k=6$) items. The following is a sample ODCT item:

Ma Yang and Professor Wang are attending an academic conference in another city. Ma Yang is going to present tomorrow. Unfortunately, Ma Yang's computer broke down. Ma Yang knows that Professor Wang brought a computer and would like to borrow it for tomorrow. Ma Yang explains the situation and says:_____.

For each ODCT item, the participants first received a mini vocabulary lesson (i.e., one or two useful Chinese words, their pinyin and their English translations were displayed on the screen). Then, the participants heard the description of a request situation in English. Meanwhile, they also saw the written description of that situation (in English) on the screen. Immediately after the audio was finished, the written description disappeared and the participants heard a beep. Upon hearing the beep, the participants were instructed to respond orally in Chinese what they would say in that situation. They only needed to say the request head act. Their spoken requests were recorded in computers. The participants clicked the “finished” button to stop the recording and went on to the next item. Two equivalent versions, ODCT-1 and ODCT-2, were created to reduce possible practice effect⁸.

Instruction and Practice

Meta-pragmatic instruction. The two pairs of form-function-context mappings were taught explicitly by the researcher in a mini-lecture that lasted for about 30 minutes. The lecture started with a brief introduction to the concept of request and its components (i.e., head act, alerter, supportive move). The participants then completed the Discourse Completion Task – Version 1 (DCT-1), which was designed to assess their initial knowledge of Chinese requests. The DCT-1 included two FL situations and two PH situations. For each situation, the participants wrote down in Chinese what they thought would be appropriate to say. Handouts were then distributed to illustrate the target mappings with some examples showing how the mappings can be used to make Chinese requests. The researcher-instructor went through the handout with the participants, and then administered the Discourse Completion Test – Version 2 (DCT-2). The DCT-2 had four situations that were comparable

to those included in DCT-1. The purpose of DCT-2 was to make sure that the participants had solid declarative knowledge before attending the practice sessions (explained below). The DCT-1 and the DCT-2 each elicited 120 request utterances (30 participants x four situations per participant). An analysis of the DCT-1 utterances showed that the learners as a group used the target request-making forms in 35 of the 120 utterances, leading to an appropriateness rate of 29.16%. In contrast, the participants achieved a pragmatic appropriateness rate of 100% in responding to the DCT-2, suggesting that they were fully aware of the target form-function-context mappings.

Structured input practice. Over two consecutive days immediately after the mini-lecture, both the regular training (RT) and the intensive training (IT) group received computerized input-based practice (two sessions for each group) in a lab on campus. The IT group received twice as much structured input practice as the RT group. The IT group read 16 dialogues (i.e., eight referential activities and eight affective activities, explained below) with target mappings and then completed the accompanying exercises; the RT group read eight dialogues (i.e., four referential activities and four affective activities, explained below) with target mappings and then completed the accompanying exercises. For both the IT group and the RT group, half of the dialogues were associated with FL situations, and the other half were associated with PH situations. By doing so, the IT group encountered eight instances for each target mapping and the RT group encountered four. Generally, it took about 30 minutes for the IT group and about 20 minutes for the RT group to complete each practice session.

The computerized program was developed with the software named Revolution (Media Version) (2009). The input-based practice program consisted of two types of activities: a

referential activity and an affective activity. Both are important components of Processing Instruction (Wong, 2004), and both were implemented in Takimoto's (2008, 2009) studies on teaching English request downgraders. The referential activities were designed to help the participants make form-function-context mappings (See Appendix 1 for a sample dialogue). The participants first read the description of a request situation in English on the screen. Then, they judged the degree of imposition involved in the request and the social distance between the two interlocutors by clicking relevant buttons (e.g., choosing between "relatively easy" and "relatively difficult" for imposition judgment). Only after correct choices were made⁹ could the participants go on to the next screen showing a dialogue that accompanied the request situation they just read. The participants read the dialogue and chose for each underlined part the appropriate request-making form out of two types: a pragmatic appropriate and grammatically accurate utterance, and a pragmatic inappropriate and grammatically accurate utterance. After the correct choices were made¹⁰, the participants proceeded to the next screen showing the dialogue with situationally appropriate request forms (highlighted). They read and listened to the dialogue twice.

Each referential activity was followed by an affective activity designed to strengthen the target form-function-context mappings by eliciting affective responses (see Appendix 2 for a sample dialogue). Following Takimoto (2008, 2009), affective response was operationalized as an appropriateness rating task. The participants first read a request-making situation in English and an accompanying dialogue in Chinese. The dialogue included four grammatically correct request utterances, among which two were situationally appropriate and the other two were not. The participants clicked the relevant number on a six-point scale to indicate the

level of appropriateness for each of the four utterances, with one meaning “least appropriate” and six meaning “most appropriate”. For a situationally appropriate request form, when the participants clicked number five or six, they saw a pop-up message “Right. Please go on to judge the next request utterance”. When they clicked number one, two, three, or four for the situationally appropriate request form, the message “Please reconsider” popped up, and they could not proceed until the correct number(s) was/were clicked. For a situationally inappropriate request form, the pattern was reversed, that is, only after the participants clicked number one or two were they allowed to move on to judge the next request utterance. Otherwise, they were asked to re-make their judgment until they made the correct choice(s).

Procedures

On Day One, all participants attended the mini-lecture (i.e., meta-pragmatic instruction), followed by a pretest (i.e., ODCT-1 and PLJT-1). After the pretest, the participants were randomly assigned to one of the three groups: the IT group, the RT group, and the control group. On Day Two and Day Three, the IT group and the RT group attended two practice sessions in a lab. The control group did not attend the practice sessions. On Day Four, all groups took a posttest (i.e., ODCT-2, and PLJT-2). Two weeks later, a delayed posttest (i.e., ODCT-1 and PLJT-1) was administered to all participants. For all three tests, the ODCT was administered before the PLJT.

Data Analysis

The PLJT and ODCT data were analyzed for speedy and accurate recognition and production of the target request-making forms. Due to the small sample size, non-parametric procedures were used. Effect sizes of the non-parametric procedures were calculated

following the formulas in Hatch and Lazaraton (1991).

The PLJT data were analyzed for accurate and speedy recognition of target request-making forms in applicable contexts. Accuracy of recognition was operationalized as correct judgment of heard request utterances. One point was assigned to each correct answer, leading to a score range of zero and 20 (one point per item x 20 items). Following Taguchi (2008), speed of recognition was operationalized as response times and was calculated by averaging the number of seconds taken to answer items correctly. For each learner group, the PLJT accuracy scores and the PLJT response times were submitted to two separate *Friedman* tests to see if there was any significant development before and after instruction. In case of significant gains, follow-up *Wilcoxon* tests were performed. The alpha level of the *Wilcoxon* tests was adjusted to .016 for three paired comparisons (i.e., pretest vs. posttest; pretest vs. delayed posttest; posttest vs. delayed posttest). Between-group comparisons in accuracy scores and response times were made with separate *Kruskal-Wallis* tests. In case of significant differences, follow-up *Mann Whitney U* tests were performed. Again, the alpha level was adjusted to .016 for three paired comparisons (RT vs. IT; RT vs. control; and IT vs. control).

The ODCT data were analyzed for accurate and speedy production of target request-making forms in applicable contexts. Accuracy of production was operationalized as ODCT accuracy scores based on a scoring rubric. The rubric adopted pragmatic appropriateness as the primary criterion and grammatical accuracy as the secondary criterion. Specifically, four points were awarded to a pragmatically appropriate and grammatically accurate utterance; three points were given to a pragmatically appropriate and grammatically

inaccurate utterance; two points were assigned to a pragmatically inappropriate and grammatically accurate utterance; and one point was given to a pragmatically inappropriate and grammatically inaccurate utterance. A zero was awarded if there was no response.

Among the 900 target ODCT utterances (10 utterances per participant per test x 30 participants x three tests), 300 were scored independently by this author and another Chinese native speaker trained in applied linguistics. The results showed that the two raters agreed on 286 out of the 300 utterances (95.33%). Discrepancies were discussed between the two raters and consensus reached. The remaining 600 ODCT utterances were scored by this author. The ODCT accuracy score range was between zero and 40 for each participant (four points per item x 10 items). The speed of production was measured in terms of planning time and speech rate. Following Taguchi (2007b), planning time was operationalized as time taken to prepare for ODCT responses. It was measured by calculating the averaged seconds taken to produce pragmatically appropriate request utterances (i.e., using target request-making form(s) in its applicable contexts). Speech rate was measured by the averaged number of Chinese syllables spoken per minute when producing pragmatically appropriate request utterances, excluding false starts, repetitions, partial repetitions, and repairs. For each learner group, the ODCT accuracy scores, the ODCT planning times, and the ODCT speech rates were submitted to three separate *Friedman* tests. Follow-up *Wilcoxon* tests with the adjusted alpha level of .016 (for three paired comparisons) were performed in cases of significant difference. Between-group comparisons in accuracy scores, planning times, and speech rates were made with separate *Kruskal-Wallis* tests. Subsequent *Mann Whitney U* tests with the adjusted alpha level of .016 (for three paired comparisons) were performed to locate the source of

difference.

To summarize, 36 learners of Chinese participated in this study and 30 were retained for final data analysis. The participants were randomly assigned into three groups (IT, RT and control). All groups received the same meta-pragmatic information regarding the target form-function-context mappings (Table 1) on Day One. On Day Two and Day Three, the IT group and the RT group attended two input-based practice sessions (i.e., referential activity and affective activity), and the IT group practiced twice as much as the RT group. The control group did not practice. A Pragmatic Listening Judgment Task (PLJT) with 26 items and an Oral Discourse Completion Task (ODCT) with 12 items were administered one day before, one day after, and two weeks after the input-based practice period to elicit the participants' performance on recognizing and producing target request-making forms. The ODCT data and the PLJT data were analyzed in speed (i.e., PLJT response times, ODCT response times, ODCT speech rates) and accuracy (i.e., PLJT accuracy, and ODCT accuracy) of pragmatic performance. Non-parametric statistical procedures were used for data analysis.

Results

Pragmatic Listening Judgment Task Accuracy

Table 3 displays the means and standard deviations of the PLJT accuracy scores across three learner groups. The results of the *Friedman* tests showed that no group made significant gains over time: RT, $\chi^2(2, n = 10) = 4.32, p = .11, \eta^2 = .15$; IT, $\chi^2(2, n = 10) = 0.74, p = .73, \eta^2 = .03$; and control, $\chi^2(2, n = 10) = 3.50, p = .19, \eta^2 = .12$. The *Kruskal-Wallis* tests also revealed no difference between the three groups at any time point: pretest, $\chi^2(2, N=30) = 3.05, p = .22, \eta^2 = .11$; posttest, $\chi^2(2, N=30) = 1.80, p = .42, \eta^2 = .06$; and delayed posttest, χ^2

(2, $N=30$) = 2.17, $p = .35$, $\eta^2 = .08$.

Insert Table 3 about here

In responding to the PLJT, the learners were asked to judge three types of request utterances (i.e., Option A. pragmatically appropriate and grammatically accurate, Option B. pragmatically appropriate and grammatically inaccurate, and Option C. pragmatically inappropriate and grammatically accurate). A *post hoc* analysis was performed to see which type of request utterances the learners had difficulty with. Table 4 displays the mean number of correct answers for each request type and the associated accuracy rates. Clearly, all three groups scored consistently high in judging Option A and Option C types of request utterances. By contrast, the learners performed poorly in dealing with Option B type of request utterances across pretest, posttest, and delayed posttest.

Insert Table 4 about here

Following the suggestion from one reviewer, an additional *post hoc* analysis was conducted to examine the participants' tendency to misjudge Option B type request utterances. The purpose was to find out how likely participants were to choose Option A (+P, -G) or Option C (-P, +G) when they heard a pragmatically appropriate and grammatically inaccurate request utterance (+P, -G). Table 5 shows the mean error rates for each learner group on pretest, posttest, and delayed posttest. For instance, among all the errors that the RT group

made in judging Option B type of request utterances on pretest, Option A constituted 81.33% of the misjudgment while Option C took up 18.70%. As Table 5 shows, when there was an error in judging Option B type of request utterances, all groups were more likely to choose Option A than Option C, and the pattern remained the same over time. Thus, these results showed that the participants, in general, were more sensitive to pragmatic appropriateness than to grammatical accuracy.

Insert Table 5 about here

Pragmatic Listening Judgment Task Response Times

Table 6 displays the means and standard deviations of the PLJT response times of the three learner groups. The results of the *Friedman* tests showed that neither the RT group nor the control group significantly speeded up response times over time: RT, $\chi^2(2, n = 10) = 2.60$, $p = .32$, $\eta^2 = .09$; and control, $\chi^2(2, n = 10) = 5.60$, $\eta^2 = .19$. The response times of the IT group, however, speeded up significantly after practice, $\chi^2(2, n = 10) = 11.40$, $p = .002$, $\eta^2 = .39$. Subsequent *Wilcoxon* tests showed that the IT group improved from pretest to posttest, $Z = -2.39$, $p = .014$, $\eta^2 = .64$, and retained the gains from posttest to delayed posttest, $Z = -1.988$, $p = .049$, $\eta^2 = .44$. Furthermore, the IT group's performance was significantly faster on delayed posttest than on pretest, $Z = -2.80$, $p = .002$, $\eta^2 = .87$. On the other hand, although the IT group made significant gains after practice, it failed to outperform the other two groups. The *Kruskal-Wallis* tests revealed no difference among the three groups over time: pretest: $\chi^2(2, N=30) = 0.17$, $p = .92$, $\eta^2 = .006$; posttest, $\chi^2(2, N=30) = 1.55$, $p = .46$, $\eta^2 = .05$; and

delayed posttest, $\chi^2(2, N=30) = 4.99, p = .08, \eta^2 = .17$. Thus, the results showed a weak effect of practice on speeding up PLJT response times for the IT group.

Insert Table 6 about here

Oral Discourse Completion Task Accuracy

Table 7 shows the means and standard deviations of the ODCT accuracy scores of the three learner groups. The *Friedman* tests revealed that both the RT group and the IT group made significant improvement over time: RT, $\chi^2(2, n = 10) = 9.14, p = .008, \eta^2 = .32$; and IT, $\chi^2(2, n = 10) = 14.35, p < .001, \eta^2 = .49$. The control group did not improve, $\chi^2(2, n = 10) = 0.65, p = .77, \eta^2 = .03$. Follow-up *Wilcoxon* tests showed that the IT group and the RT group both improved significantly from pretest to posttest: RT, $Z = -2.49, p = .012, \eta^2 = .69$; IT, $Z = -2.67, p = .004, \eta^2 = .79$, and that both retained their gains from posttest to delayed posttest: RT, $Z = -0.10, p = .99, \eta^2 = .001$; IT, $Z = -0.38, p = 1.00, \eta^2 = .016$. Furthermore, their delayed posttest scores were significantly higher than their pretest scores: RT, $Z = -2.54, p = .008, \eta^2 = .715$; IT, $Z = -2.677, p = .004, \eta^2 = .79$.

The observed effects of practice were further corroborated by the results of the *Kruskal-Wallis* tests, which showed no difference among the three groups on pretest, $\chi^2(2, N=30) = 0.03, p = .99, \eta^2 = .001$, but significant group difference on posttest, $\chi^2(2, N=30) = 8.89, p = .008, \eta^2 = .31$, as well as on delayed posttest, $\chi^2(2, N=30) = 7.96, p = .015, \eta^2 = .27$. Subsequent *Mann-Whitney U* tests revealed that the IT group outperformed the control group on posttest, $Z = -2.66, p = .007, \eta^2 = .37$, and on delayed posttest, $Z = -2.72, p = .005,$

$\eta^2 = .39$. However, it did not outperform the RT group on posttest, $Z = -1.66$, $p = .12$, $\eta^2 = .14$, or on delayed posttest, $Z = -0.85$, $p = .45$, $\eta^2 = .04$. On the other hand, the RT group failed to outperform the control group on either posttest, $Z = -1.89$, $p = .06$, $\eta^2 = .19$, or delayed posttest, $Z = -1.91$, $p = .06$, $\eta^2 = .19$. In summary, the developmental trajectory of the IT group demonstrated a strong effect of practice, while the RT group's performance showed a weak effect of practice.

Insert Table 7 about here

Oral Discourse Completion Task Planning Times

Table 8 presents the means and standard deviations of the planning times for producing situationally appropriate request utterances. Three separate *Friedman* tests showed that none of the learner groups made any significant development over time: RT, $\chi^2(2, n = 10) = 3.20$, $p = .22$, $\eta^2 = .11$; IT, $\chi^2(2, n = 10) = 3.80$, $p = .19$, $\eta^2 = .13$; and control, $\chi^2(2, n = 10) = 0.80$, $p = .71$, $\eta^2 = .03$. The *Kruskal-Wallis* tests also revealed no difference among the groups over time: pretest: $\chi^2(2, N=30) = 1.73$, $p = .44$, $\eta^2 = .06$; posttest, $\chi^2(2, N=30) = 4.48$, $p = .11$, $\eta^2 = .16$; and delayed posttest: $\chi^2(2, N=30) = 1.24$, $p = .54$, $\eta^2 = .04$. Hence, the results showed no effect of practice on reducing ODCT planning times across the groups.

Insert Table 8 about here

Oral Discourse Completion Task Speech Rates

Table 9 shows the means and standard deviations of the speech rates of situationally appropriate request utterances across the learner groups. Three separate *Friedman* tests revealed that all groups made significant development over time: RT, $\chi^2(2, n = 10) = 8.60$, $p = .012$, $\eta^2 = .29$; IT, $\chi^2(2, n = 10) = 7.80$, $p = .018$, $\eta^2 = .27$; and control, $\chi^2(2, n = 10) = 6.20$, $p = .046$, $\eta^2 = .21$. Follow-up *Wilcoxon* tests showed that the RT group made significant gains from pretest to delayed posttest, $Z = -2.49$, $p = .01$, $\eta^2 = .69$, but there was no difference between pretest and posttest, $Z = -1.58$, $p = .13$, $\eta^2 = .28$, nor was there any difference between posttest and delayed posttest, $Z = -0.87$, $p = .43$, $\eta^2 = .08$. For the IT group, the only significant development occurred between posttest and delayed posttest, $Z = -2.80$, $p = .002$, $\eta^2 = .87$. No difference was found between pretest and posttest, $Z = -0.76$, $p = .49$, $\eta^2 = .07$. For the comparison between pretest and delayed posttest, again the result was not significant based on the adjusted alpha level, $Z = -2.19$, $p = .027$, $\eta^2 = .53$, although the p value of .027 suggested a trend that the participants improved over time. As for the control group, none of the paired comparisons reached significant level: pretest vs. posttest, $Z = -1.78$, $p = .08$, $\eta^2 = 0.35$; pretest vs. delayed posttest, $Z = -2.19$, $p = .027$, $\eta^2 = .53$; and posttest vs. delayed posttest, $Z = -0.76$, $p = .49$, $\eta^2 = .07$. However, there was a trend that the control group gained from pretest to delayed posttest, as indicated by the p value of .027.

The results of the *Kruskal-Wallis* tests showed no difference on posttest, $\chi^2(2, N=30) = 3.27$, $p = .19$, $\eta^2 = .11$, and delayed posttest, $\chi^2(2, N=30) = 4.09$, $p = .13$, $\eta^2 = .14$. However, the three groups differed on pretest, $\chi^2(2, N=30) = 6.60$, $p = .04$, $\eta^2 = .23$. Follow-up *Mann-Whitney U* tests revealed that there was a trend for the IT group to outperform the RT group, $Z = -2.21$, $p = .027$, $\eta^2 = .26$, and the control group, $Z = -2.19$, $p = .029$, $\eta^2 = .25$,

although neither comparison reached significant level. The RT and the control group did not differ in speech rate on pretest either, $Z = -0.45$, $p = .68$, $\eta^2 = .01$.

In summary, there were clear trends showing that all three groups improved from pretest to delayed posttest. Since the control group – who did not receive input-based practice – also gained in ODCT speech rates, one cannot attribute the improvements made by the RT group and the IT group to input-based practice. In other words, the effects of instructional practice were negligible.

Insert Table 9 about here

Collectively, the results showed the effects of input-based practice in the following two aspects. First, regarding the measure of PLJT response time, the IT group significantly reduced the time needed to make correct judgments from pretest to posttest and retained these gains on delayed posttest. However, the IT group did not outperform the other two groups on posttest and delayed posttest. These two lines of evidence revealed a weak effect of practice on speeding up PLJT response times for the IT group. Second, regarding ODCT accuracy, the IT group not only showed significant improvement over time but also outperformed the control group on both posttest and delayed posttest. On the other hand, the RT group, although showing significant gains over time, failed to outperform the control group. The different developmental trajectories displayed by the IT group and the RT group point to the differential effects that different amounts of input-based practice had on promoting ODCT accuracy. No other effects of instructional practice were observed.

Discussion

This study investigated the effects of different amounts of instructional practice on the development of accurate and speedy recognition and production of request-making forms in L2 Chinese. The findings regarding the accuracy measures and the speed measures will be discussed in turn.

Regarding the PLJT accuracy measure, those who participated in the practice sessions did not become more accurate in recognizing target request-making forms, nor did they outperform the control group at any time. These findings suggest that practice was not effective in promoting the learners' ability to recognize different types of request utterances. In addition, the results of the *post hoc* analysis (Table 4) showed that the learners were much better at identifying Option A and Option C types of request utterances than recognizing Option B type of request utterances.

Why, then, would the learners' ability to recognize Option A and Option C types of request utterances be so different from their ability to recognize Option B type of request utterances? Moreover, why did the instructional practice fail to enable the learners to become more competent in identifying request utterances that were pragmatically appropriate and grammatically inaccurate? In judging Option A and Option C types of request utterances, the learners only had to pay attention to pragmatic appropriateness (i.e., whether the target forms were used in applicable situations) without having to attend to the grammaticality of the heard utterances. Since there were only two situation types (i.e., friend - low imposition, and professor - high imposition), it was probably not difficult for the learners to decide whether a heard request utterance belonged to Option A or Option C, especially after being taught

explicitly about the target mappings. This meant that meta-pragmatic instruction might have been enough to enable the participants to correctly judge the pragmatic appropriateness of the heard request utterances. Since all learners received the same meta-pragmatic instruction, there should be little difference among the groups in judging Option A and Option C types of request utterances.

In judging option B type of request utterances, however, the learners had to further decide the grammaticality of the heard utterances. Grammaticality was not covered in the meta-pragmatic instruction session (i.e., the mini lecture). The practice sessions did not offer any opportunity for practicing grammaticality judgment either, because the dialogues in the practice sessions were all grammatically correct. Grammatical accuracy was not emphasized and practiced (in terms of grammaticality judgment) and pragmatic appropriateness was highlighted and practiced. In hindsight, therefore, it was to be expected that the learners would show greater sensitivity to pragmatic appropriateness than to grammatical accuracy when dealing with Option B type of request utterances. The results in Table 5 support this interpretation, as all participants were more likely to choose Option A (pragmatically appropriate and grammatically accurate) than Option C (pragmatically inappropriate and grammatically accurate) when making errors in judging Option B type of request utterances. Taken together, these observations help explain why the IT group and the RT group showed negligible improvement after attending the practice sessions.

Although the practice effect was minimal on PLJT accuracy development, the practice sessions were effective in promoting the learners' ability to produce pragmatically appropriate and grammatically accurate requests over time. The RT group and the IT group

both significantly increased their ODCT accuracy scores from pretest to posttest and maintained the gains through the delayed posttest. The fact that the gains occurred only from pretest to posttest and that there was barely any improvement from posttest to delayed posttest suggest that the gains were attributable to the input-based practice sessions. These findings corroborate Takimoto's (2008, 2009) findings, which also showed significant increase in DCT and role play accuracy scores after engaging in input-based activities. Broadly speaking, the findings also echo previous studies that documented positive instructional effects on promoting pragmatic accuracy in production tasks (e.g., Alcón-Soler, 2005; House, 1996; Kakegawa, 2009; Kondo, 2008; Rose & Ng, 2001; Takahashi, 2001; Tateyama, 2009). Unlike previous research in which pretests were placed before any pedagogical intervention, the pretest of this study was administered after the meta-pragmatic instruction. Therefore, the gains in ODCT accuracy scores reflected the unique contribution of input-based practice, over and above meta-pragmatic instruction.

In what way, then, was the practice helpful in improving the learners' pragmatic performance accuracy as reflected by the gains in ODCT accuracy score? As Table 7 shows, the mean scores of the three groups were all above 30 (out of a maximum of 40). Based on the scoring rubric, the relatively high mean scores suggest that the learners were able to use the target request-making forms as early as when they took the pretest. What the RT group and the IT group improved over time was their ability to produce grammatically correct request-making forms in applicable contexts. In other words, the role of the practice sessions was to fine-tune the learners' declarative knowledge of how to construct grammatically correct request utterances based on the explicitly taught request-making forms.

Particularly interesting in the results is how the different amounts of practice affected the gains in ODCCT accuracy scores. The IT group outperformed the control group at both posttest and delayed posttest. By contrast, the RT group did not outperform the control group at any time. These findings suggest that eight instances of processing target pragmatic mappings were sufficient to enable the learners in the IT group to make significant improvement over time as well as to outperform the control group. On the other hand, four instances of processing were also enough for the learners in the RT group to make significant improvement, but not sufficient for them to perform better than the control group. Because the IT group processed target mappings twice as often as the RT group, their declarative knowledge was refined to a greater extent than that of the RT group. This observation was further corroborated by the larger effect sizes of the IT group than the RT group. Hence, the findings point to a positive association between the amount of instructional practice and the magnitude of improvement in pragmatic performance accuracy for production tasks.

Although practice was effective in promoting the learners' pragmatic performance accuracy, its role in enhancing pragmatic performance speed was limited. Regarding ODCCT speech rates, all three groups made significant progress over time. Since the major increase in speech rates for all groups occurred either between pretest and delayed posttests (with a three-week interval) or between posttest and delayed posttest (with a two-week interval), the unanimous improvement can be interpreted as a result of natural development, rather than as a function of instructional practice. By contrast, no group made significant gains in ODCCT planning time. On the other hand, on the measure of PLJT response times, the IT group did make significant gains, yet the effect of practice was weak, since the IT group did not

outperform the control group. Interestingly, although the mean PLJT response times showed that, to varying degrees, all groups became faster over time (Table 6), it is likely that the varied pace of development reflected the effects of differential amounts of practice. With four instances of processing target mappings, the RT group was not able to even surpass its own performance; with eight instances of processing, the IT group was able to make significant gains over time, but not enough to outperform the other two groups. Hence, it might be argued that a certain threshold in the amount of instructional practice is needed before any significant development in the speed of making pragmatic judgment can be observed.

From the perspective of skill acquisition theory, it is not difficult to understand why, on the one hand, there were negligible gains in pragmatic performance speed in the productive task (i.e., the ODCT) and, on the other hand, there was at least a tendency toward improvement in pragmatic performance speed in the receptive task (i.e., the PLJT). The crux of the matter lies in the development of different types of procedural knowledge as a function of input-based practice. According to skill acquisition theory, the development of procedural knowledge requires repeated applications of explicitly learnt declarative knowledge in target behaviors (i.e., practice). But the transition from declarative knowledge to procedural knowledge in different skill domains (e.g., productive and receptive) requires skill-specific practice. In other words, practice in receptive skills will not benefit productive skills (for developing procedural knowledge) and vice versa. In this study, the input-based practice repeatedly asked the learners to identify pragmatically appropriate request utterances through referential and affective activities. In so doing, it clearly aimed at building the procedural knowledge associated with pragmatic performance in receptive tasks. On the other hand, the

input-based practice could not lead to the development of procedural knowledge associated with pragmatic performance in productive tasks. Because the PLJT response time was an indication of the procedural knowledge associated with receptive tasks, whereas the ODCT planning time and the ODCT speech rate reflected the procedural knowledge associated with productive tasks, it is not surprising that the measure of PLJT response time demonstrated signs of improvement but that both ODCT speed measures failed to show any effect of practice. These findings echo previous research on L2 grammar learning that shows a lack of transfer between receptive (i.e., comprehension) and productive (i.e., production) skills at the level of both proceduralized and automatized knowledge (e.g., DeKeyser, 1997). At the level of declarative knowledge, however, the effects of input-based practice can be transferred in the sense that practice in one skill can help fine-tune the declarative knowledge of the opposite skill (DeKeyser, 2007b). This is why both the RT group and the IT group made significant improvement in their oral production accuracy (i.e., ODCT accuracy score) even though the practice sessions were essentially aimed at training the learners' receptive skills. The findings of the study thus partially confirmed the predication set by skill acquisition theory in the area of L2 pragmatics learning.

The present findings also add to our understanding of the effects of processing instruction. Based on VanPatten's theory of input processing, processing instruction is aimed at helping learners develop underlying linguistic systems that can be accessed during comprehension and production. The results of this study show transfer effects at the level of pragmatic performance accuracy but not pragmatic performance speed. Hence, it can be argued that, after receiving input-based instruction, L2 learners can access their underlying

linguistic systems in both comprehension and production tasks, but the efficiency of access varies across different skill domains.

Limitations and Implications for Future Research

This study investigated the effects of input-based practice on pragmatic development of requests in L2 Chinese. Different from previous instructional studies that typically have focused on the combined effects of meta-pragmatic information sessions and practice activities, the findings of this study demonstrated the unique contribution of practice to L2 pragmatic development, over and above explicit meta-pragmatic instruction. In addition, it was found that the amount of input-based practice can influence the development of speed and accuracy of L2 pragmatic performance, albeit to different degrees. As one reviewer pointed out, the overall effects of practice were moderate even for the IT group. However, since no instructional ILP study has investigated the relationship between amount of practice and gains in pragmatic performance speed and accuracy, this study can be seen as the first attempt to explore the issue. It is hoped that the findings will invite future research in this area.

This study has several limitations. A noticeable one is the small sample size, which made it inappropriate to apply parametric statistics. The small sample size might have also increased the chance of a Type II error. To reduce the negative effects of small sample size, non-parametric procedures were used. Another limitation is that the participants were from three Chinese programs, and their L1 backgrounds were mixed. Efforts were made in this study to ensure that the three Chinese programs were evenly represented in each group and that the proportion of non-native speakers of English was comparable across the three groups.

Yet future research should target a more homogenous learner group with a larger sample size.

Based on the findings of this study, there is a need to further explore the role of differential amounts of practice in promoting the development of declarative and procedural knowledge in L2 pragmatics. The findings of this study show that, given a fixed amount of practice, declarative knowledge is more amenable to improvement than the corresponding procedural knowledge. Thus, while eight instances of processing target pragmatic features were enough to enable the IT group to make significant gains in declarative knowledge (as measured by ODCCT accuracy scores), that amount of practice was just enough to show some signs of improvement in procedural knowledge (as measured by the PLJT response time). Hence, it remains to be examined how much practice is needed to promote procedural knowledge development.

A related issue is what kind(s) of practice can promote speedy and accurate pragmatic performance in different skill domains. The input-based practice adopted in this study, though effective in certain ways, did not seem to be effective in promoting pragmatic performance speed on production tasks. Considering the skill specificity issue of practice, it would be desirable to explore the potential of output-based practice as well.

Finally, there is also a need to expand the target of instruction in order to increase the generalizability of the observed effects of practice. This study focused on the request head act because it is the core component of any request sequence. However, there are other components (e.g., request modifications) that can also determine the appropriateness of a request. Future research should explore the effects of instructional practice on the learning of other elements of request in L2 Chinese, in conjunction with request head acts. Furthermore,

researchers should also investigate the effects of practice on the acquisition of other pragmatic features (e.g., other speech acts such as complaint and/or apology). It has been a common practice in the field of L2 pragmatics instruction to focus on only one speech act at a time (e.g., Alcón-Soler, 2005; Martínez-Flor & Fukuya, 2005; Ghobadi & Fahim, 2009), and the issue of external validity has largely been neglected. This study focused on the speech act of requesting because the primary purpose was to examine the effects of differential amounts of practice. However, just as one reviewer pointed out, effective ILP instruction can only be designed if findings are generalizable across a wide variety of pragmatic features. Given that the effects of differential amounts of practice were documented in this study, future research should incorporate multiple target pragmatic features to address the issue of generalization.

In summary, this study documented the development of speed and accuracy in L2 pragmatic performance as a function of differential amounts of input-based practice. The results demonstrated the feasibility and the potential of extending what is known to be effective in promoting L2 grammar learning to L2 pragmatics instruction. In so doing, this study extended the insights of skill acquisition theory to the area of L2 pragmatics teaching and learning. It also showed the extent to which processing instruction can contribute to L2 pragmatic development.

References

- Alcón-Soler, E. (2005). Does instruction work for learning pragmatics in the EFL context? *System*, 33, 417-435.
- Anderson, J. R. (1993). *Rules of the mind*. Hillsdale, New Jersey: Lawrence Erlbaum Associations.
- Bardovi-Harlig, K. (2001). Evaluating the empirical evidence: Grounds for instruction in pragmatics? In G. Kasper & K. R. Rose (Eds.) *Pragmatics in language teaching* (pp.13-32). New York, NY: Cambridge University Press.
- Belz, J. A., & Vyatkina, N. (2005). Learner corpus analysis and the development of L2 pragmatic competence in networked intercultural language study: The case of German modal particles. *The Canadian Modern Language Review*, 62, 17-48.
- Billmyer, K. (1990). "I really like your lifestyle: ESL learners learning how to compliment. *Penn Working Papers in Educational Linguistics*, 6, 31-48.
- Blum-Kulka, S., House, J., & Kasper, G., (1989). *Cross-cultural pragmatics: Requests and apologies*. Norwood, NJ: Ablex.
- Bouton, L. (1994). Can NNS skill in interpreting implicatures in American English be improved through explicit instruction? In L. Bouton & Y. Kachru (Eds.), *Pragmatics and language learning, Monograph Series 5* (pp.88-109). Division of English as an International Language, University of Illinois, Urbana-Champaign, IL.
- Brown, P., & Levinson, S. C. (1987). *Politeness – some universals in language usage*. Cambridge: Cambridge University Press.
- Codina-Espurz, V. (2008). The immediate vs. delayed effect of instruction on mitigators in

relation to the learner's language proficiency in English. In Alcón-Soler (Ed.), *Learning how to request in an instructed language learning context* (pp.228-256). Berlin: Peter Lang.

Da Silvia, A. J. B. (2003). The effects of instruction on pragmatic development: Teaching polite refusals in English. *Second Language Studies*, 22, 55-106.

DeKeyser, R. (1996). Exploring automatization processes. *TESOL Quarterly*, 30, 349-357.

DeKeyser, R. (1997). Beyond explicit rule learning: Automatizing second language morphosyntax. *Studies in Second Language Acquisition*, 19, 195-221.

DeKeyser, R. (1998). Beyond focus on form. In C. Doughty & J. Williams (Eds.), *Focus on form in classroom second language acquisition* (pp. 42-63). New York, NY: Cambridge University Press.

DeKeyser, R. (2001). Automaticity and automatization. In P. Robinson (Ed.), *Cognition and second language instruction* (pp.125-151). Cambridge: Cambridge University Press.

DeKeyser, R. (2007a). Introduction: Situating the concept of practice. In R. DeKeyser (Ed.), *Practice in a second language: Perspectives from applied linguistics and cognitive psychology* (pp. 1-18). Cambridge: Cambridge University Press.

DeKeyser, R. (2007b). The future of practice. In R. DeKeyser (Ed.), *Practice in a second language: Perspectives from applied linguistics and cognitive psychology* (pp. 287-304). Cambridge: Cambridge University Press.

DeKeyser, R. (2007c). Skill acquisition theory. In B. VanPatten & J. Williams (Eds.), *Theories in Second language acquisition: An introduction* (pp.94-113). Mahwah, Jersey: Lawrence Erlbaum Associates, Publishers.

- Fukuya, Y., & Clark, M. K. (2001). A comparison of input enhancement and explicit instruction of mitigators. In L. Bouton (Ed.), *Pragmatics and language learning, Monograph Series 10* (pp. 111 – 130). Urbana, IL: University of Illinois at Urbana-Champaign.
- Fukuya, Y., & Zhang, Y. (2002). Effects of recasts on EFL learners' acquisition of pragmalinguistic conventions of request. *Second Language Studies, 21*, 1-47.
- Ghobadi, A., & Fahim, M. (2009). The effect of explicit teaching of English “thanking formulas” on Iranian EFL intermediate level students at English language institutes. *System, 37*, 526–537.
- Hatch, E., & Lazaraton, A. (1991). *The research manual: Design and statistics for applied linguistics*. Boston, MA: Heinle & Heinle Publishers.
- House, J. (1996). Developing pragmatic fluency in English as a foreign language. *Studies in Second Language Acquisition, 18*, 225-252.
- HSK center. (2006). Test of Practical Chinese. Available at <http://www.c-test.org.cn/ListDown2.asp?ClassType=%E8%AF%95%E9%A2%98>
- Ishida, K. (2007). Developing understanding of how the *desu / masu* and plain forms express one's stance. In D. R. Yoshimi & H. Wang (Eds.). *Selected papers from the conference on pragmatics in the CJK classroom: The state of the art* (pp. 181-202). <http://nflrc.hawaii.edu/FJKProceedings>.
- Jeon, E. H., & Kaya, T. (2006). Effects of L2 instruction on interlanguage pragmatic development. In J. M. Norris & L. Ortega (Eds.), *Synthesizing research on language learning and teaching* (pp.165-211). Amsterdam / Philadelphia: John Benjamins

Publishing Company.

Kasper, G. (1997). Can pragmatic competence be taught? *Second Language Teaching & Curriculum Center*. <http://www.nflrc.hawaii.edu/networks/NW06/>.

Takegawa, T. (2009). Development of the use of Japanese sentence-final particles through email correspondence. In N. Taguchi (Ed.), *Pragmatic competence* (pp.301-333). Berlin: Mouton de Gruyter.

Kasper, G. (2001). Classroom research on interlanguage pragmatics. In G. Kasper & K. R. Rose (Eds.), *Pragmatics in language teaching* (pp.33-60). New York, NY: Cambridge University Press.

Kasper, G., & Rose, K. R. (2002). *Pragmatic development in a second language*, Blackwell Publishing.

Koike, D. A., & Pearson, L. (2005). The effect of instruction and feedback in the development of pragmatic competence. *System*, 33, 481-501.

Kubota, M. (1995). Teachability of conversational implicatures to Japanese EFL learners. *IRLT Bulletin*, 9, 35-67.

Li, S. (2007). *A study on the pragmatic development of making requests by American learners of Chinese* (Unpublished master's thesis). Beijing Language and Culture University, Beijing, China.

Liddicoat, A., & Crozet, C. (2001). Acquiring French interactional norms through instruction. In K. R. Rose & G. Kasper (Eds.), *Pragmatics in language teaching* (pp.125-144). Cambridge, UK: Cambridge University Press.

Lyster, R. (1994). The effect of functional-analytic teaching on aspects of French immersion

students' sociolinguistic competence. *Applied Linguistics*, 15, 263-287.

Martínez-Flor, A. (2008). The effect of an inductive-deductive teaching approach to develop learners' use of request modifiers in the EFL classroom. In Alcón-Soler (Ed.), *Learning how to request in an instructed language learning context* (pp.191-226). Berlin: Peter Lang.

Martínez-Flor, A., & Fukuya, Y. (2005). The effects of instruction on learners' production of appropriate and accurate suggestions. *System*, 33, 463-480.

Pearson, L. (2006). Patterns of development in Spanish L2 pragmatic acquisition: An analysis of novice learners' production of directives. *The Modern Language Journal*, 90, 473-495.

Ranta, L., & Lyster, R. (2007). A cognitive approach to improving immersion students' oral language abilities: The awareness-practice-feedback sequence. In R. DeKeyser (Ed.), *Practice in a second language: Perspectives from applied linguistics and cognitive psychology* (pp. 141-160). Cambridge: Cambridge University Press.

Revolution (Media Version) (2009). [Computer software] . Edinburgh, UK: Runtime Revolution Ltd.

Rose, K. R. (2005). On the effects of instruction in second language pragmatics. *System*, 22, 385-399.

Rose, K. R., & Ng. K. F. (2001). Inductive and deductive teaching of compliments and compliment responses. In K. R. Rose & G. Kasper (Eds.), *Pragmatics in language teaching* (pp.145-169). Cambridge, UK: Cambridge University Press.

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

- Schmidt, R. (1993). Consciousness, learning and interlanguage pragmatics. In G. Kasper, & S. Blum-Kulka, (Eds.) *Interlanguage pragmatics* (43-57). New York, NY: Oxford University Press.
- Takahashi, S. (2001). The role of input enhancement in developing pragmatic competence. In K. R. Rose & G. Kasper (Eds.), *Pragmatics in language teaching* (pp.171-199). Cambridge, UK: Cambridge University Press.
- Takahashi, S. (2010). The effect of pragmatic instruction on speech act performance. In A. Martínez-FLor & E. Uso-Juan (Eds.), *Speech act performance. Theoretical, empirical and methodological issues* (pp.126-142). Amsterdam: John Benjamins Publishing Company.
- Takimoto, M. (2006). The effects of explicit feedback on the development of pragmatic proficiency. *Language Teaching Research*, 10, 393-417.
- Takimoto, M. (2008). The effects of deductive and inductive Instruction on the development of language learners' pragmatic competence. *The Modern Language Journal*, 92, 369-386.
- Takimoto, M. (2009). The effects of input-based tasks on the development of learners' pragmatic proficiency. *Applied Linguistics*, 30, 1-25.
- Taguchi, N. (2007a). Development of speed and accuracy in pragmatic comprehension in English as a foreign language. *TESOL Quarterly*, 41, 313-338.
- Taguchi, N. (2007b). Task difficulty in oral speech act production. *Applied Linguistics*, 28, 113-135.
- Taguchi, N. (2008). Cognition, language contact, and the development of pragmatic comprehension in a study-abroad context. *Language Learning*, 58, 33-71.
- Tateyama, Y. (2009). Requesting in Japanese: The effect of instruction on JFL learners'

pragmatic competence. In N. Taguchi (ed.), *Pragmatic competence* (pp. 129-166). Berlin: Mouton de Gruyter.

Thomas, J. (1995). *Meaning and interaction: An introduction to pragmatics*. London: Longman.

VanPatten, B. (2004). Input processing in SLA. In B. VanPatten (Ed.), *Processing instruction: Theory, research, and commentary* (pp. 5-31). Mahwah, NJ: Lawrence Erlbaum Associates.

VanPatten, B. (2007). Input processing in adult second language acquisition. In B. VanPatten & J. Williams (Eds.), *Theories in Second language acquisition: An introduction* (pp.115-135). Mahwah, Jersey: Lawrence Erlbaum Associates, Publishers.

Vyatkina, N. & Belz, J. (2006). A learner corpus-driven intervention for the development of L2 pragmatic competence. In K. Bardovi-Harlig, C. Felix-Brasdefer & A. Omar (Eds.), *Pragmatics and Language Learning* (pp. 315-357). University of Hawaii Second Language Teaching and Curriculum Center.

Wishnoff, J. R. (2000). Hedging you bets: L2 learners' acquisition of pragmatic devices in academic writing and computer-mediated discourse. *Second Language Studies*, 19, 127-157.

Wong, W, (2004). The nature of processing instruction. In B. VanPatten (Ed.), *Processing instruction: Theory, research, and commentary* (pp. 33-63). Mahwah, NJ: Lawrence Erlbaum Associates.

Yoshimi, D. R. (2001). Explicit instruction and JFL learners' use of interactional discourse makers. In K. R. Rose & G. Kasper (Eds.), *Pragmatics in language teaching* (pp.223-244).

Pragmatics Instruction 48

Cambridge, UK: Cambridge University Press.

Appendix One

Sample Referential Activity

(A friend - low imposition situation)

Li Xiaochen and Wang Ning are good friends studying in the same university. This semester, they both take the English literature course. Li Xiaochen did not come to the class because he/she was ill. So Li Xiaochen wants to borrow and study Wang Ning's notes. They meet in the classroom.

wángníng
李晓晨：王宁！

xiǎochén nǐ lái shàngkè lā nǐ de bìng hǎole ma
王宁：晓晨，你来上课啦。你的病好了吗？

hǎo le wǒ zhè jǐ tiān méi shàng kè néng bù néng jiè wǒ kàn nǐ
李晓晨：好了。我这几天没上课。(a) 能不能借我看你
de bǐjì (b) 借我看一下你的笔记吧。

nǐ yào nǎ mén kè de bǐjì
王宁：你要哪门课的笔记？

bǎ yīngguó wénxué kè de bǐjì jiè wǒ kàn yíxià ba néng jiè wǒ
李晓晨：(a) 把英国文学课的笔记借我看一下吧。(b) 能借我
kàn yīngguó wénxuékè de bǐjì ma
看英国文学课的笔记吗？

hǎo xiàkè yǐhòu gěi nǐ
王宁：好，下课以后给你。

English translation of the dialogue (not provided to the learners)

Li Xiaochen: Wang Ning!

Wang Ning: Xiaochen. So you are coming to class. Have you recovered?

Li Xiaochen: Yes, I have. I did not come to class these days. **(a) Can you lend me your notes? (b) Lend me your notes.**

Wang Ning: For which course?

Li Xiaochen: **(a) Lend me the notes of the English literature course. (b) Can you lend me the notes of the English literature course?**

Wang Ning: Ok, I will give it to you after class.

Appendix Two

Sample affective activity

(A friend - low imposition situation)

There will be a talent show in the evening. The students are decorating the classroom. Li

Xiaochen wants to ask Wang Ning, who is standing nearby, to pass a pen which is close to

Wang Ning. Li Xiaochen and Wang Ning know each other very well.

李晓晨: wángníng nǐ kàndào zhuōzǐ shàng biān de bǐ le ma
王宁, 你看到桌子上边的笔了吗?

(a) bāng wǒ bǎ bǐ ná yíxià ba. (b) néngbunéng bāng wǒ ná
帮我**把**笔**拿**一下**吧**。 (b) 能**不**能**帮**我**拿**
yíxià bǐ
一下**笔**?

王宁: shì zhè zhī bǐ ma
是这支笔吗?

李晓晨: duì. (c) néng bāng wǒ ná yíxià ma. (d) bāng wǒ ná yíxià ba.
对。(c) 能**帮**我**拿**一下**吗**? (d) **帮**我**拿**一下**吧**。

王宁: xíng, gěi nǐ.
行, 给你。

Please indicate the appropriateness level of the two requests

(a) least appropriate 1-----2-----3-----4-----5-----6 most appropriate

(b) least appropriate 1-----2-----3-----4-----5-----6 most appropriate

(c) least appropriate 1-----2-----3-----4-----5-----6 most appropriate

(d) least appropriate 1-----2-----3-----4-----5-----6 most appropriate

English translation of the dialogue (not provided to the learners)

Li Xiaochen: Wang Ning, do you see that pen on the desk? **(a) Help me get the pen. (b) Can you help get the pen for me?**

Wang Ning: This pen?

Li Xiaochen: Yes. **(c) Can you help get it for me? (d) Help me get it.**

Wang Ning: Sure. Here you are.

Author Note

Shuai Li, Department of Modern Languages, Carnegie Mellon University.

This study was funded by the Department of Modern Languages, Carnegie Mellon University. An earlier version of this paper was presented at the 2010 AAAL Conference in Atlanta, GA. I would like to specifically thank Dr. Naoko Taguchi for her guidance. I also appreciate the comments from the SLA faculty and graduate students at Carnegie Mellon University. My deep appreciation goes also to the three anonymous reviewers and the journal editors for their insightful and constructive suggestions. Finally, I thank Dr. Marc Siskin for his assistance in creating the computer programs and Ms. Jing Wang for her help in data collection and analysis.

Correspondence concerning this article should be addressed to Shuai Li, BH 160, Department of Modern Languages, Carnegie Mellon University, Pittsburgh, PA 15213, USA.
E-mail: shuaili@andrew.cmu.edu.

Footnotes

1. In the EI+SI group, the participants first received meta-pragmatic instruction about the target pragmatic features (i.e., downgraders in English requests), followed by structured input activities (i.e., referential activities and affective activities). In the SI group, the participants completed the same structured input activities without meta-pragmatic instruction. In the PS group, the participants finished a series of consciousness-raising exercises (e.g., comparing request-making forms, analyzing request-making situations, etc.).

2. The *bǎ* structure is a language-specific construction in Chinese. The prototypical structure is “Subject (agent) + *bǎ* + object (patient) + verb-complement”. The structure conveys a strong sense of disposal, that is, it usually indicates that the object is disposed of, dealt with, or affected by the subject. Functionally, the *bǎ* structure can be used to form imperative sentences in Chinese. Note that this disposal *bǎ* (把) is different from the modal particle *ba* (吧). The modal particle *ba* (吧) can serve as a lexical downgrader to internally modify the illocutionary force of a request. It should also be noted that the *bǎ* structure does not equal direct strategy. In fact, it can occur in a conventionally indirect request. However, the Chinese native speaker data collected by Li (2007) showed that the disposal *bǎ* occurred only in three out of a total of 92 conventionally indirect requests in situations where Chinese university students were asked to make high-imposition request to their professors. Given the low frequency of co-occurrence, this study did not include instances where the disposal *bǎ* occurred in a conventionally indirect request.

3. As explained in the outcome measures section later, this study adopted five measures to examine the effects of practice (i.e., PLJT accuracy, PLJT response times, ODCCT accuracy,

ODCT planning times, and ODCT speech rates). The 33 participants' scores were converted into Z scores for each measure. Outliers were operationalized as those whose Z score(s) of a particular measure was larger than $|\pm 3.29|$ ($p < .001$). Three outliers were identified. One participant was found to be an outlier for two measures (i.e., PLJT accuracy and ODCT accuracy), and the other two participants were found to be outliers for one measure (i.e., PLJT accuracy, and ODCT planning time, respectively).

4. Three distracters were situations in which the speaker made suggestions to his/her friends, and the other three situations in which the speaker made low-imposition requests to his/her Professors.

5. The options were counterbalanced for the 12 "new" items but not for the eight "old" items. This was because the RT group only received practice on eight situations/dialogues and counterbalancing for the three options among the eight "old" situations was not possible.

6. The PLJT had two versions, PLJT-1 and PLJT-2, in order to reduce the potential effects of test practice. The two versions differed in three aspects: (1) the names of the role characters, (2) item order, and (3) the Chinese request utterance accompanying each situation. For example, suppose that a request situation was accompanied by a pragmatically appropriate and grammatically accurate request utterance in one version, the same situation would be accompanied in the other version by either (1) a pragmatically inappropriate and grammatically accurate request utterance or (2) a pragmatically appropriate and grammatically inaccurate request utterance.

7. One distractor item was a situation in which the speaker made suggestions to his/her friends, and the other was a situation in which the speaker made small requests to his/her

professors.

8. The two versions had comparable request-making situations, but the names of the role characters were different. The situations included in both ODCT versions were from the PLJT.

9. When the participants made a correct choice, they saw a pop-up window with the message “Right”; otherwise, they saw the message “Please re-consider”.

10. Again, two types of feedback, i.e., “Right” or “Please re-consider”, were provided following the participants’ choices.

Table 1

Target Form-Function-Context Mappings

Form	Function: Directness level	Context
1. Imperatives	Direct strategy	Equal power status (=P)
2. <i>Bǎ</i> structure		Small social distance (-D) Low ranking of imposition (-R)
3. <i>Néngbùnéng...?</i>	Conventionally	Speaker lower power status (-P)
4. <i>Néng...ma?</i>	indirect strategy	Small social distance (-D) High ranking of imposition (+R)

Table 2

Demographic Information of the Three Learner Groups

	Program	Gender	Non-native speakers of English
Regular training (<i>n</i> =10)	5 US program	4 males	3 Koreans
	3 Shanghai program	6 females	1 Thai
	2 Beijing program		
Intensive training (<i>n</i> =10)	4 US program	4 males	2 Koreans
	3 Shanghai program	6 females	1 Frenchman
	3 Beijing program		1 Japanese
Control (<i>n</i> =10)	4 US program	5 males	2 Koreans
	3 Shanghai program	5 females	1 Spaniard
	3 Beijing program		

Table 3

Means and Standard Deviations of PLJT Accuracy Scores

		Regular training (<i>n</i> =10)	Intensive training (<i>n</i> =10)	Control (<i>n</i> =10)
Pretest	Mean	14.70	15.80	16.20
	<i>SD</i>	1.88	1.81	1.61
Posttest	Mean	15.40	16.20	16.00
	<i>SD</i>	1.26	1.31	1.63
Delayed posttest	Mean	16.20	15.80	16.80
	<i>SD</i>	1.39	2.09	0.91

Note. The score range is 0 – 20.

Table 4

Means and Accuracy Rates of the Three PLJT Options

Option	Test		Regular Training (<i>n</i> =10)	Intensive Training (<i>n</i> =10)	Control (<i>n</i> =10)
A (<i>k</i> =8)	Pretest	<i>n</i>	6.90	7.30	7.30
		%	86.25%	91.25%	91.25%
	Posttest	<i>n</i>	7.10	7.50	7.30
%		88.75%	93.75%	91.25%	
	Delayed posttest	<i>n</i>	7.60	7.60	7.50
		%	95.00%	95.00%	93.75%
B (<i>k</i> =6)	Pretest	<i>n</i>	2.30	3.10	3.50
		%	38.33%	51.66%	58.33%
	Posttest	<i>n</i>	2.60	2.70	2.70
%		43.33%	45.00%	45.00%	
	Delayed posttest	<i>n</i>	2.70	3.20	3.60
		%	45.00%	53.33%	60.00%
C (<i>k</i> =6)	Pretest	<i>n</i>	5.50	5.40	5.40
		%	91.66%	90.00%	90.00%
	Posttest	<i>n</i>	5.70	6.00	6.00
%		95.00%	100.00%	100.00%	
	Delayed posttest	<i>n</i>	5.90	5.00	5.70
		%	98.33%	83.33%	95.00%

Note. A refers to utterances that are pragmatically appropriate and grammatically accurate, B refers to utterances that are pragmatically appropriate and grammatically inaccurate, and C refers to utterances that are pragmatically inappropriate and grammatically accurate.

Table 5

Tendency to Misjudge Option B Type of Request Utterances

	Pretest		Posttest		Delayed posttest	
	Option A	Option C	Option A	Option C	Option A	Option C
Regular training	81.33%	18.70%	86.33%	13.67%	90.50%	9.50%
Intensive training	79.17%	20.80%	89.17%	10.83%	95.00%	5.00%
Control	90.00%	10.00%	92.50%	7.50%	90.83%	9.17%

Table 6

Means and Standard Deviations of PLJT Response Times

		Regular training (<i>n</i> =10)	Intensive training (<i>n</i> =10)	Control (<i>n</i> =10)
Pretest	Mean	4.53	4.36	4.48
	<i>SD</i>	1.32	1.02	1.90
Posttest	Mean	4.19	3.17	3.69
	<i>SD</i>	1.91	1.16	1.47
Delayed posttest	Mean	3.77	2.51	3.60
	<i>SD</i>	1.11	0.86	1.75

Note. Response times refer to the average number of seconds taken to answer each item correctly.

Table 7

Means and Standard Deviations of ODCT Scores

		Regular training (n=10)	Intensive training (n=10)	Control (n=10)
Pretest	Mean	34.30	33.70	33.90
	<i>SD</i>	2.31	4.00	3.28
Posttest	Mean	37.30	38.10	35.50
	<i>SD</i>	1.25	1.10	2.17
Delayed posttest	Mean	37.30	38.00	35.00
	<i>SD</i>	1.82	1.33	2.53

Note. The score range is 0 – 40.

Table 8

Means and Standard Deviations of ODCT Planning Times

		Regular training (<i>n</i> =10)	Intensive training (<i>n</i> =10)	Control (<i>n</i> =10)
Pretest	Mean	1.81	1.44	2.60
	<i>SD</i>	0.88	0.79	2.49
Posttest	Mean	1.39	0.92	1.84
	<i>SD</i>	0.83	0.51	1.38
Delayed posttest	Mean	1.25	1.05	1.78
	<i>SD</i>	0.45	0.56	1.23

Note. Planning times refers to the average number of seconds taken to produce pragmatically appropriate request utterances.

Table 9

Means and Standard Deviations of ODCT Speech Rates

		Regular training (<i>n</i> =10)	Intensive training (<i>n</i> =10)	Control (<i>n</i> =10)
Pretest	Mean	107.11	144.96	112.23
	<i>SD</i>	31.02	33.73	28.17
Posttest	Mean	118.92	149.83	126.17
	<i>SD</i>	30.13	39.62	16.98
Delayed posttest	Mean	129.80	165.92	130.99
	<i>SD</i>	16.65	43.48	24.84

Note. Speech rates refer to the average number of Chinese syllables spoken per minute when producing pragmatically appropriate request utterances, excluding false starts, repetitions, partial repetitions, and items contained in repairs.